Logic, Algebra & Geometry

Achim Blumensath

ACHIM BLUMENSATH

Logic, Algebra and Geometry

BRNO 2024

Achim Blumensath
blumens@fi.muni.cz

This document was last updated 2024-04-09.
The latest version can be found at

www.fi.muni.cz/~blumens

Contents

Part A.

Set Theory

A1. Basic set theory

1. Sets and classes

In mathematics there are basically two ways to define the objects under consideration. On the one hand, one can explicitly construct them from already known objects. For instance, the rational numbers and the real numbers are usually introduced in this way. On the other hand, one can take the axiomatic approach, that is, one compiles a list of desired properties and one investigates any object meeting these requirements. Some well known examples are groups, fields, vector spaces, and topological spaces.

Since set theory is meant as foundation of mathematics there are no more basic objects available in terms of which we could define sets. Therefore, we will follow the axiomatic approach. We will present a list of six axioms and any object satisfying all of them will be called a *model of set theory.* Such a model consists of two parts: (1) a collection $\mathbb{S}$ of objects that we will call *sets,* and (2) some method which, given two sets a and b, tells us whether a is *an element of* b.

We will not care what exactly the objects in $\mathbb{S}$ are or how this method looks like. For example, one could imagine a model of set theory consisting of natural numbers. If we define that a natural number a is *an element of* the natural number b if and only if the a-th bit in the binary encoding of b is 1, then all but one of our axioms will be satisfied. It is conceivable that a similar but more involved definition might yield a model that satisfies all of them.

We will introduce our axioms in a stepwise fashion during the following sections. To help readers trying to look up a certain axiom we include a complete list below even if most of the needed definitions are still missing.

Axiom of Extensionality. Two sets a and b are equal if, and only if, we have $x \in a \Leftrightarrow x \in b$, for all sets x.

Axiom of Separation. If a is a set and φ a property then $\{ x \in a \mid \varphi \}$ is a set.

Axiom of Creation. For every set a there is a set S such that S is a stage and $a \in S$.

Axiom of Infinity. There exists a set that is a limit stage.

Axiom of Choice. For every set A there exists a well-order R over A.

Axiom of Replacement. If F is a function and $\operatorname{dom} F$ is a set then so is $\operatorname{rng} F$.

Asking whether these axioms are *true* does make as much sense as the question of whether the field axioms are true, or those of a vector space. Instead, what we are concerned with is their *consistency* and *completeness.* That is, there should *exist* at least one object satisfying these axioms and all such objects should *look alike.* Unfortunately, one can prove that there is no complete axiom system for set theory. Hence, we will have to deal with the fact that there are many different models of set theory and there is no way to choose one of them as the 'canonical one'. In particular, there is no such thing as 'the real model of set theory'.

More seriously, it is even impossible to prove that our axiom system is consistent. That is, it might be the case that there is *no* model of set theory and we have wasted our time studying a nonsensical theory.

The first problem is dealt with rather easily. It does not matter which of these models we are given since any theorem that we can derive from the axioms holds in every model. But the second problem is serious. All we can do is to restrict ourselves to as few axioms as possible and to hope that no one will ever be able to derive a contradiction. Of course, the weaker the axioms the more different models we might get and the fewer theorems we will be able to prove.

In the following we will assume that $\mathbb{S}$ is an arbitrary but fixed model of set theory. That is, $\mathbb{S}$ is a collection of objects that satisfies all the axioms we will introduce below. $\mathbb{S}$ will be called the *universe* and its elements are called *sets.* Note that $\mathbb{S}$ itself is not a set since we will prove below that no set is an element of itself. By convention, if below we say that some set *exists* then we mean that it is contained in $\mathbb{S}$. Similarly, we say that *all* sets have some property if all elements of $\mathbb{S}$ do so.

Intuitively, a set is a collection of objects called its *elements.* If a and b are sets, i.e., elements of $\mathbb{S}$, we write $a \in b$ if a is *an element of* b and we define

$$a \subseteq b \quad :\text{iff} \quad \text{every element } x \in a \text{ is also an element } x \in b\,.$$

If $a \subseteq b$, we call a a *subset* of b, and we say that a is *included* in b, or that b is a *superset* of a. We use the usual abbreviations such as $a \subset b$ for $a \subseteq b$ and $a \neq b$; $a \ni b$ for $b \in a$; and $a \notin b$ if $a \in b$ does not hold.

Since a set is a collection of objects it is natural to require that a set is uniquely determined by its elements. Our first axiom can therefore be regarded as the definition of a set.

Axiom of Extensionality. *Two sets a and b are equal if, and only if,*

$$x \in a \quad \text{iff} \quad x \in b\,, \quad \textit{for all sets } x\,.$$

Lemma 1.1. *Two sets a and b are equal if and only if $a \subseteq b$ and $b \subseteq a$.*

In order to define a set we have to say what its elements are. If the set is finite we can just enumerate them. Otherwise, we have to find some property φ such that an object x is an element of a if, and only if, it has the property φ.

Definition 1.2. (a) Let φ be a property. $\{\, x \mid \varphi \,\}$ denotes the set a such that, for all sets x, we have

$$x \in a \quad \text{iff} \quad x \text{ has property } \varphi\,.$$

If $\mathbb{S}$ does not contain such an object then the expression $\{ x \mid \varphi \}$ is undefined.

(b) Let $b_0, \dots, b_{n-1}$ be sets. We define

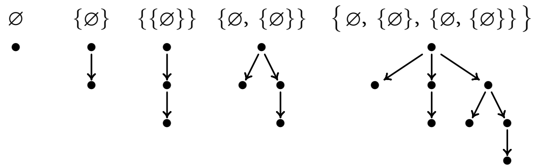
$$\{b_0, \dots, b_{n-1}\} := \{ x \mid x = b_i \text{ for some } i < n \} .$$

(c) The *empty set* is $\emptyset := \{ x \mid x \neq x \}$.

Note that, by the Axiom of Extensionality, if the set $\{ x \mid \varphi \}$ exists, it is unique.

In a model of set theory nothing but sets exists. But how can we have sets without some objects that serve as elements? The answer of course is to construct sets of other sets. First of all, there is one set that we can form even if we do not have any suitable elements: the empty set $\emptyset$. So we already have one object and we use it as element of other sets. In the next step we can form the set $\{\emptyset\}$, then we can form the sets $\{ \{\emptyset\} \}$ and $\{ \emptyset, \{\emptyset\} \}$ and so on.

Sometimes it is helpful to imagine such sets as trees. The empty set $\emptyset$ corresponds to a single vertex •. To a nonempty sets a we associate the tree consisting of a root to which we attach, for every element $b \in a$ the tree corresponding to b. For example, we have



To better understand this inductive construction of sets we introduce a toy version of set theory which has the advantage that it can be defined explicitly. It consists of all sets that one can construct from the empty set in finitely many steps.

Definition 1.3. We construct a sequence $\mathrm{HF}_0 \subseteq \mathrm{HF}_1 \subseteq \dots$ of sets as follows. We start with the empty set $\mathrm{HF}_0 := \emptyset$. When the set HF_n has already been defined, the next stage

$$\mathrm{HF}_{n+1} := \{\, x \mid x \subseteq \mathrm{HF}_n \,\}$$

consists of all sets that we can construct from elements of HF_n.

A set is called *hereditary finite* if it is an element of some HF_n. The set of all hereditary finite sets is

$$\mathrm{HF} := \{\, x \mid x \in \mathrm{HF}_n \text{ for some } n \,\}.$$

Note that we cannot prove at the moment that HF really is a set. Since the empty universe $\mathbb{S} = \varnothing$ trivially satisfies the Axiom of Extensionality, we even cannot show that the empty set exists without additional axioms. Let us assume for the moment that HF does exists. Its first stages are

$$\begin{aligned}
\mathrm{HF}_0 &= \varnothing \\
\mathrm{HF}_1 &= \{\varnothing\} \\
\mathrm{HF}_2 &= \{\varnothing, \{\varnothing\}\} \\
\mathrm{HF}_3 &= \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\} \\
&\dots
\end{aligned}$$

By induction on n, one can prove that $\mathrm{HF}_n \subseteq \mathrm{HF}_{n+1}$ and each set $a \in \mathrm{HF}_{n+1}$ is of the form $a = \{b_0, \dots, b_{k-1}\}$, for finitely many elements $b_0, \dots, b_{k-1} \in \mathrm{HF}_n$. Note that each stage HF_n is hereditary finite since $\mathrm{HF}_n \in \mathrm{HF}_{n+1} \subseteq \mathrm{HF}$, but their union HF is not because $\mathrm{HF} \notin \mathrm{HF}$.

Exercise 1.1. Prove the following statements by induction on n. (Although we have not defined the natural numbers yet, you may assume for this exercise that they are available and that their usual properties hold.)

(a) $\mathrm{HF}_n \subseteq \mathrm{HF}_{n+1}$.

(b) HF_n has finitely many elements.

(c) Every set $a \in \mathrm{HF}_{n+1}$ is of the form $a = \{b_0, \dots, b_{k-1}\}$, for finitely many elements $b_0, \dots, b_{k-1} \in \mathrm{HF}_n$.

HF can be regarded as an approximation to the class of all sets. In fact, all but one of the usual axioms of set theory hold for HF. The only exception is the Axiom of Infinity which states that there exists an infinite set.

We can encode natural numbers by special hereditary finite sets.

Definition 1.4. To each natural number n we associate the set

$$[n] := \{[0], \dots, [n-1]\}\,.$$

The set of all natural numbers is

$$\mathbb{N} := \{\, [n] \mid n \text{ a natural number} \,\}\,.$$

Note that $[n] \in \mathrm{HF}_{n+1}$ but $[n] \notin \mathrm{HF}_n$, and $\mathbb{N} \notin \mathrm{HF}$. This construction can be used to define the natural numbers in purely set theoretic terms. In the following by a *natural number* we will always mean a set of the form $[n]$.

It would be nice if there were a universe $\mathbb{S}$ that contains all sets of the form $\{\, x \mid \varphi \,\}$. Unfortunately, such a universe does not exists, that is, if we add the axiom that claims that $\{\, x \mid \varphi \,\}$ is defined for all φ, we obtain a theory that is inconsistent, i.e., it contradicts itself. In fact, we can even show that there are properties φ such that *no* model of set theory contains a set of the form $\{\, x \mid \varphi \,\}$. And we can do so without using a single axiom of set theory.

Theorem 1.5 (Zermelo-Russell Paradox). *$\{\, x \mid x \notin x \,\}$ is not a set.*

Proof. Suppose that the set $a := \{\, x \mid x \notin x \,\}$ exists. Let x be an arbitrary set. By definition, we have $x \in a$ if and only if $x \notin x$. In particular, for $x = a$, we obtain $a \in a$ iff $a \notin a$. A contradiction. □

To better understand what is going on, let us see what happens if we restrict ourselves to hereditary finite sets. The set $\{\, x \in \mathrm{HF} \mid x \notin x \,\}$ equals HF since no hereditary finite set contains itself. But $\mathrm{HF} \notin \mathrm{HF}$ is not hereditary finite. The same happens in real set theory. The condition

$x \notin x$ is satisfied by all sets and we have $\{ x \mid x \notin x \} = \mathbb{S}$, which is not a set.

In general, an expression of the form $\{ x \mid \varphi \}$ denotes a collection $X \subseteq \mathbb{S}$ that may or may not be a set, i.e., an element $X \in \mathbb{S}$. We will call objects of the form $\{ x \mid \varphi \}$ *classes.* Classes that are not sets will be called *proper classes.* If $X = \{ x \mid \varphi \}$ and $Y = \{ x \mid \psi \}$ are classes and a is a set, we write

$$\begin{aligned} a \in X \quad &:\text{iff} \quad a \text{ has property } \varphi\,, \\ X \subseteq Y \quad &:\text{iff} \quad \text{every set with property } \varphi \text{ also has property } \psi\,, \\ \text{and} \quad X = Y \quad &:\text{iff} \quad X \subseteq Y \text{ and } Y \subseteq X\,. \end{aligned}$$

If X is a proper class then we define $X \notin Y$, for every Y. Note that, if X and Y are sets then these definitions coincide with the ones above. Finally, we remark that every set a is a class since we can write a as $\{ x \mid x \in a \}$.

When defining classes we have to be a bit careful about what we call a property. Let us define a property to be a statement that is build up from basic propositions of the form $x \in y$ and $x = y$ by

- logical conjunctions like 'and', 'or', 'not', 'if-then';
- constructs of the form 'there exists a set x such that …' and 'for all sets x it holds that …'.

(Such statements will be defined in a more formal way in Chapter C1 where we will call them 'first-order formulae'.) Things we are not allowed to say include statements of the form 'There exists a property φ such that …' or 'For all classes X it holds that …'.

We have defined a class to be an object of the form $\{ x \mid \varphi \}$ where φ is a statement about sets. What happens if we allow statements about arbitrary classes? Note that, if φ is a property referring to a class $X = \{ x \mid \psi \}$ then we can transform φ into an equivalent statement only talking about sets by replacing all propositions $y \in X$, $X \in y$, $X = y$, etc. by their respective definitions.

Example. Let $X = \{ x \mid \varnothing \notin x \}$. We can write the class

$$\{ y \mid y \neq \varnothing \text{ and } y \subseteq X \}$$

in the form

$$\{ y \mid y \neq \varnothing \text{ and } \varnothing \notin x \text{ for all } x \in y \} .$$

The situation is analogous to the case of the complex numbers which are obtained from the real numbers by adding imaginary elements. We can translate any statement about complex numbers $x + iy$ into one about pairs $\langle x, y \rangle$ of real numbers. Consequently, it does not matter whether we allow classes in the definition of other classes.

Intuitively, the reason for a proper class such as $\mathbb{S}$ not being a set is that it is too 'large'. For instance, when considering HF we see that a set $a \subseteq \text{HF}$ is hereditary finite if, and only if, it has only finitely many elements. Hence, if we can show that a class $X = \{ x \mid \varphi \}$ is 'small', it should form a set. What do we mean by 'small'? Clearly, we would like every set to be small. Furthermore, it is natural to require that, if Y is small and $X \subseteq Y$ then X is also small. Therefore, we define a class X to be small if it is a subclass $X \subseteq a$ of some set a.

Definition 1.6. For a class A and a property φ we define

$$\{ x \in A \mid \varphi \} := \{ x \mid x \in A \text{ and } x \text{ has property } \varphi \} .$$

This definition ensures that every class of the form $X = \{ x \in a \mid \varphi \}$ where a is a set is small. Conversely, if $X = \{ x \mid \varphi \}$ is small then $X \subseteq a$, for some set a, and we have $X = \{ x \in a \mid \varphi \}$. Our second axiom states that every small class is a set.

Axiom of Separation. *If a is a set and φ a property then the class*

$$\{ x \in a \mid \varphi \}$$

is a set.

With this axiom we still cannot prove that there is any set. But if we have at least one set a, we can deduce, for instance, that also the empty set $\varnothing = \{ x \in a \mid x \neq x \}$ exists.

Definition 1.7. Let A and B be classes.

(a) The *intersection* of A is the class

$$\bigcap A := \{ x \mid x \in y \text{ for all } y \in A \} .$$

(b) The *intersection* of A and B is

$$A \cap B := \{ x \mid x \in A \text{ and } x \in B \} .$$

(c) The *difference* between A and B is

$$A \setminus B := \{ x \in A \mid x \notin B \} .$$

Lemma 1.8. *Let a be a set and B a class. Then $a \cap B$ and $a \setminus B$ are sets. If B contains at least one element then $\bigcap B$ is a set.*

Proof. The fact that $a \cap B = \{ x \in a \mid x \in B \}$ and $a \setminus B$ are sets follows immediately from the Axiom of Separation. If B contains at least one element $c \in B$ then we can write

$$\bigcap B = \{ x \in c \mid x \in y \text{ for all } y \in B \} . \qquad \square$$

Note that $\bigcap \varnothing = \mathbb{S}$ is not a set.

2. *Stages and histories*

The construction of HF above can be extended to one of the class $\mathbb{S}$ of all sets. We define $\mathbb{S}$ as the union of an increasing sequence of sets S_α, called the *stages* of $\mathbb{S}$. Again, we start with the empty set $S_0 := \varnothing$. If S_α is defined then the next stage $S_{\alpha+1}$ contains all subsets of S_α. But this time, we do not stop when we have defined S_α for all natural numbers α. Instead, every time we have defined an infinite sequence of stages we continue by taking their union to form the next stage. So our sequence starts with

$$S_0 = \mathrm{HF}_0\,, \quad S_1 = \mathrm{HF}_1\,, \quad S_2 = \mathrm{HF}_2\,, \quad \dots$$

The next stage after all the finite ones is $S_\omega := \mathrm{HF}$ and we continue with

$$S_{\omega+1} = \{\, x \mid x \subseteq \mathrm{HF} \,\}\,, \quad S_{\omega+2} = \{\, x \mid x \subseteq S_{\omega+1} \,\}\,, \quad \dots$$

After we have defined $S_{\omega+n}$ for all natural numbers n we again take the union

$$S_{\omega+\omega} = \{\, x \mid x \in S_{\omega+n} \text{ for some } n \,\}\,,$$

and so on.

Unfortunately, making this construction precise turns out to be quite technical since we cannot define the numbers α yet that we need to index the sequence S_α. This has to wait until Section A3.2. Instead, we start by giving a condition for some set S to be a *stage*, i.e., one of the S_α. If we order all such sets by inclusion then we obtain the desired sequence

$$S_0 \subseteq S_1 \subseteq \cdots \subseteq S_\omega \subseteq S_{\omega+1} \subseteq \cdots\,,$$

without the need to refer to its indices.

First, we isolate some characteristic properties of the sets HF_n which we would like that our stages S_α share. Note that, at the moment, we cannot prove that any of the sets mentioned below actually exists.

Definition 2.1. Let A be a class.

(a) We call A *transitive* if $x \in y \in A$ implies $x \in A$.

(b) We call A *hereditary* if $x \subseteq y \in A$ implies $x \in A$.

(c) The *accumulation* of A is the class

$$\mathrm{acc}(A) := \{\, x \mid \text{there is some } y \in A \text{ such that } x \in y \text{ or } x \subseteq y \,\}\,.$$

Note that each stage HF_n of HF is hereditary and transitive.

Exercise 2.1. By induction on n, show that the set $[n]$ is transitive. Give an example of a number n such that $[n]$ is not hereditary.

The next lemmas follow immediately from the definitions.

Lemma 2.2. *Let A be a class, and b, c sets. The following statements are equivalent:*

(a) $c \in b \in A$ *implies* $c \in A$, *that is,* A *is transitive.*

(b) $b \in A$ *implies* $b \subseteq A$.

(c) $b \in A$ *implies* $b \cap A = b$.

Lemma 2.3. *Let A and B be classes.*

(a) $A \subseteq \mathrm{acc}(A)$

(b) *If B is hereditary and transitive and if $A \subseteq B$, then* $\mathrm{acc}(A) \subseteq B$.

(c) *A is hereditary and transitive if, and only if,* $\mathrm{acc}(A) = A$.

Lemma 2.4. *If A and B are transitive classes then so is $A \cap B$.*

Exercise 2.2. Prove Lemmas 2.2, 2.3, and 2.4.

Definition 2.5. Let A be a class.

(a) A *minimal element* of A is an element $b \in A$ such that $b \cap A = \varnothing$, that is, there is no element $c \in A$ with $c \in b$.

(b) A set a is *founded* if every set $b \ni a$ has a minimal element.

(c) The *founded part* of A is the set

$$\mathrm{fnd}(A) := \{\, x \in A \mid x \text{ is founded} \,\}\,.$$

Example. The empty set $\varnothing$ and the set $\{\varnothing\}$ are founded. To see that $\{\varnothing\}$ is founded, consider a set $b \ni \{\varnothing\}$. If $\{\varnothing\}$ is not a minimal element of b, then $b \cap \{\varnothing\} \neq \varnothing$. Hence, $\varnothing \in b$ is a minimal element of b.

Exercise 2.3. Prove that every hereditary finite set is founded.

We will introduce an axiom below which implies that every class has a minimal element. Hence, every set is founded and we have $\mathrm{fnd}(A) = A$, for all classes A. Although the notions of a founded set and the founded part of a set will turn out to be trivial, we still need them to define stages and to formulate the axiom.

Lemma 2.6. *If B is a hereditary class and $a \in B$ then* $\mathrm{fnd}(a) \in \mathrm{fnd}(B)$.

Proof. For a contradiction suppose that $\mathrm{fnd}(a) \notin \mathrm{fnd}(B)$. Since B is hereditary and $\mathrm{fnd}(a) \subseteq a \in B$, we have $\mathrm{fnd}(a) \in B$. Consequently, $\mathrm{fnd}(a) \notin \mathrm{fnd}(B)$ implies that there is some set $x \ni \mathrm{fnd}(a)$ without minimal element. In particular, $\mathrm{fnd}(a)$ is not a minimal element of x, that is, there exists some set $y \in x \cap \mathrm{fnd}(a)$. But $y \in \mathrm{fnd}(a)$ implies that y is founded. Therefore, from $y \in x$ it follows that x has a minimal element. A contradiction. □

In the language of Section A3.1 the next theorem states that the membership relation $\in$ is well-founded on every class of transitive, hereditary sets.

Theorem 2.7. *Let A be a nonempty class. If every element $x \in A$ is hereditary and transitive, then A has a minimal element.*

Proof. Choose an arbitrary element $c \in A$ and set

$$b := \{ \mathrm{fnd}(x) \mid x \in c \cap A \} .$$

If $b = \varnothing$ then $c \cap A = \varnothing$ and c is a minimal element of A. Therefore, we may assume that $b \neq \varnothing$. Since $c \in A$ is hereditary, it follows from Lemma 2.6 that $b \subseteq \mathrm{fnd}(c)$. Fix some $x \in b \subseteq \mathrm{fnd}(c)$. Then x is founded and $x \in b$ implies that b has a minimal element y. By definition of b, we have $y = \mathrm{fnd}(z)$, for some $z \in c \cap A$.

We claim that z is a minimal element of A. Suppose otherwise. Then there exists some element $u \in z \cap A$. Since c is transitive we have $u \in c$. Hence, $u \in c \cap A$ implies $\mathrm{fnd}(u) \in b$. On the other hand, since $z \in A$ is hereditary it follows from Lemma 2.6 that $\mathrm{fnd}(u) \in \mathrm{fnd}(z)$. Hence,

$\mathrm{fnd}(u) \in \mathrm{fnd}(z) \cap b \neq \varnothing$ and $y = \mathrm{fnd}(z)$ is not a minimal element of b. A contradiction. □

We would like to define that a set S is a stage if it is hereditary and transitive. Unfortunately, this definition is too weak to show that the stages can be arranged in an increasing sequence $S_0 \subseteq S_1 \subseteq \cdots \subseteq S_\alpha \subseteq \cdots$. Therefore, our definition will be slightly more involved. To each stage S_α we will associate its *history*

$$H(S_\alpha) = \{ S_\beta \mid \beta < \alpha \},$$

and we will call a set S a *stage* if $S = \mathrm{acc}(H(S))$. Note that, for HF_n, we have

$$H(\mathrm{HF}_n) = \{\mathrm{HF}_0, \dots, \mathrm{HF}_{n-1}\} \quad \text{and} \quad \mathrm{HF}_n = \mathrm{acc}\big(H(\mathrm{HF}_n)\big).$$

Of course, to avoid a vicious cycle we have to define a history without mentioning stages.

Definition 2.8. (a) A class H is a *history* if every element $a \in H$ is hereditary, transitive, and satisfies

$$a = \mathrm{acc}(H \cap a).$$

(b) If H is a history, we call the class $S := \mathrm{acc}(H)$ the *stage* with history H.

Let us show that these definitions have the desired effect.

Lemma 2.9. *Let S be a stage with history H.*

(a) $H \subseteq S$.

(b) *Every set $a \in H$ is a stage with history $H \cap a$.*

(c) *S is hereditary and transitive.*

(d) $S = \{ x \mid x \subseteq s$ *for some stage* $s \in S \}$.

(e) $H(S) := \{ s \in S \mid s$ *is a stage* $\}$ *is a history of* S.

Proof. (a) $a \subseteq a \in H$ implies $a \in \mathrm{acc}(H) = S$.

(b) By definition of a history, we have $a = \mathrm{acc}(H \cap a)$. Hence, if we can show that $H \cap a$ is a history then its stage is a. Clearly, every element of $H \cap a \subseteq H$ is hereditary and transitive. Let $b \in H \cap a$. Then $b \subseteq \mathrm{acc}(H \cap a) = a$. It follows that $H \cap b = (H \cap a) \cap b$. Furthermore, since H is a history we have

$$b = \mathrm{acc}(H \cap b) = \mathrm{acc}((H \cap a) \cap b)\,,$$

which shows that $H \cap a$ is a history.

(c) Let $b \in S$. The class

$$a := \{\, s \in H \mid b \in s \text{ or } b \subseteq s \,\}$$

is nonempty because $b \in S = \mathrm{acc}(H)$. By Theorem 2.7, it has a minimal element $s \in a$.

If $b \in s = \mathrm{acc}(H \cap s)$, there is some set $z \in H \cap s$ such that $b \in z$ or $b \subseteq z$. It follows that $z \in a$. But $z \in s \cap a$ implies that s is not a minimal element of a. Contradiction.

Therefore, $b \notin s$ which implies, by definition of a, that $b \subseteq s$. For transitivity, note that $x \in b$ implies

$$x \in b \subseteq s = \mathrm{acc}(H \cap s) \subseteq \mathrm{acc}(H) = S\,.$$

For hereditarity, let $x \subseteq b$. Then $x \subseteq b \subseteq s \in H$, which implies $x \in \mathrm{acc}(H) = S$.

(d) By (c) we know that $x \subseteq s \in S$ implies $x \in S$. For the other direction, suppose that $x \in S = \mathrm{acc}(H)$. There is some set $s \in H$ such that $x \in s$ or $x \subseteq s$. By (a), (b), and (c) it follows that $s \in S$, s is a stage, and s is hereditary and transitive. By transitivity, if $x \in s$ then $x \subseteq s$. Consequently, we have $x \subseteq s \in S$ in both cases and the claim follows.

(e) By (d), we have $S = \mathrm{acc}(H(S))$. It remains to show that $H(S)$ is a history. By (c), every element $s \in H(S)$ is hereditary and transitive. Furthermore, since S is transitive we have $s \subseteq S$ and it follows that

$$H(S) \cap s = \{\, x \in s \mid x \text{ is a stage } \}\,.$$

Since s is a stage we know by (d) that $s = \mathrm{acc}(H(S) \cap s)$. □

Note that, by (a) and (b) above, we have $H \subseteq H(S)$, for all histories H of S. In fact, $H(S)$ is the only history of S but we need some further results before we can prove this.

Exercise 2.4. Prove, by induction on n, that $\{\mathrm{HF}_0, \dots, \mathrm{HF}_{n-1}\}$ is a history with stage HF_n.

Exercise 2.5. Construct a hereditary transitivity set a that is not a stage. *Hint.* It is sufficient to consider sets $\mathrm{HF}_n \subset a \subset \mathrm{HF}_{n+1}$, for a small n.

After we have seen how to define stages we now prove that they form a strictly increasing sequence $S_0 \subseteq S_1 \subseteq \dots$. Together with Theorem 2.7 it follows that the class of all stages is well-ordered by the membership relation $\in$ (see Section A3.1).

Theorem 2.10. *If S and T are stages that are sets then we have*

$$S \in T \quad \textit{or} \quad S = T \quad \textit{or} \quad T \in S\,.$$

Proof. Suppose that there are stages S and T such that

$$(*) \quad S \notin T\,, \quad S \neq T\,, \quad \text{and} \quad T \notin S\,.$$

Define

$$A := \{\, s \mid s \text{ is a stage and there is some stage } t \text{ such that } s \text{ and } t \text{ satisfy } (*) \,\}\,.$$

By Theorem 2.7, the class A has a minimal element S_0. Define

$$B := \{\, t \mid t \text{ is a stage such that } S_0 \text{ and } t \text{ satisfy } (*) \,\}\,.$$

Again there is a minimal element $T_0 \in B$.

If we can show that $H(S_0) = H(T_0)$, it follows that

$$S_0 = \mathrm{acc}(H(S_0)) = \mathrm{acc}(H(T_0)) = T_0$$

in contradiction to our choice of S_0 and T_0.

Let $s \in S_0$ be a stage. Then $s \neq T_0$ since $T_0 \notin S_0$. Furthermore, we have $T_0 \notin s$ since, otherwise, transitivity of S_0 would imply that $T_0 \in S_0$. By minimality of S_0 it follows that s and T_0 do not satisfy $(*)$. Therefore, we have $s \in T_0$.

We have shown that $H(S_0) \subseteq H(T_0)$. A symmetric argument shows that $H(T_0) \subseteq H(S_0)$. Hence, we have $H(S_0) = H(T_0)$ as desired. □

Lemma 2.11. *Let S and T be stages that are sets.*

(a) $S \notin S$

(b) $S \subseteq T$ *if and only if* $S \in T$ *or* $S = T$.

(c) $S \subseteq T$ *or* $T \subseteq S$.

(d) $S \subset T$ *if, and only if,* $S \in T$.

Proof. (a) Suppose otherwise. Let X be the class of all stages s such that $s \in s$. By Theorem 2.7, X has a minimal element s, that is, an element such that $s \cap X = \varnothing$. But $s \in s \cap X$. Contradiction.

(b) If $S = T$ then $S \subseteq T$, and if $S \in T$ then $S \subseteq T$, by transitivity of T. Conversely, if neither $S = T$ nor $S \in T$ then Theorem 2.10 implies that $T \in S$. If $S \subseteq T$ then $T \in S \subseteq T$ would contradict (a).

(c) If $S \nsubseteq T$ then (b) implies that $S \notin T$ and $S \neq T$. By Theorem 2.10, it follows that $T \in S$ which, again by (b), implies $T \subseteq S$.

(d) We have $S \subset T$ iff $S \subseteq T$ and $S \neq T$. By (a) and (b), the latter is equivalent to $S \in T$. □

3. *The cumulative hierarchy*

In the previous section we have seen that we can arrange all stages in an increasing sequence

$$S_0 \subset S_1 \subset \cdots \subset S_\alpha \subset \cdots,$$

which we will call the *cumulative hierarchy.* If $S \in T$ are stages then we will say that S is *earlier* than T, or that T is *later* than S.

From the axioms we have available we cannot prove that there actually are any stages. We introduce a new axiom which ensures that enough stages are available.

Axiom of Creation. *For every set a there is a set $S \ni a$ which is a stage.*

In particular, this axiom implies that

- for every stage S that is a set, there exists a later stage $T \ni S$ that is also a set.
- the universe $\mathbb{S}$ is the union of all stages.

Of course, even with this new axiom we might still have $\mathbb{S} = \varnothing$. But if at least one set exists, we can now prove that HF $\subseteq \mathbb{S}$. In particular, $\mathbb{S} =$ HF satisfies all axioms we have introduced so far.

Exercise 3.1. Prove that $\mathbb{S}$ is a stage with history

$$H(\mathbb{S}) = \{ S \mid S \text{ is a stage} \} .$$

Definition 3.1. (a) We say that a stage T is the *successor* of the stage S if $S \in T$ and there exists no stage T' such that $S \in T' \in T$. A nonempty stage is a *limit* if it is not the successor of some other stage.

(b) Let A be a class. We denote by $S(A)$ the earliest stage such that $A \subseteq S(A)$.

Note that $S(A)$ is well-defined by Theorem 2.7. We have $S(s) = s$, for every stage s, in particular, $S(\varnothing) = \varnothing$. The stages $\mathbb{S}$ and HF are limits and HF_{n+1} is the successor of the stage HF_n.

Lemma 3.2. *$a \in b$ implies $S(a) \in S(b)$.*

Proof. Since $a \in b \subseteq S(b) = \mathrm{acc}(H(S(b)))$ it follows that there is some stage $s \in S(b)$ such that $a \in s$ or $a \subseteq s$. In particular, $S(a)$ is not later than s which implies that $S(a) \subseteq s \in S(b)$. As $S(b)$ is hereditary we therefore have $S(a) \in S(b)$. □

Lemma 3.3. *$\mathbb{S}$ is the only stage that is a proper class.*

Proof. Let S be a stage. If $S \neq \mathbb{S}$, there is some set $a \in \mathbb{S} \setminus S$. Hence, $S(a) \notin S$ which implies that

$$T \notin H(S), \quad \text{for all stages } T \supseteq S(a).$$

By Lemma 2.9 (e) and Theorem 2.10, we have

$$H(S) \subseteq \{ T \mid T \text{ is a stage with } T \in S(a) \} = H(S(a)).$$

In particular, $H(S)$ is a set, which implies that so is $S = \operatorname{acc}(H(S))$. □

Lemma 3.4. *Let A be a class. The following statements are equivalent:*

(1) *A is a proper class.*

(2) *$S(A)$ is a proper class.*

(3) *$S(A) = \mathbb{S}$.*

Proof. (3) ⇒ (1) By the Axiom of Creation, if A is a set then so is $S(A)$.

(1) ⇒ (2) If $S(A)$ is a set then $A \subseteq S(A)$ implies that

$$A = \{ x \in S(A) \mid x \in A \}$$

is also a set.

(2) ⇒ (3) follows by Lemma 3.3. □

With the Axiom of Creation we are finally able to prove most 'obvious' properties of sets such that no set is an element of itself or that the union of sets is a set.

Lemma 3.5. *If a is a set then $a \notin a$.*

Proof. Suppose that there exists some set such that $a \in a$. Then $a \in a \subseteq S(a)$ and, by Lemma 2.9 (d), there is some stage $s \in S(a)$ with $a \subseteq s$. This contradicts the minimality of $S(a)$. □

Theorem 3.6. *Every nonempty class A has a minimal element.*

Proof. By Theorem 2.7, we can choose some element $b \in A$ such that $S(b)$ is minimal. We claim that b is a minimal element of A. Suppose otherwise. Then there exists some element $x \in A \cap b$. Since $x \in b \subseteq S(b)$, Lemma 2.9 (d) implies that there is some stage $s \in S(b)$ such that $x \subseteq s$. Hence, x is an element of A with $S(x) \in S(b)$ in contradiction to the choice of b. □

We will see in Section A3.1 that Theorem 3.6 implies that there are no infinite descending sequences $a_0 \ni a_1 \ni \dots$ of sets. (If such a sequence exists then the set $\{a_0, a_1, \dots\}$ has no minimal element.)

Example. By induction on n, it trivially follows that, if $a_0 \ni \dots \ni a_{k-1}$ is a sequence of sets starting with $a_0 \in \mathrm{HF}_n$, then $k < n$. What happens if $a_0 = \mathrm{HF}$? Then $a_1 \in \mathrm{HF}_n$, for some n, and the sequence is of length $k \leq n$. But note that, for every n, we can find a sequence of length n starting with $a_0 = \mathrm{HF}$. So there is no one bound that works for all sequences.

Definition 3.7. Let A and B be classes.

(a) The *union* of A is the class

$$\bigcup A := \{\, x \mid x \in b \text{ for some } b \in A \,\}\,.$$

(b) The *union* of A and B is

$$A \cup B := \{\, x \mid x \in A \text{ or } x \in B \,\}\,.$$

(c) The *power set* of A is the class

$$\wp(A) := \{\, x \mid x \subseteq A \,\}\,.$$

Remark. Note that, by definition, a class contains only sets. In particular, the power set $\wp(A)$ of a proper class contains only the *subsets* of A, not all subclasses. For instance, we have $\wp(\mathbb{S}) = \mathbb{S}$.

Lemma 3.8. *If a and b are sets then so are $\bigcup a$, $a \cup b$, $\{a\}$, and $\wp(a)$.*

Proof. Let S_0 and S_1 be stages such that $a \in S_0$ and $b \in S_1$. We know that $S_0 \subseteq S_1$ or $S_1 \subseteq S_0$. By choosing either S_0 or S_1 we can find a stage S such that $S_0 \subseteq S$ and $S_1 \subseteq S$. By transitivity of S it follows that

$$\bigcup a = \{\, x \in S \mid x \in b \text{ for some } b \in a \,\},$$
$$a \cup b = \{\, x \in S \mid x \in a \text{ or } x \in b \,\},$$
$$\{a\} = \{\, x \in S \mid x = a \,\},$$
and $$\mathcal{P}(a) = \{\, b \in S \mid b \subseteq a \,\}.$$ □

Corollary 3.9. *If $a_0, \dots, a_{n-1}$ are sets then so is*

$$\{a_0, \dots, a_{n-1}\} = \{a_0\} \cup \dots \cup \{a_{n-1}\}.$$

In particular, every finite class is a set.

The next definition provides a useful tool which sometimes allows us to replace a proper class A by a set a. Instead of taking every element $x \in A$ we only consider those such that $S(x)$ is minimal.

Definition 3.10. The *cut* of a class A is the set

$$\operatorname{cut} A := \{\, x \in A \mid S(x) \subseteq S(y) \text{ for all } y \in A \,\}.$$

Exercise 3.2. What are $\operatorname{cut} \mathbb{S}$ and $\operatorname{cut}\{\, x \mid a \in x \,\}$?

Lemma 3.11. *Every class of the form* $\operatorname{cut} A$ *is a set.*

Proof. If $A = \varnothing$ then $\operatorname{cut} A = \varnothing$. Otherwise, choose an arbitrary set $a \in A$. Then $\operatorname{cut} A \subseteq S(a)$ which implies that $\operatorname{cut} A$ is a set. □

The following lemmas clarify the structure of the cumulative hierarchy.

Lemma 3.12. *The successor of a stage S is $\mathcal{P}(S)$.*

Proof. By Theorem 2.7, there exists a minimal stage T with $S \in T$. We have to prove that $T = \mathcal{P}(S)$. $a \subseteq S \in T$ implies $a \in T$ since T is hereditary. Hence, $\mathcal{P}(S) \subseteq T$.

Conversely, if $s \in T$ is a stage then $S \notin s$ because T is the successor of S. By Theorem 2.10, it follows that $s \in S$ or $s = S$. This implies $s \subseteq S$.

We have shown that $s \in T$ iff $s \subseteq S$, for all stages s. It follows by Lemma 2.9 (d) that

$$\begin{aligned} T &= \{\, x \mid x \subseteq s \text{ for some stage } s \in T \,\} \\ &= \{\, x \mid x \subseteq s \text{ for some stage } s \subseteq S \,\} = \{\, x \mid x \subseteq S \,\} = \wp(S)\,. \end{aligned}$$ □

Lemma 3.13. *Let S be a nonempty stage. The following statements are equivalent:*

(1) *S is a limit stage.*

(2) $S = \bigcup H(S)$.

(3) *For every set $a \in S$, there exists some stage $s \in S$ with $a \in s$.*

(4) *If $a \in S$ then $\wp(a) \in S$.*

(5) *If $a \in S$ then $\{a\} \in S$.*

(6) *If $a \subseteq S$ then* cut $a \in S$.

Proof. (2) ⇒ (1) Suppose that S is the successor of a stage T. Then we have

$$H(S) = \{T\} \cup H(T)\,.$$

Since $s \subseteq T$, for all $s \in H(T)$, it follows that

$$\bigcup H(S) = T \neq S.$$

(1) ⇒ (2) Suppose that S is a limit stage. By Lemma 2.9 (d), we have

$$\begin{aligned} S &= \bigcup \{\, \wp(s) \mid s \in H(S) \,\} \\ &= \bigcup \{\, t \mid t \text{ is the successor of some stage } s \in H(S) \,\} \\ &= \bigcup \{\, t \mid t \in H(S) \,\} \\ &= \bigcup H(S)\,. \end{aligned}$$

(1) ⇒ (3) Suppose that S is a limit and let $a \in S$. By Lemma 2.9 (d), there is some stage $s \in S$ with $a \subseteq s$. Hence, $a \in \wp(s)$. Since $T := \wp(s)$ is the successor of s we have $T \in S$.

(3) ⇒ (4) For each $a \in S$, there is some stage $s \in S$ with $a \in s$. Since s is transitive it follows that $x \subseteq a$ implies $x \in s$. Hence, $\wp(a) \subseteq s$. By transitivity of S, we obtain $\wp(a) \in S$.

(4) ⇒ (5) If $a \in S$ then $\{a\} \subseteq \wp(a) \in S$. Since S is hereditary, it follows that $\{a\} \in S$.

(5) ⇒ (1) If S is no limit, there is some stage $T \in S$ such that $S = \wp(T)$. By assumption, $\{T\} \in S = \wp(T)$. Hence, $\{T\} \subseteq T$ which implies that $T \in T$. A contradiction.

(3) ⇒ (6) Let $b := \operatorname{cut} a$. If $a = \varnothing$ then $b = \varnothing$ and we are done. If there is some element $x \in a$ then, by assumption, we can find a stage $s \in S$ with $x \in s$. By definition, $b \subseteq s$, and it follows that $b \in S$.

(6) ⇒ (5) Let $a \in S$ and set $b := \{ x \in S \mid a \subseteq x \}$. Clearly, $b \subseteq S$. By assumption, we therefore have $c := \operatorname{cut} b \in S$. Hence, $\{a\} \subseteq c$ implies $\{a\} \in S$. □

So far, we still might have $\mathbb{S} = \varnothing$ or $\mathbb{S} = \mathrm{HF}$. To exclude these cases we introduce a new axiom which states that $\mathrm{HF} \in \mathbb{S}$.

Axiom of Infinity. *There exists a set that is a limit stage.*

We call the theory consisting of the four axioms

- Axiom of Extensionality
- Axiom of Creation
- Axiom of Separation
- Axiom of Infinity

basic set theory. Every model of this theory consist of a hierarchy of stages

$$S_0 \subset S_1 \subset \cdots \subset S_\omega \subset S_{\omega+1} \subset \dots$$

where $S_n = \mathrm{HF}_n$, for finite n. The differences between two such models can be classified according to two axes: the length of the hierarchy and the size of each stage.

Let $\mathbb{S}$ and $\mathbb{S}'$ be two models with stages $(S_\alpha)_{\alpha<\kappa}$ and $(S'_\alpha)_{\alpha<\lambda}$, respectively. We know that their lengths κ and λ are at least what we will call $\omega + \omega$ in Section A3.2. But our current axioms do not tell us whether the process of creation stops there or whether we again take the union of all stages and continue taking power sets until we reach $\omega + \omega + \omega$. At this point we again have to decide whether to stop or to continue, and so on.

The second possible difference stems from the fact that the power-set operation is ambiguous. We know that $S_n = \mathrm{HF}_n = S'_n$, for all finite n. But we might have $S_\alpha \neq S'_\alpha$, for infinite α. The reason is that there is no way to express that *all* subsets of S_α are contained in $S_{\alpha+1}$. We have the Axiom of Separation which states that all subsets exist that we can explicitly define. But there are much more possible subsets than there are definitions.

A2. *Relations*

1. *Relations and functions*

With basic set theory available we can define most of the concepts used in mathematics. The simplest one is the notion of an ordered pair. The characteristic property of such pairs is that $\langle a, b\rangle = \langle c, d\rangle$ implies $a = c$ and $b = d$.

Definition 1.1. (a) Let a and b be sets. The *ordered pair* $\langle a, b\rangle$ is the set

$$\langle a, b\rangle := \big\{\{a\}, \{a, b\}\big\}.$$

(b) Let A and B be classes. The *cartesian product* of A and B is the class

$$A \times B := \{\, c \mid c = \langle a, b\rangle \text{ for some } a \in A \text{ and } b \in B \,\}.$$

Let us show that ordered pairs have the desired property.

Lemma 1.2. *If* $\{a, b\} = \{a, c\}$ *then* $b = c$.

Proof. We have $b \in \{a, b\} = \{a, c\}$. Hence, $b = a$ or $b = c$. In the latter case we are done. Otherwise, we have $c \in \{a, c\} = \{a, b\} = \{b\}$ which implies that $c = b$. ☐

Lemma 1.3. *If* $\langle a, b\rangle = \langle c, d\rangle$ *then* $a = c$ *and* $b = d$.

Proof. Suppose that $\langle a, b\rangle = \langle c, d\rangle$.

$$\{a\} \in \big\{\{a\}, \{a, b\}\big\} = \big\{\{c\}, \{c, d\}\big\}$$

implies $\{a\} = \{c\}$ or $\{a\} = \{c, d\}$. In the latter case, we have $a = c = d$. In both cases, we therefore have $\{a\} = \{c\}$. By the preceding lemma, it follows that $\{a, b\} = \{c, d\}$ and, applying the lemma again, we obtain $b = d$. □

Remark. The above definition of an ordered pair $\langle a, b\rangle$ does only work for sets. Nevertheless we will use also pairs $\langle A, B\rangle$ where A or B are proper classes. There are several ways to make such an expression well-defined. A simple one is to define

$$\langle A, B\rangle := (\{[0]\} \times A) \cup (\{[1]\} \times B) \quad (= A \uplus B)$$

whenever at least one of A and B is a proper class. (The operation $\uplus$ will be defined more generally in the next section.) It is easy to check that with this definition the term $\langle A, B\rangle$ has the properties of an ordered pair, that is, $A \uplus B = C \uplus D$ implies $A = C$ and $B = D$.

Definition 1.4. (a) For sets $a_0, \dots, a_n$ we define inductively

$$\langle\rangle := \varnothing\,, \qquad \langle a_0\rangle := a_0\,,$$

and $\quad \langle a_0, \dots, a_n\rangle := \langle\langle a_0, \dots, a_{n-1}\rangle, a_n\rangle\,.$

We call $\langle a_0, \dots, a_{n-1}\rangle$ a *tuple of length n*. $\langle\rangle$ is the *empty tuple*.

(b) For a class A, we define its *n-th power* by

$$A^0 := \{\langle\rangle\}\,, \quad A^1 := A\,, \quad \text{and} \quad A^{n+1} := A^n \times A\,, \quad \text{for } n > 1\,.$$

Definition 1.5. A *relation,* or a *predicate,* of *arity n* is a subclass $R \subseteq \mathbb{S}^n$. If $R \subseteq A^n$, for some class A, we say that R is *over A*.

Note that $\varnothing$ and $\{\langle\rangle\}$ are the only relations of arity 0. In logic they are usually interpreted as *false* and *true.* A relation of arity 1 over A is just a subclass $R \subseteq A$.

Definition 1.6. Let R be a binary relation. The *domain* of R is the class

$$\operatorname{dom} R := \{\, a \mid \langle a, b\rangle \in R \text{ for some } b \,\}\,,$$

and its *range* is

$$\operatorname{rng} R := \{\, b \mid \langle a, b\rangle \in R \text{ for some } a \,\}\,.$$

The *field* of R is $\operatorname{dom} R \cup \operatorname{rng} R$.

In particular, $\operatorname{dom} R$ and $\operatorname{rng} R$ are the smallest classes such that

$$R \subseteq \operatorname{dom} R \times \operatorname{rng} R\,.$$

Definition 1.7. (a) A binary relation R is called *functional* if, for every $a \in \operatorname{dom} R$, there exists exactly one set b such that $\langle a, b\rangle \in R$. We denote this unique element b by $R(a)$. Hence, we can write R as

$$R = \{\, \langle a, R(a)\rangle \mid a \in \operatorname{dom} R \,\}\,.$$

A functional relation $R \subseteq A \times B$ is also called a *partial function* from A to B.

(b) A *function* from A to B is a functional relation $f \subseteq A \times B$ with $\operatorname{dom} f = A$ and $\operatorname{rng} f \subseteq B$. Functions are also called *maps* or *mappings*. We write $f : A \to B$ to denote the fact that f is a function from A to B.

A function of *arity n* is a function of the form

$$f : A_0 \times \dots \times A_{n-1} \to B\,.$$

We will write $x \mapsto y$ to denote the function f such that $f(x) = y$. (Usually, y will be an expression depending on x.)

(c) For a set a and a class B, we denote by B^a the class of all functions $f : a \to B$.

Remark. A 0-ary function $f : A^0 \to B$ is uniquely determined by the value $f(\langle\rangle)$. We will call such functions *constants* and identify them with their only value.

Sometimes we write binary relations and functions in infix notation, that is, for a relation $R \in A \times A$, we write $a\ R\ b$ instead of $\langle a, b\rangle \in R$ and, for $f : A \times A \to A$, we write $a\ f\ b$ instead of $f(a, b)$.

Definition 1.8. (a) For every class A, we define the *identity function* $\mathrm{id}_A : A \to A$ by $\mathrm{id}_A(a) := a$.

(b) If $R \subseteq A \times B$ and $S \subseteq B \times C$ are relations, we can define their *composition* $S \circ R : A \times C$ by

$$S \circ R := \{\, \langle a, c \rangle \mid \text{there is some } b \in B \text{ such that } \langle a, b \rangle \in R \text{ and } \langle b, c \rangle \in S \,\}\,.$$

(Note the reversal of the ordering.) In particular, if $f : A \to B$ and $g : B \to C$ are functions then

$$(g \circ f)(x) := g(f(x))\,.$$

(c) The *inverse* of a relation $R \subseteq A \times B$ is the relation

$$R^{-1} := \{\, \langle b, a \rangle \mid \langle a, b \rangle \in R \,\}\,.$$

In particular, a function $g : B \to A$ is the inverse of the function $f : A \to B$ if

$$g(f(a)) = a \quad \text{and} \quad f(g(b)) = b\,, \quad \text{for all } a \in A \text{ and } b \in B\,,$$

that is, if $g \circ f = \mathrm{id}_A$ and $f \circ g = \mathrm{id}_B$.

For $b \in B$, we will write

$$R^{-1}(b) := \{\, a \mid \langle a, b \rangle \in R \,\}\,.$$

Note that, if R^{-1} is a function, we have already defined

$$R^{-1}(b) := a \quad \text{for the unique } a \text{ such that } \langle a, b \rangle \in R\,.$$

It should always be clear from the context which of these two definitions we have in mind when we write $R^{-1}(b)$.

(d) The *restriction* of a relation $R \subseteq A \times B$ to a class C is the relation

$$R|_C := R \cap (C \times C)\,.$$

Its *left restriction* is

$$R \restriction C := R \cap (C \times B)\,.$$

(e) The *image* of a class C under a binary relation $R \subseteq A \times B$ is the class

$$R[C] := \operatorname{rng}(R \restriction C)\,.$$

Remark. The set A^A together with the operation $\circ$ forms a *monoid,* that is, $\circ$ is *associative*

$$f \circ (g \circ h) = (f \circ g) \circ h\,, \qquad \text{for all } f, g, h \in A^A,$$

and there exists a *neutral element*

$$\mathrm{id}_A \circ f = f \quad \text{and} \quad f \circ \mathrm{id}_A = f \qquad \text{for all } f \in A^A.$$

Exercise 1.1. Is it true that $R^{-1} \circ R = \mathrm{id}_A$, for all relations $R \subseteq A \times B$?

Exercise 1.2. Prove that $\circ$ is associative and that id_A is a neutral element.

Definition 1.9. Let $f : A \to B$ be a function.

(a) f is *injective* if there is no pair $a, a' \in A$ of distinct elements such that $f(a) = f(a')$.

(b) f is *surjective* if $\operatorname{rng} f = B$.

(c) f is called *bijective* if it is both injective and surjective.

Lemma 1.10. *Let $f : A \to B$ be a function.*

(a) *The following statements are equivalent:*

  (1) *f is bijective.*

  (2) *f^{-1} is a function $B \to A$.*

  (3) *There exists a function $g : B \to A$ such that $g \circ f = \mathrm{id}_A$ and $f \circ g = \mathrm{id}_B$.*

(b) *The following statements are equivalent:*

(1) *f is injective.*

(2) *$f \circ g = f \circ h$ implies $g = h$, for all functions $g, h : C \to A$.*

(3) *$A = \varnothing$ or there exists some function $g : B \to A$ such that $g \circ f = \mathrm{id}_A$.*

(4) *$f^{-1}[f[X]] = X$, for all $X \subseteq A$.*

(c) *The following statements are equivalent:*

(1) *f is surjective.*

(2) *$g \circ f = h \circ f$ implies $g = h$, for all functions $g, h : B \to C$.*

(3) *$f[f^{-1}[Y]] = Y$, for all $Y \subseteq B$.*

(d) *If there exists some function $g : B \to A$ such that $f \circ g = \mathrm{id}_B$ then f is surjective.*

Proof. (a) (1) ⇒ (2) Let $b \in B$. Since f is surjective there exists some $a \in A$ such that $f(a) = b$. If $a' \in A$ is some element with $f(a') = b$ then the injectivity of f implies that $a' = a$. We have shown that, for every element $b \in B$, there is a unique $a \in A$ such that $f^{-1}(b) = a$. Hence, f^{-1} is functional and $\mathrm{dom}\, f^{-1} = B$.

(2) ⇒ (3) $f^{-1} : B \to A$ is a function and we have $f^{-1} \circ f = \mathrm{id}_A$ and $f \circ f^{-1} = \mathrm{id}_B$.

(3) ⇒ (1) If $f(a) = f(b)$, for $a, b \in A$, then

$$a = \mathrm{id}_A(a) = (g \circ f)(a) = (g \circ f)(b) = \mathrm{id}_A(b) = b\,.$$

Consequently, f is injective. To show that it is also surjective let $b \in B$. Setting $a := g(b)$ we have

$$f(a) = (f \circ g)(b) = \mathrm{id}_B(b) = b\,.$$

Hence, $b \in \mathrm{rng}\, f$.

(b) (1) ⇒ (4) Let $X \subseteq A$. For every $a \in X$, we have $f(a) \in f[X]$ and, therefore, $a \in f^{-1}[f[X]]$. Consequently, $X \subseteq f^{-1}[f[X]]$. Conversely, suppose that $a \in f^{-1}[f[X]]$ and set $b := f(a)$. Since $b \in f[X]$ there is some $c \in X$ with $f(c) = b$. As f is injective this implies that $a = c \in X$.

(4) ⇒ (3) If $A = \varnothing$ then there is nothing to do. Hence, assume that $A \neq \varnothing$. We define g as follows. For every $b \in \operatorname{rng} f$, there is some element $a \in A$ with $f(a) = b$. Since $f^{-1}(b) = f^{-1}[f[\{a\}]] = \{a\}$ it follows that this element a is unique. Hence, fixing $a_0 \in A$ we can define g by

$$g(b) := \begin{cases} a & \text{if } f^{-1}(b) = \{a\}\,, \\ a_0 & \text{if } b \notin \operatorname{rng} f\,. \end{cases}$$

(3) ⇒ (2) If $A = \varnothing$, there are no functions $C \to A$ and the claim holds trivially. Hence, assume that $A \neq \varnothing$ and let k be a function such that $k \circ f = \mathrm{id}_A$. Then $f \circ g = f \circ h$ implies

$$g = \mathrm{id}_A \circ g = k \circ f \circ g = k \circ f \circ h = \mathrm{id}_A \circ h = h\,.$$

(2) ⇒ (1) Suppose that f is not injective. Then there are two elements $a, b \in A$ with $a \neq b$ such that $f(a) = f(b)$. Let $C := [1] = \{0\}$ be a set with a single element and define $g, h : C \to A$ by $g(0) := a$ and $h(0) := b$. Then $g \neq h$ but $f \circ g = f \circ h$.

(c) (1) ⇒ (2) Suppose that $g \neq h$. There is some element $b \in B$ with $g(b) \neq h(b)$. Since f is surjective we can find an element $a \in A$ with $f(a) = b$. Hence, $(g \circ f)(a) = g(b) \neq h(b) = (h \circ f)(a)$.

(2) ⇒ (1) Suppose that f is not surjective. Then there is some element $b \in B \setminus \operatorname{rng} f$. Let $C := [2] = \{0, 1\}$ be a set with two elements and define $g, h : B \to C$ by

$$g(x) := \begin{cases} 1 & \text{if } x = b\,, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad h(x) := 0\,, \qquad \text{for all } x \in B\,.$$

Then we have $g \neq h$ but $g \circ f = h \circ f$.

(3) ⇒ (1) $f[f^{-1}[B]] = B$ implies that $\operatorname{rng} f = B$.

(1) ⇒ (3) Let $Y \subseteq B$. If $b \in f[f^{-1}[Y]]$ then there is some $a \in f^{-1}[Y]$ with $f(a) = b$. Hence, $a \in f^{-1}[Y]$ implies that $b = f(a) \in Y$. Consequently, we have $f[f^{-1}[Y]] \subseteq Y$.

For the converse, let $b \in Y$. Since f is surjective there is some $a \in A$ with $f(a) = b$. Hence, $a \in f^{-1}[Y]$ and it follows that $b = f(a) \in f[f^{-1}[Y]]$.

(d) Let k be a function such that $f \circ k = \mathrm{id}_B$. Then $g \circ f = h \circ f$ implies

$$g = g \circ \mathrm{id}_B = g \circ f \circ k = h \circ f \circ k = h \circ \mathrm{id}_B = h\,.$$

By (c), it follows that f is surjective. □

Remark. The converse of (d) also holds but we cannot prove it without the Axiom of Choice, which we will introduce in Section A4.1 below. Actually one can prove that the Axiom of Choice is equivalent to the claim that, for every surjective function f, there exists some function g with $f \circ g = \mathrm{id}$.

Remark. The subset of all bijective functions $f : A \to A$ forms a *group* since, by the preceding lemma, every element f has an *inverse* f^{-1}.

Exercise 1.3. Let $f : A \to B$ and $g : B \to C$ be functions. Prove that, if f and g are (a) injective, (b) surjective, or (c) bijective then so is $g \circ f$.

We conclude this section with two important results about the existence of functions. The first one can be used to prove that there exists a bijection between two given sets without constructing this function explicitly.

Lemma 1.11. *Let $A \subseteq B \subseteq C$ be sets. If there exists a bijective function $f : C \to A$, there is also a bijection $g : C \to B$.*

Proof. Let

$$Z := \bigcap \{\, X \subseteq C \mid C \setminus B \subseteq X \text{ and } f[X] \subseteq X \,\}\,.$$

Then $C \setminus B \subseteq Z$ and $f[Z] \subseteq Z$. We claim that

$$g(x) := \begin{cases} f(x) & \text{if } x \in Z\,, \\ x & \text{otherwise,} \end{cases}$$

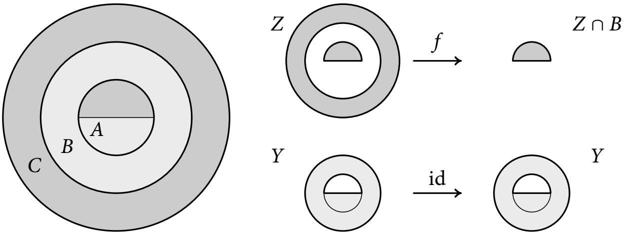


Figure 1.. The proof of Lemma 1.11.

is the desired bijection $g : C \to B$.

Let $Y := C \smallsetminus Z$ be the complement of Z. Since $g[Y] \subseteq Y$ and $g[Z] \subseteq Z$ it is sufficient to show that the restrictions $g \restriction Y : Y \to Y$ and $g \restriction Z : Z \to Z \cap B$ are bijections. Clearly, $g \restriction Y = \mathrm{id}_Y$ is bijective and $g \restriction Z = f \restriction Z$ is injective. Therefore, we only need to prove that $f[Z] = Z \cap B$.

By definition of Z, we have $f[Z] \subseteq Z \cap \operatorname{rng} f \subseteq Z \cap B$. For the other inclusion, suppose that there exists some element $a \in (Z \cap B) \smallsetminus f[Z]$. Since $a \in B$ the set $X := Z \smallsetminus \{a\}$ satisfies $C \smallsetminus B \subseteq X$ and $f[X] \subseteq X$. By definition of Z, it follows that $Z \subseteq X$. Contradiction. □

Theorem 1.12 (Bernstein). *If there are injective functions $f : A \to B$ and $g : B \to A$ then there exists a bijective function $h : A \to B$.*

Proof. We have $g[f[A]] \subseteq g[B] \subseteq A$. Since f and g are injective so is their composition $g \circ f$. When regarded as function $g \circ f : A \to g[f[A]]$ it is also surjective. Hence, by the preceding lemma, there exists a bijective mapping $h : A \to g[B]$. Since $k := g^{-1} \restriction g[B] : g[B] \to B$ is bijective it follows that so is $k \circ h : A \to B$. □

The second result deals with functions between a set and its power set.

Theorem 1.13 (Cantor). *For every set a, there exists an injective function $a \to \wp(a)$ but no surjective one.*

Proof. The function $f : a \to \wp(a)$ with $f(x) := \{x\}$ is injective.

For a contradiction, suppose that there is also a surjective function $f : a \to \wp(a)$. We define the set

$$z := \{\, x \in a \mid x \notin f(x) \,\} \subseteq a\,.$$

Since f is surjective there is some element $b \in a$ with $f(b) = z$. By definition of z, we have

$$b \in z \quad \text{iff} \quad b \notin f(b) = z\,.$$

A contradiction. □

Corollary 1.14. *For all sets a, there are no injective functions $\wp(a) \to a$.*

Proof. Suppose that $f : \wp(a) \to a$ is injective. We define a function $g : a \to \wp(a)$ by

$$g(x) := \begin{cases} f^{-1}(x) & \text{if } x \in \operatorname{rng} f\,, \\ \varnothing & \text{otherwise}\,. \end{cases}$$

Note that g is well-defined since f is injective. Furthermore, we have $g \circ f = \mathrm{id}_{\wp(A)}$. Hence, Lemma 1.10 (d) implies that g is surjective. This contradicts the Theorem of Cantor. □

2. *Products and unions*

So far, we have defined cartesian products of finitely many sets and tuples of finite length. In this section we will show how to generalise these definitions to infinitely many factors.

Remark. (a) There is a canonical bijection $\pi : A^{[n]} \to A^n$ between the set $A^{[n]}$ of all functions $[n] \to A$ and the n-th power A^n of A. π maps a function $f : [n] \to A$ to the tuple

$$\pi(f) := \langle f(0), \dots, f(n-1) \rangle\,,$$

and its inverse π^{-1} maps a tuple $\langle a_0, \dots, a_{n-1}\rangle$ to the function $f : [n] \to A$ with $f(i) = a_i$.

(b) There is also a canonical bijection $\pi : (A \times B) \times C \to A \times (B \times C)$ defined by

$$\pi\langle\langle a, b\rangle, c\rangle := \langle a, \langle b, c\rangle\rangle .$$

(c) Finally, let us define a canonical bijection $\pi : A^{B\times C} \to (A^C)^B$ that maps a function $f : B \times C \to A$ to the function $g : B \to A^C$ with

$$g(b) := h_b \quad \text{where} \quad h_b(c) := f(b, c), \qquad \text{for } b \in B,\ c \in C .$$

In the theory of programming languages this transformation of a function $B \times C \to A$ into a function $B \to A^C$ is called *currying.*

Part (a) of the above remark gives a hint on how to generalise finite tuples. A tuple of length n corresponds to a map $[n] \to A$. Therefore, we define an infinite tuple as map $\mathbb{N} \to A$.

Definition 2.1. (a) Let A be a class and I a set. A function $f : I \to A$ is called a *sequence,* or *family,* over I. If $f(i) = a_i$ then we also write f in the form $(a_i)_{i\in I}$.

(b) Let I be a set, $(A_i)_{i\in I}$ a sequence of sets, and $A := \bigcup\{ A_i \mid i \in I \}$ their union. The *product* of $(A_i)_{i\in I}$ is the class

$$\prod_{i\in I} A_i := \{ f \in A^I \mid f(i) \in A_i \text{ for all } i \} .$$

(c) Let $(A_i)_{i\in I}$ be a sequence of sets and $k \in I$. The *projection* to the k-th coordinate is the map

$$\mathrm{pr}_k : \prod_{i\in I} A_i \to A_k \quad \text{with} \quad \mathrm{pr}_k(f) := f(k) .$$

Remark. (a) If $A_i = A$, for all $i \in I$, then $\prod_{i\in I} A_i = A^I$.

(b) As we have seen above there is a canonical bijection between $A_0 \times A_1$ and $\prod_{i\in[2]} A_i$. In the following we will not distinguish between these sets.

Let us introduce some notation and conventions regarding sequences. To indicate that a certain variable refers to a sequence we will write it with a bar $\bar{a}$. If the sequence is over I, the components of $\bar{a}$ will always be $(a_i)_{i\in I}$. Sometimes we will not distinguish between a sequence $\bar{a} = (a_i)_{i\in I}$ and its range $\operatorname{rng} \bar{a} = \{ a_i \mid i \in I \}$. In particular, we write $\bar{a} \cup \bar{b}$ instead of $\operatorname{rng} \bar{a} \cup \operatorname{rng} \bar{b}$ and, if we do not want to specify the index set I, we will write $\bar{a} \subseteq A$ instead of $\bar{a} \in A^I$. Finally, for a function $f : A \to B$, we write $f(\bar{a})$ to denote the sequence $(f(a_i))_{i\in I}$.

Lemma 2.2. *Let A be a set and $(B_i)_{i\in I}$ a sequence of sets. For every sequence $(f_i)_{i\in I}$ of functions $f_i : A \to B_i$ there exists a unique function $g : A \to \prod_i B_i$ such that*

$$\mathrm{pr}_i \circ g = f_i\,, \qquad \text{for all } i \in I\,.$$

Proof. The function

$$g(a) := (f_i(a))_{i\in I}$$

has obviously the desired properties. We have to show that it is unique. Let $h : A \to \prod_i B_i$ be another such function. If $g \neq h$, there is some element $a \in A$ such that $g(a) \neq h(a)$. Let $(b_i)_{i\in I} := h(a)$. For every $i \in I$, we have

$$b_i = (\mathrm{pr}_i \circ h)(a) = f_i(a)\,.$$

Hence $g(a) = (f_i(a))_i = (b_i)_i = h(a)$. A contradiction. □

Definition 2.3. The *disjoint union* of a sequence $(A_i)_{i\in I}$ of sets is the class

$$\dot{\bigcup_{i\in I}} A_i := \{ \langle i, a\rangle \mid i \in I,\ a \in A_i \}\,.$$

Similarly, if A and B are classes then we can define their disjoint union as

$$A \mathbin{\dot\cup} B := (\{[0]\} \times A) \cup (\{[1]\} \times B)\,.$$

The k-th *insertion* is the canonical map

$$\mathrm{in}_k : A_k \to \dot{\bigcup_{i \in I}} A_i \quad \text{with} \quad \mathrm{in}_k(a) := \langle k, a \rangle .$$

Remark. If $A_i = A$, for all $i \in I$, then $\dot\bigcup_{i \in I} A_i = I \times A$.

Lemma 2.4. *Let B be a set and $(A_i)_{i \in I}$ a sequence of sets. For every sequence $(f_i)_{i \in I}$ of functions $f_i : A_i \to B$ there exists a unique function $g : \dot\bigcup_i A_i \to B$ such that*

$$g \circ \mathrm{in}_i = f_i , \qquad \textit{for all } i \in I .$$

Proof. The function

$$g\langle i, a \rangle := f_i(a)$$

has obviously the desired properties. We have to show that it is unique. Let $h : \dot\bigcup_i A_i \to B$ be another such function. If $g \neq h$ then there is some element $\langle k, a \rangle \in \dot\bigcup_i A_i$ such that $g\langle k, a \rangle \neq h\langle k, a \rangle$. We have

$$h\langle k, a \rangle = (h \circ \mathrm{in}_k)(a) = f_k(a) = g\langle k, a \rangle .$$

A contradiction. ☐

3. *Graphs and partial orders*

When considering relations it is frequently necessary to specify the sets they are over.

Definition 3.1. A *graph* is a pair $\langle A, R \rangle$ where $R \subseteq A \times A$ is a binary relation on A.

More generally one can consider sets together with several relations and functions. This will lead to the notion of a structure in Chapter B1.

Definition 3.2. Let $\langle A, R \rangle$ be a graph.

(a) R is *reflexive* if $\langle a, a \rangle \in R$, for all $a \in A$.

(b) R is *irreflexive* if $\langle a, a \rangle \notin R$, for all $a \in A$.

(c) R is *symmetric* if we have $\langle a, b \rangle \in R$ if, and only if, $\langle b, a \rangle \in R$, for all $a, b \in A$.

(d) R is *antisymmetric* if $\langle a, b \rangle \in R$ and $\langle b, a \rangle \in R$ implies $a = b$.

(e) R is *transitive* if $\langle a, b \rangle \in R$ and $\langle b, c \rangle \in R$ implies $\langle a, c \rangle \in R$, for all $a, b, c \in A$.

Note that, for the definition of reflexivity, it is important to specify the set A. If $\langle A, R \rangle$ is reflexive and $A \subset B$ then $\langle B, R \rangle$ is not reflexive.

Example. (a) The relation $A \times A$ is reflexive, symmetric, and transitive. It is irreflexive if, and only if, $A = \varnothing$, and it is antisymmetric if, and only if, A contains at most one element.

(b) The diagonal $\mathrm{id}_A = \{ \langle a, a \rangle \mid a \in A \}$ is reflexive, symmetric, antisymmetric, and transitive. It is irreflexive if, and only if, $A = \varnothing$.

(c) The empty relation $\varnothing \subseteq A \times A$ is irreflexive, symmetric, antisymmetric, and transitive. It is reflexive if, and only if, $A = \varnothing$.

Definition 3.3. (a) A (non-strict) *partial order* is a graph $\langle A, \leq \rangle$ where $\leq$ is reflexive, transitive, and antisymmetric.

(b) A *strict partial order* is a graph $\langle A, < \rangle$ where $<$ is irreflexive and transitive.

(c) A partial order $\langle A, \leq \rangle$ is *linear,* or *total,* if

$$a \leq b \text{ or } b \leq a\,, \quad \text{for all } a, b \in A\,.$$

(d) Instead of saying that $\langle A, R \rangle$ is a partial or linear order we also say that R is a partial/linear order on A, or that R *orders* A partially/linearly.

(e) If $\mathfrak{A} = \langle A, \leq \rangle$ is a partial order, we denote by $\mathfrak{A}^{\mathrm{op}} := \langle A, \leq^{-1} \rangle$ the graph where the order relation is reversed. $\mathfrak{A}^{\mathrm{op}}$ is called the *opposite order.*

Remark. (a) To each non-strict partial order $\leq$ on A we can associate the strict partial order

$$a < b \quad :\text{iff} \quad a \leq b \text{ and } a \neq b\,.$$

Similarly, if $<$ is a strict partial order on A, we can define a non-strict version by

$$a \leq b \quad :\text{iff} \quad a < b \text{ or } a = b\,.$$

(b) If $\mathfrak{A}$ is a partial order then so is $\mathfrak{A}^{\mathrm{op}}$.

Example. (a) The subset relation $\subseteq$ is a partial order on $\mathbb{S}$.
(b) The usual ordering $\leq$ is a linear order on the rational numbers $\mathbb{Q}$.
(c) The divisibility relation

$$a \mid b \quad :\text{iff} \quad b = ac \text{ for some } c$$

is a partial order on the natural numbers $\mathbb{N}$.

Definition 3.4. Let $\mathfrak{A} = \langle A, \leq \rangle$ be a partial order.
(a) An *initial segment* of A is a subset $I \subseteq A$ such that $a \in I$ and $b \leq a$ implies $b \in I$. Similarly, a *final segment* of A is a subset $F \subseteq A$ such that $a \in F$ and $b \geq a$ implies $b \in F$.
(b) A set $X \subseteq A$ *generates* the segments

$$\Downarrow_{\mathfrak{A}} X := \{\, a \in A \mid a \leq b \text{ for some } b \in X \,\}\,,$$
$$\text{and} \quad \Uparrow_{\mathfrak{A}} X := \{\, a \in A \mid a \geq b \text{ for some } b \in X \,\}\,.$$

For $X = \{x\}$, we also write $\Downarrow_{\mathfrak{A}} x$ and $\Uparrow_{\mathfrak{A}} x$. Similarly, we define

$$\downarrow_{\mathfrak{A}} X := \{\, a \in A \mid a < b \text{ for some } b \in X \,\}\,,$$
$$\text{and} \quad \uparrow_{\mathfrak{A}} X := \{\, a \in A \mid a > b \text{ for some } b \in X \,\}\,.$$

Finally, we set

$$[a, b]_{\mathfrak{A}} := \Uparrow_{\mathfrak{A}} a \cap \Downarrow_{\mathfrak{A}} b \quad \text{and} \quad (a, b)_{\mathfrak{A}} := \uparrow_{\mathfrak{A}} a \cap \downarrow_{\mathfrak{A}} b\,.$$

(c) Let $X \subseteq A$ and $a \in X$. We call a the *greatest element* of X if $x \leq a$, for all $x \in X$. And we say that a is *maximal* if there is no $x \in X$ with $a < x$. *Least* and *minimal* elements are defined analogously. We denote the greatest element of X by $\max_{\mathfrak{A}} X$ and the least element by $\min_{\mathfrak{A}} X$, provided these elements exist.

(d) Let $X \subseteq A$. We say that a is an *upper bound* of X if $x \leq a$, for all $x \in X$. If a is an upper bound of X and $a \leq b$, for every other upper bound b of X, then a is the *least upper bound,* or *supremum,* of X. If the least upper bound of X exists, we denote it by $\sup_{\mathfrak{A}} X$.

The notion of a (*greatest*) *lower bound* is defined analogously. The greatest lower bound is also called the *infimum* of X. We denote it by $\inf_{\mathfrak{A}} X$. If the order $\mathfrak{A}$ is understood we will omit the subscript $\mathfrak{A}$ and we just write $\sup X$ and $\inf X$.

(e) A linearly ordered subset $C \subseteq A$ is called a *chain.*

Example. (a) Let $\mathfrak{Q} := \langle \mathbb{Q}, \leq \rangle$. The set $I := \{ x \in \mathbb{Q} \mid x < \sqrt{2} \}$ is an initial segment of $\mathfrak{Q}$. Every rational number $y > \sqrt{2}$ is an upper bound of I but I has no least upper bound.

(b) Consider $\langle \mathbb{N}, \mid \rangle$. Its least element is the number 1 and its greatest element is 0. The least upper bound of two elements $[k], [m] \in \mathbb{N}$ is their least common multiple $\operatorname{lcm}(k, m)$, and their greatest lower bound is their greatest common divisor $\gcd(k, m)$. The set $P \subseteq \mathbb{N}$ of all prime numbers has the least upper bound 0 and the greatest lower bound 1. The set $\{ 2^n \mid n \in \mathbb{N} \}$ of all powers of two forms a chain.

Exercise 3.1. Consider $\langle B, \subseteq \rangle$ where

$$B := \{ X \subseteq \mathbb{N} \mid X \text{ is finite or } \mathbb{N} \setminus X \text{ is finite} \} .$$

(a) Construct a set $X \subseteq B$ that has no minimal element.

(b) Construct a set $X \subseteq B$ with lower bounds but without infimum.

Lemma 3.5. *Let $\langle A, \leq \rangle$ be a partial order. If A is a set, the following statements are equivalent:*

(1) *Every subset $X \subseteq A$ has a supremum.*

(2) *Every subset $X \subseteq A$ has an infimum.*

Proof. We only prove (1) $\Rightarrow$ (2). The other direction follows in exactly the same way. Let $X \subseteq A$ and set

$$C := \{ a \in A \mid a \text{ is a lower bound of } X \} .$$

By assumption, $c := \sup C$ exists. We claim that $\inf X = c$. Let $b \in X$. By definition, we have $a \leq b$, for all $a \in C$. Hence, b is an upper bound of C and we have $b \geq \sup C = c$. As b was arbitrary it follows that c is a lower bound of X. If a is an arbitrary lower bound of X, we have $a \in C$, which implies that $a \leq c$. Consequently, c is the greatest lower bound of X. □

Definition 3.6. A partial order $\langle A, \leq \rangle$ is *complete* if every subset $X \subseteq A$ has an infimum and a supremum.

Remark. Every complete partial order has a least element $\bot := \sup \varnothing$ and a greatest element $\top := \inf \varnothing$.

Example. (a) Let A be a set. The partial order $\langle \wp(A), \subseteq \rangle$ is complete. If $X \subseteq \wp(A)$ then

$$\sup X = \bigcup X \in \wp(A) \quad \text{and} \quad \inf X = \bigcap X \in \wp(A) .$$

(b) The order $\langle \mathbb{R}, \leq \rangle$ is complete. $\langle \mathbb{Q}, \leq \rangle$ is not since the set

$$\{ x \in \mathbb{Q} \mid x \leq \pi \}$$

has no least upper bound in $\mathbb{Q}$.

(c) The order $\langle \mathbb{N}, \leq \rangle$ is not complete since $\inf \varnothing$ and $\sup \mathbb{N}$ do not exist.

(d) Let $\mathfrak{A} = \langle A, \leq \rangle$ be an arbitrary partial order. We can construct a complete partial order $\mathfrak{C} = \langle C, \subseteq \rangle$ containing $\mathfrak{A}$ as follows. Let $C \subseteq \wp(A)$ be the set of all initial segments of A ordered by inclusion. The desired embedding $f : A \to C$ is given by $f(a) := \Downarrow_A a$.

Next we turn to the study of functions between partial orders. In particular, we will consider functions $f : A \to A$ mapping one partial order into itself. To simplify notation, we will write

$$f : \mathfrak{A} \to \mathfrak{B},$$

for partial orders $\mathfrak{A} = \langle A, \leq_A \rangle$ and $\mathfrak{B} = \langle B, \leq_B \rangle$, to denote that f is a function $f : A \to B$.

Definition 3.7. Let $\mathfrak{A} = \langle A, \leq_A \rangle$ and $\mathfrak{B} = \langle B, \leq_B \rangle$ be partial orders.
(a) A function $f : A \to B$ is *increasing* if

$$a \leq_A b \quad \text{implies} \quad f(a) \leq_B f(b), \qquad \text{for all } a, b \in A,$$

and f is *strictly increasing* if

$$a <_A b \quad \text{implies} \quad f(a) <_B f(b), \qquad \text{for all } a, b \in A.$$

(b) A function $f : A \to B$ is an *embedding* if we have

$$a \leq_A b \quad \text{iff} \quad f(a) \leq_B f(b), \qquad \text{for all } a, b \in A.$$

A bijective embedding is called an *isomorphism*. If there exists an isomorphism $f : A \to B$ then we say that $\mathfrak{A}$ and $\mathfrak{B}$ are *isomorphic* and we write $\mathfrak{A} \cong \mathfrak{B}$.

Remark. Every isomorphism is strictly increasing.

Exercise 3.2. Define a function that is

(a) increasing but not strictly increasing;

(b) strictly increasing but not an embedding;

(c) an embedding but not an isomorphism.

Exercise 3.3. Construct a strictly increasing function

$$f : \langle \mathbb{N}, \mid \rangle \to \langle \wp(\mathbb{N}), \subseteq \rangle.$$

Lemma 3.8. *Let $\langle A, \leq_A \rangle$ and $\langle B, \leq_B \rangle$ be partial orders and $h : A \to B$ an increasing function. Let $C \subseteq A$ and $a \in A$.*

(a) *If a is an upper bound of C then $h(a)$ is an upper bound of $h[C]$.*

(b) *If a is a lower bound of C then $h(a)$ is a lower bound of $h[C]$.*

Lemma 3.9. *Let $\langle A, \leq_A \rangle$ and $\langle B, \leq_B \rangle$ be partial orders and $h : A \to B$ an embedding. Let $C \subseteq A$ and $a \in A$.*

(a) $h(a) = \sup h[C]$ *implies* $a = \sup C$.

(b) $h(a) = \inf h[C]$ *implies* $a = \inf C$.

Proof. (a) Since h is an embedding it follows that $h(c) \leq_B h(a)$ implies $c \leq_A a$, for $c \in C$. Hence, a is an upper bound of C. To show that it is the least one, suppose that b is another upper bound of C. Then $c \leq_A b$, for $c \in C$, implies $h(c) \leq_B h(b)$. Hence, $h(b)$ is an upper bound of $h[C]$. Since $h(a)$ is the least such bound it follows that $h(a) \leq_B h(b)$. Consequently, we have $a \leq_A b$, as desired.

(b) h is also an embedding of $\langle A, \geq_A \rangle$ into $\langle B, \geq_B \rangle$. Hence, (b) follows from (a) by reversing the orders. □

Corollary 3.10. *Let $\langle F, \subseteq \rangle$ be a partial order with $F \subseteq \wp(A)$ and $C \subseteq F$.*

(a) $\bigcup C \in F$ *implies* $\sup C = \bigcup C$.

(b) $\bigcap C \in F$ *implies* $\inf C = \bigcap C$.

Proof. We can apply Lemma 3.9 to the inclusion map $F \to \wp(A)$. □

Corollary 3.11. *Let $\mathfrak{A} = \langle A, \leq \rangle$ be a partial order. If $B \subseteq A$ is a nonempty set such that*

$$\inf_{\mathfrak{A}} X \in B \quad \textit{and} \quad \sup_{\mathfrak{A}} X \in B\,, \quad \textit{for every nonempty } X \subseteq B\,,$$

then $\mathfrak{B} := \langle B, \leq \rangle$ is a complete partial order where, for every nonempty subset $X \subseteq B$, we have

$$\inf_{\mathfrak{B}} X = \inf_{\mathfrak{A}} X \quad \textit{and} \quad \sup_{\mathfrak{B}} X = \sup_{\mathfrak{A}} X\,.$$

Proof. If $X \subseteq B$ is nonempty then, applying Lemma 3.9 to the inclusion map $\mathfrak{B} \to \mathfrak{A}$, it follows that

$$\inf_{\mathfrak{B}} X = \inf_{\mathfrak{A}} X \quad \text{and} \quad \sup_{\mathfrak{B}} X = \sup_{\mathfrak{A}} X\,.$$

In particular, $\inf_{\mathfrak{B}} X$ and $\sup_{\mathfrak{B}} X$ exist. For the empty set, it follows similarly that

$$\begin{aligned} &\inf_{\mathfrak{B}} \varnothing = \sup_{\mathfrak{B}} B = \sup_{\mathfrak{A}} B \in B\,, \\ \text{and} \quad &\sup_{\mathfrak{B}} \varnothing = \inf_{\mathfrak{B}} B = \inf_{\mathfrak{A}} B \in B\,. \end{aligned}$$

Consequently, $\mathfrak{B}$ is complete. □

We have seen that although increasing functions preserve the ordering of elements they do not necessarily preserve supremums and infimums. Let us take a look at functions that do.

Definition 3.12. Let $\langle A, \leq_A \rangle$ and $\langle B, \leq_B \rangle$ be partial orders. A function $f : A \to B$ is *continuous* if, whenever a nonempty chain $C \subseteq A$ has a supremum then $f[C]$ also has a supremum and we have

$$\sup f[C] = f(\sup C)\,.$$

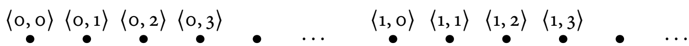
f is called *strictly continuous* if it is continuous and injective.

Remark. Every (strictly) continuous function is (strictly) increasing.

Exercise 3.4. Prove that continuous functions are increasing.

Example. (a) Let $\langle A, \leq \rangle$ be the linear order where $A = \mathbb{N} \cup \mathbb{N}$ and

$$\langle i, a \rangle \leq \langle k, b \rangle \quad :\text{iff} \quad i < k\,, \text{ or } i = k \text{ and } a \leq b\,.$$



The function $f : A \to A : \langle i, a\rangle \mapsto \langle i, a+1\rangle$ is not continuous. Consider the initial segment $X := \{0\} \times \mathbb{N} = {\downarrow}\langle 1, 0\rangle \subseteq A$. We have $\sup X = \langle 1, 0\rangle$ but

$$\sup f[X] = \langle 1, 0\rangle < \langle 1, 1\rangle = f(\langle 1, 0\rangle)\,.$$

(b) Let A be a set and $\langle F, \subseteq\rangle$ the partial order with

$$F := \{\, X \subseteq A \mid A \setminus X \text{ is finite} \,\}\,.$$

For every bijective function $\sigma : A \to A$ we obtain a continuous mapping $f : F \to F$ by setting

$$f(X) := \{\, \sigma(x) \mid x \in X \,\}\,.$$

Lemma 3.13. *Every isomorphism $f : \mathfrak{A} \to \mathfrak{B}$ is strictly continuous.*

Proof. Let $C \subseteq A$ be a nonempty chain with supremum. For every $a \in C$, we have $a \leq \sup C$, which implies that $f(a) \leq f(\sup C)$. Hence,

$$\sup f[C] \leq f(\sup C)\,.$$

For the converse, let $b := \sup f[C]$. By Lemma 3.9, it follows that $\sup C = f^{-1}(b)$. □

4. *Fixed points and closure operators*

Many objects can be defined as solution to an equation of the form $x = f(x)$. Such solutions are called *fixed points* of the function f. For example, the solutions of a system of linear equations $Ax = b$ are exactly the fixed points of the function

$$f(x) := Ax + x - b\,.$$

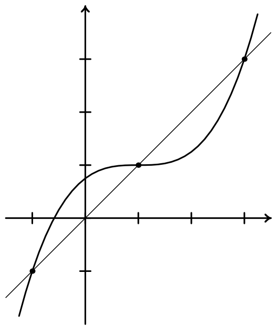


Figure 2.. Fixed points of $f(x) = \frac{1}{4}x^3 - \frac{3}{4}x^2 + \frac{3}{4}x + \frac{3}{4}$

Definition 4.1. Let $f : A \to A$ be a function. An element $a \in A$ with $f(a) = a$ is called a *fixed point* of f. The class of all fixed points of f is denoted by

$$\operatorname{fix} f := \{ a \in A \mid f(a) = a \} .$$

We denote the *least* and *greatest* fixed point of f, if it exists, by

$$\operatorname{lfp} f := \min \operatorname{fix} f \quad \text{and} \quad \operatorname{gfp} f := \max \operatorname{fix} f .$$

Example. (a) Let $\langle \mathbb{R}, < \rangle$ be the order of the real numbers. The function

$$f(x) := \tfrac{1}{4}x^3 - \tfrac{3}{4}x^2 + \tfrac{3}{4}x + \tfrac{3}{4}$$

has 3 fixed points: $\operatorname{fix} f = \{-1, 1, 3\}$.

(b) Consider $\langle \mathbb{N}, \leq \rangle$. The function $f : \mathbb{N} \to \mathbb{N}$ with $f(n) := n + 1$ has no fixed points.

(c) Consider $\langle \wp[2], \subseteq \rangle$. The function $f : \wp[2] \to \wp[2]$ with

$$f(x) := \begin{cases} \{0\} & \text{if } x = \varnothing , \\ x & \text{otherwise} , \end{cases}$$

has the fixed points $\{0\}$, $\{1\}$, $\{0, 1\}$. It has no least fixed point.

(d) Consider $\langle F, \subseteq \rangle$ where

$$F := \{ X \subseteq \mathbb{N} \mid X \text{ or } \mathbb{N} \setminus X \text{ is finite} \} .$$

The function $f : F \to F$ defined by

$$f(X) := \begin{cases} X \cup \{1 + \max X\} & \text{if } X \text{ is finite,} \\ X & \text{otherwise,} \end{cases}$$

has fixed points

$$\operatorname{fix} f = \{ X \subseteq \mathbb{N} \mid \mathbb{N} \setminus X \text{ is finite} \} ,$$

but no least one.

Exercise 4.1. Let $\mathfrak{A} = \langle \mathscr{P}(\mathbb{N}), \subseteq \rangle$. Construct a function $f : \mathfrak{A} \to \mathfrak{A}$ that has a least fixed point but no greatest one.

Not every function has fixed points. The next theorem presents an important special case where we always have a least fixed point. In Section A3.3 we will collect further results about the existence of fixed points and methods to compute them.

Theorem 4.2 (Knaster, Tarski). *Let $\langle A, \leq \rangle$ be a complete partial order where A is a set. Every increasing function $f : A \to A$ has a least fixed point and we have*

$$\operatorname{lfp} f = \inf \{ a \in A \mid f(a) \leq a \} .$$

Proof. Set $B := \{ a \in A \mid f(a) \leq a \}$ and $b := \inf B$. For every $a \in B$, $b \leq a$ implies $f(b) \leq f(a) \leq a$, since f is increasing. Hence, $f(b)$ is a lower bound of B and it follows that $f(b) \leq \inf B = b$. This implies that $f(f(b)) \leq f(b)$ and, by definition of B, it follows that $f(b) \in B$. Hence, $f(b) \geq \inf B = b$. Consequently, we have $f(b) = b$ and b is a fixed point of f.

Let a be another fixed point of f. Then $f(a) = a$ implies $a \in B$ and we have $b = \inf B \leq a$. Hence, b is the least fixed point of f. □

Theorem 4.3. *Let $\langle A, \leq \rangle$ be a complete partial order where A is a set and let $f : A \to A$ be increasing. The set $F := \operatorname{fix} f$ is nonempty and $\mathfrak{F} := \langle F, \leq \rangle$ forms a complete partial order where, for $X \subseteq F$,*

$$\inf_{\mathfrak{F}} X = \sup_{\mathfrak{A}} \{ a \in A \mid a \leq \inf_{\mathfrak{A}} X \text{ and } f(a) \geq a \},$$
$$\sup_{\mathfrak{F}} X = \inf_{\mathfrak{A}} \{ a \in A \mid a \geq \sup_{\mathfrak{A}} X \text{ and } f(a) \leq a \}.$$

Proof. We have already shown in the preceding theorem that $F \neq \emptyset$. It remains to prove that $\mathfrak{F}$ is complete. For $X \subseteq A$, let $U := \Uparrow \sup_{\mathfrak{A}} X \subseteq A$ be the set of all upper bounds of X. If $Z \subseteq U$ then

$$\sup_{\mathfrak{A}} Z \geq \sup_{\mathfrak{A}} X \quad \text{and} \quad \inf_{\mathfrak{A}} Z \geq \sup_{\mathfrak{A}} X .$$

It follows that the partial order $\langle U, \leq \rangle$ is complete. Furthermore, if $a \in U$ and $x \in X$ then $a \geq x$ implies $f(a) \geq f(x)$. Hence, $f \restriction U$ is an increasing function $U \to U$. By Theorem 4.2, it follows that

$$\sup_{\mathfrak{F}} X = \operatorname{lfp}(f \restriction U) = \inf_{\mathfrak{A}} \{ a \in U \mid f(a) \leq a \},$$

as desired. The claim for $\inf_{\mathfrak{F}} X$ follows by applying the equation for $\sup_{\mathfrak{F}} X$ to the opposite order $\mathfrak{A}^{\mathrm{op}}$. □

Example. Consider a closed interval $[a, b] \subseteq \mathbb{R}$ of the real line.

(a) Since the order $\langle [a, b], < \rangle$ is complete, it follows by the Theorem of Knaster and Tarski that every increasing function $f : [a, b] \to [a, b]$ has a fixed point.

(b) Let $f : [0, 2] \to [0, 2]$ be the polynomial function

$$f(x) := \tfrac{1}{4}x^3 - \tfrac{3}{4}x^2 + \tfrac{3}{4}x + \tfrac{3}{4}$$

from Figure 2. We have $\{ x \mid f(x) \leq x \} = [1, 2]$ and $\operatorname{lfp} f = 1$.

(c) The order $\langle \mathbb{R}, < \rangle$ is not complete. Again, let $f : \mathbb{R} \to \mathbb{R}$ by the function from Figure 2. We have already seen that its fixed points are -1, 1, and 3. But the set

$$\{ x \mid f(x) \leq x \} = (-\infty, -1] \cup [1, 3]$$

has no minimal element.

As a special case of Theorem 4.3 we consider complete partial orders obtained via closure operators.

Definition 4.4. Let A be a class.

(a) A *closure operator* on A is a function $c : \wp(A) \to \wp(A)$ such that, for all $x, y \in \wp(A)$,

- $x \subseteq c(x)$,
- $c(c(x)) = c(x)$, and
- $x \subseteq y$ implies $c(x) \subseteq c(y)$.

(b) A set $x \subseteq A$ is *c-closed* if $c(x) = x$.

(c) A closure operator c has *finite character* if, for all sets $x \subseteq A$, we have

$$c(x) = \bigcup \{ c(x_0) \mid x_0 \subseteq x \text{ is finite} \} .$$

If c has finite character we also say that c is *algebraic*.

(d) A closure operator c is *topological* if we have

- $c(\varnothing) = \varnothing$ and
- $c(x \cup y) = c(x) \cup c(y)$, for all $x, y \in \wp(A)$.

Remark. Let c be a closure operator on A.

(a) The class of c-closed sets is $\operatorname{fix} c = \operatorname{rng} c$.

(b) If the class A is a set then it is c-closed.

Example. (a) Let V be a vector space. For $X \subseteq V$, let $\langle\langle X \rangle\rangle$ be the subspace of V spanned by X. The function $X \mapsto \langle\langle X \rangle\rangle$ is a closure operator with finite character.

(b) Let X be a topological space. For $A \subseteq X$, let $c(A)$ be the topological closure of A in X. Then c is a topological closure operator.

(c) Let A be a set and $a \in A$. The functions $c, d : \wp(A) \to \wp(A)$ with

$$c(X) := X \quad \text{and} \quad d(X) := X \cup \{a\}$$

are closure operators on A.

Exercise 4.2. Let $\mathfrak{A} = \langle A, \leq \rangle$ be a partial order. For $X \subseteq A$, we define

$$c(X) := \{ \sup C \mid C \subseteq X \text{ is a nonempty chain with supremum} \} .$$

(a) Prove that the function c is a topological closure operator on A.

(b) Let $\mathfrak{B}$ be a second partial order and d the corresponding closure operator. Prove that a function $f : \mathfrak{A} \to \mathfrak{B}$ is continuous if, and only if, every d-closed set $X \in \operatorname{fix} d$ has a c-closed preimage $f^{-1}[X] \in \operatorname{fix} c$.

Exercise 4.3. Let $\langle A, \leq \rangle$ be a partial order. For sets $X \subseteq A$, we define

$$U(X) := \{ a \in A \mid a \text{ is an upper bound of } X \} ,$$
$$L(X) := \{ a \in A \mid a \text{ is a lower bound of } X \} .$$

Prove that the function $c : X \mapsto L(U(X))$ is a closure operator on A.

Lemma 4.5. *Let c be a closure operator on A and $x, y \subseteq A$ sets.*

(a) $c(x) \cup c(y) \subseteq c(x \cup y)$.

(b) $c(x \cup y) = c(c(x) \cup c(y))$.

Proof. (a) By monotonicity of c, we have $c(x) \subseteq c(x \cup y)$ and $c(y) \subseteq c(x \cup y)$.

(b) It follows from $x \cup y \subseteq c(x) \cup c(y)$ and (a) that

$$c(x \cup y) \subseteq c(c(x) \cup c(y)) \subseteq c(c(x \cup y)) = c(x \cup y) . \qquad \square$$

Lemma 4.6. *Let c be a closure operator on A with finite character. For every chain $C \subseteq \operatorname{fix} c$, we have*

$$c(\bigcup C) = \bigcup C .$$

Proof. By definition, we have $\bigcup C \subseteq c(\bigcup C)$. For the converse, let $x_0 \subseteq \bigcup C$ be finite. Since C is linearly ordered by $\subseteq$ there exists some element $x \in C$ with $x_0 \subseteq x$. Hence, we have $c(x_0) \subseteq c(x) = x \subseteq \bigcup C$. It follows that

$$c(\bigcup C) = \bigcup \{ c(x_0) \mid x_0 \subseteq \bigcup C \text{ finite} \} \subseteq \bigcup C . \qquad \square$$

If c is a closure operator, the set $\mathcal{C} := \operatorname{fix} c$ of c-closed sets has the following properties.

Definition 4.7. A set $\mathcal{C} \subseteq \wp(A)$ is called a *system of closed sets* if we have

- $A \in \mathcal{C}$ and
- $\bigcap Z \in \mathcal{C}$, for every $Z \subseteq \mathcal{C}$.

A pair $\langle A, \mathcal{C} \rangle$ where $\mathcal{C} \subseteq \wp(A)$ is a system of closed sets is called a *closure space.*

Lemma 4.8. (a) *If c is a closure operator on A then* fix c *forms a system of closed sets.*

(b) *If $\mathcal{C} \subseteq \wp(A)$ is a system of closed sets then the mapping*

$$c : X \mapsto \bigcap \{ C \in \mathcal{C} \mid X \subseteq C \}$$

defines a closure operator on A with $\operatorname{fix} c = \mathcal{C}$.

The following theorem states that the family of c-closed sets forms a complete partial order. We can use this result to prove that a given partial order $\mathfrak{A}$ is complete by defining a closure operator whose closed sets are exactly the elements of $\mathfrak{A}$. An example of such a proof is provided in Corollary 4.17.

Theorem 4.9. *Let A be a set and c a closure operator on A. The graph $\langle F, \subseteq \rangle$ with $F :=$* fix c *forms a complete partial order with*

$$\inf X = \bigcap X \quad \textit{and} \quad \sup X = c\big(\bigcup X\big)\,, \quad \textit{for all } X \subseteq F\,.$$

Proof. Since closure operators are increasing we can apply Theorem 4.3. By Lemma 4.8 (b), it follows that

$$\begin{aligned}
\sup X &= \bigcap \{ Z \subseteq A \mid Z \supseteq \bigcup X \text{ and } c(Z) \subseteq Z \} \\
&= \bigcap \{ Z \subseteq A \mid Z \supseteq \bigcup X \text{ and } c(Z) = Z \} \\
&= c\big(\bigcup X\big)\,,
\end{aligned}$$

$$\begin{aligned}\text{and}\quad \inf X &= \bigcup \{\, Z \subseteq A \mid Z \subseteq \bigcap X \text{ and } c(Z) \supseteq Z \,\} \\ &= \bigcup \{\, Z \subseteq A \mid Z \subseteq \bigcap X \,\} \\ &= \bigcap X \,. \end{aligned}$$ □

Corollary 4.10. *Let c be a closure operator on A and set F := fix c. The operator c is continuous if we consider it as a function*

$$c : \langle \wp(A), \subseteq \rangle \to \langle F, \subseteq \rangle \,.$$

Proof. For a nonempty chain $X \subseteq \wp(A)$, we have

$$\begin{aligned} c(\sup X) = c(\bigcup X) \subseteq c(\bigcup c[X]) &= \sup c[X] \\ &\subseteq \sup \{c(\sup X)\} = c(\sup X) \,. \end{aligned}$$ □

As an application of closure operators we consider equivalence relations.

Definition 4.11. (a) A binary relation $\sim \subseteq A \times A$ is an *equivalence relation* on A if it is reflexive, symmetric, and transitive.

(b) Let $\sim \subseteq A \times A$ be an equivalence relation. If A is a set, we define the $\sim$-*class* of an element $a \in A$ by

$$[a]_\sim := \{\, b \in A \mid b \sim a \,\} \,.$$

For proper classes A, we set

$$[a]_\sim := \text{cut}\,\{\, b \in A \mid b \sim a \,\} \,.$$

Note that, despite the name, a $\sim$-class is always a set. We denote the class of all $\sim$-classes by

$$A/{\sim} := \{\, [a]_\sim \mid a \in A \,\} \,.$$

Example. (a) The diagonal id_A is the smallest equivalence relation on A. The largest one is the full relation $A \times A$.

(b) The isomorphism relation $\cong$ is an equivalence relation on the class of all partial orders.

Lemma 4.12. *Let $\sim$ be an equivalence relation on A and $a, b \in A$. Then*

$$a \sim b \quad \text{iff} \quad [a]_\sim = [b]_\sim \quad \text{iff} \quad [a]_\sim \cap [b]_\sim \neq \varnothing .$$

Remark. Let A be a set. A *partition* of A is a set $P \subseteq \wp(A)$ of nonempty subsets of A such that $A = \bigcup P$ and $p \cap q = \varnothing$, for all $p, q \in P$ with $p \neq q$.

If $\sim$ is an equivalence relation on A then $A/\sim$ forms a partition on A. Conversely, given a partition P of A, we can define an equivalence relation $\sim_P$ on A with $A/\sim_P = P$ by setting

$$a \sim_P b \quad :\text{iff} \quad \text{there is some } p \in P \text{ with } a, b \in p .$$

Definition 4.13. Let A be a set and $R \subseteq A \times A$ a binary relation on A. The *transitive closure* of R is the relation

$$\text{TC}(R) := \bigcap \{ S \subseteq A \times A \mid S \supseteq R \text{ is transitive} \} .$$

Since the family of transitive relations is closed under intersections we can use Lemma 4.8 (b) to prove that TC is a closure operator.

Lemma 4.14. *Let A be a class.* TC *is a closure operator on $A \times A$.*

Exercise 4.4. Prove Lemma 4.14.

Lemma 4.15. *If $R \subseteq A \times A$ is a symmetric relation then so is* $\text{TC}(R)$.

Proof. Let $S := \text{TC}(R) \cap (\text{TC}(R))^{-1}$. Since R is symmetric we have $R \subseteq S$. We claim that S is transitive.

Let $\langle a, b\rangle, \langle b, c\rangle \in S$. Then $\langle a, b\rangle, \langle b, c\rangle \in \text{TC}(R)$ and $\langle b, a\rangle, \langle c, b\rangle \in \text{TC}(R)$. Therefore, we have $\langle a, c\rangle \in \text{TC}(R)$ and $\langle c, a\rangle \in \text{TC}(R)$. This implies that $\langle a, c\rangle \in S$, as desired.

We have shown that S is a transitive relation containing R. By the definition of TC it follows that $\text{TC}(R) \subseteq S = \text{TC}(R) \cap \text{TC}(R)^{-1}$. This implies that $\text{TC}(R)^{-1} = \text{TC}(R)$. Hence, $\text{TC}(R)$ is symmetric. □

Lemma 4.16. *Let $R \subseteq A \times A$ be a binary relation.*

(a) *The smallest reflexive relation containing R is* $R \cup \text{id}_A$.

(b) *The smallest symmetric relation containing R is $R \cup R^{-1}$.*

(c) *The smallest transitive relation containing R is $\mathrm{TC}(R)$.*

(d) *The smallest equivalence relation containing R is $\mathrm{TC}(R \cup R^{-1} \cup \mathrm{id}_A)$.*

Proof. (a) $R \cup \mathrm{id}_A$ is obviously reflexive and it contains R. Conversely, suppose that $S \supseteq R$ is reflexive. Then $\mathrm{id}_A \subseteq S$ implies that $R \cup \mathrm{id}_A \subseteq S$.

(b) is proved analogously.

(c) Let $S \supseteq R$ be transitive. Then the intersection in the definition of TC contains S. Hence, $\mathrm{TC}(R) \subseteq S$. Furthermore, we have $R \subseteq \mathrm{TC}(R)$ by definition. It remains to prove that $\mathrm{TC}(R)$ is transitive.

Let $\langle a, b \rangle, \langle b, c \rangle \in \mathrm{TC}(R)$. Then we have $\langle a, b \rangle, \langle b, c \rangle \in S$, for every transitive relation $S \supseteq R$. Hence, we have $\langle a, c \rangle \in S$, for each such relation S. This implies that $\langle a, c \rangle \in \mathrm{TC}(R)$.

(d) Set $E := \mathrm{TC}(R \cup R^{-1} \cup \mathrm{id}_A)$. Clearly, we have $R \subseteq E$ and, if $S \supseteq R$ is an equivalence relation then $E \subseteq S$. Hence, it is remains to prove that E is an equivalence relation. It is transitive by (c), symmetric by Lemma 4.15, and E is reflexive since $\mathrm{id}_A \subseteq \mathrm{TC}(R \cup R^{-1} \cup \mathrm{id}_A)$. ☐

Corollary 4.17. *Let A be a set and $F \subseteq \wp(A \times A)$ the set of all equivalence relations on A. Then $\langle F, \subseteq \rangle$ forms a complete partial order. If $X \subseteq F$ is nonempty then we have*

$$\inf X = \bigcap X \quad \textit{and} \quad \sup X = \mathrm{TC}\big(\bigcup X\big)\,.$$

Proof. By Lemma 4.16, we have $F = \mathrm{fix}\, c$ where c is the closure operator with

$$c(R) := \mathrm{TC}(R \cup R^{-1} \cup \mathrm{id}_A)\,.$$

The relation $E := \bigcup X$ is reflexive and symmetric since X is nonempty. Hence, we have $\mathrm{TC}(E \cup E^{-1} \cup \mathrm{id}_A) = \mathrm{TC}(E)$. Consequently, the claim follows from Theorem 4.9. ☐

A3. *Ordinals*

1. *Well-orders*

When defining stages we frequently used the fact that any class of stages has a minimal element. In this section we study arbitrary orders with this property.

Definition 1.1. Let $\langle A, R\rangle$ be a graph.

(a) An element $a \in A$ is *R-minimal* if $\langle b, a\rangle \in R$ implies $b = a$.

(b) A relation R is *left-narrow* if $R^{-1}(a)$ is a set, for every set $a \in \operatorname{rng} R$.

(c) R is *well-founded* if every nonempty subset $B \subseteq A$ contains an R-minimal element. A left-narrow, well-founded linear order is called a *well-order.*

Example. (a) $\langle \mathbb{N}, \leq\rangle$ is a well-order.

(b) $\langle \mathbb{N}, \mid\ \rangle$ is a well-founded partial order.

(c) The membership relation $\in$ is a well-founded partial order on $\mathbb{S}$. It is a well-order on the class of all stages.

(d) $\langle \wp(\mathbb{N}), \subseteq\rangle$ is not well-founded.

(e) A partial order $\langle A, \leq\rangle$ is left-narrow if, and only if, $\Downarrow a$ is a set, for all $a \in A$.

Exercise 1.1. Prove that $\langle \wp(\mathbb{N}), \subseteq\rangle$ is not well-founded.

Lemma 1.2. *If $\langle A, R\rangle$ is a well-founded graph and $B \subseteq A$ then $\langle B, R|_B\rangle$ is also well-founded.*

Proof. Every nonempty subset $C \subseteq B$ is also a nonempty subset of A and has an R-minimal element. □

Lemma 1.3. *If $\langle A, \leq \rangle$ is a well-founded and left-narrow partial order, there exists no infinite sequence $(a_n)_{n \in \mathbb{N}} \in A^{\mathbb{N}}$ such that $a_n \neq a_{n+1}$ and $a_{n+1} \leq a_n$, for all n.*

Proof. If there exists such an infinite sequence then the class $\operatorname{rng} \bar{a} = \{ a_n \mid n \in \mathbb{N} \}$ is nonempty and has no $\leq$-minimal element. Furthermore, $\operatorname{rng} \bar{a} \subseteq \Downarrow a_0$ is a set since the order is left-narrow. □

The reason why well-founded relations are of interest is that these are exactly those relations that admit proofs by induction. As the theorem below shows we can prove that every element of a well-founded partial order $\langle A, \leq \rangle$ satisfies a given property φ by showing that, if every element $b < a$ satisfies φ then a also satisfies φ.

Lemma 1.4. *Let $\langle A, \leq \rangle$ be a well-founded, left-narrow partial order. Every nonempty subclass $X \subseteq A$ has a minimal element.*

Proof. Let $X \subseteq A$ be nonempty and fix some element $a \in X$. $\Downarrow a$ is a set since $\leq$ is left-narrow. Hence, $Y := X \cap \Downarrow a$ is a nonempty subset of A and has a minimal element b. Note that $b \in Y \subseteq X$ and, if $c \in X$ is some element with $c \leq b \leq a$, then $c \in Y$. Therefore, it follows that b is also a minimal element of X. □

Theorem 1.5. *Let $\langle A, \leq \rangle$ be a well-founded, left-narrow partial order. If $X \subseteq A$ is a subclass such that*

$$\downarrow a \subseteq X \quad \textit{implies} \quad a \in X\,, \qquad \textit{for all } a \in A\,,$$

then $X = A$.

Proof. Let $X \subseteq A$ be a class as above. For a contradiction, suppose that $X \neq A$. Fix some element $a \in A \setminus X$. Since $\leq$ is left-narrow $B := \Downarrow a \setminus X$ is a set. Hence, B has a $\leq$-minimal element b. It follows that $\downarrow b \subseteq A \setminus B \subseteq X$, which implies that $b \in X$. Contradiction. □

Example. Consider the well-order $\langle \mathbb{N}, < \rangle$ of the natural numbers. Suppose that $X \subseteq \mathbb{N}$ is a subset such that we can show that

$$b \in X \,, \text{ for all } b < a \,, \quad \text{implies} \quad a \in X \,,$$

then we have $X = \mathbb{N}$. Proofs based on this fact are called 'proofs by induction'. The above corollary states that such proofs work not only for the natural numbers but for all well-orders.

Let $\langle A, \leq \rangle$ be a well-order. The minimal element of a given subclass $X \subseteq A$ is unique since A is linearly ordered. Therefore, if A is not empty, it has a least element $\perp$. The *successor* a^+ of an element $a \in A$ is the least element of the class $\uparrow a$. a^+ is defined for every element of A except for the greatest one. An element that is neither the least one nor a successor of some other element is called a *limit*.

It turns out that we can define a canonical well-founded order on the class Wo of all well-orders.

Remark. Note that speaking of 'the class of all well-orders' is sloppy language since, by definition, a class contains only sets. Instead, we should call Wo 'the class of all well-orders that are sets'.

Definition 1.6. Let $\mathfrak{A} = \langle A, \leq_A \rangle$ and $\mathfrak{B} = \langle B, \leq_B \rangle$ be well-orders. We define

$$\mathfrak{A} < \mathfrak{B} \quad :\text{iff} \quad A \text{ is a set and, for some } b \in B \,, \text{ there exists an isomorphism } f : A \to {\downarrow_B} b \,.$$

(Note that, if f exists, it is necessarily a set because A and ${\downarrow_B} b$ are both sets.)

To prove that this defines an order on Wo we need some technical lemmas.

Lemma 1.7. *Let $\langle A, \leq \rangle$ be a well-order. If $f : A \to A$ is a strictly increasing function then $a \leq f(a)$, for all $a \in A$.*

Proof. Suppose that there exists some $a \in A$ with $a > f(a)$. Let a_0 be the minimal such element. By minimality of a_0 we have

$$f(a_0) \leq f(f(a_0)) .$$

On the other hand, since f is strictly increasing we have

$$f(f(a_0)) < f(a_0) .$$

Contradiction. □

Lemma 1.8. *Let $\langle A, \leq \rangle$ be a well-order and $I \subseteq A$. The following statements are equivalent:*

(1) *I is a proper initial segment of A.*

(2) *$I = {\downarrow_A} a$, for some $a \in A$.*

(3) *I is an initial segment of A and I is non-isomorphic to A.*

Proof. (1) ⇒ (2) If I is a proper subclass of A then $A \setminus I$ is nonempty and has a least element a. Consequently, we have $I = {\downarrow} a$.

(2) ⇒ (3) Let $I = {\downarrow} a$. Suppose there exists an isomorphism $f : A \to I$. By Lemma 1.7, we have $f(a) \geq a$. Hence, $f(a) \notin I = \operatorname{rng} f$. Contradiction.

(3) ⇒ (1) is trivial. □

Corollary 1.9. *$<$ is a strict partial order on* Wo.

Proof. We can see immediately from the definition that $<$ is transitive. Suppose that $\mathfrak{A} < \mathfrak{A}$, for some well-order $\mathfrak{A} = \langle A, \leq \rangle$. By definition there exists an element $a \in A$ and an isomorphism $f : A \to {\downarrow_A} a$. This contradicts the preceding lemma. □

Lemma 1.10. *Let $\langle A, \leq_A \rangle$ and $\langle B, \leq_B \rangle$ be well-orders. There exists at most one isomorphism $f : A \to B$.*

Proof. Let $f, g : A \to B$ be isomorphisms. Then so is $g \circ f^{-1} : B \to B$. In particular, $g \circ f^{-1}$ is strictly increasing. By Lemma 1.7, we obtain

$$f(a) \leq (g \circ f^{-1})(f(a)) = g(a), \qquad \text{for all } a \in A.$$

Similarly, we derive $g(a) \leq f(a)$, for all a. It follows that $f = g$. ☐

We still have to prove that $<$ is linear. Unfortunately, this is not true. The following theorem states that, for all well-orders $\mathfrak{A}$ and $\mathfrak{B}$, exactly one of the following conditions holds $\mathfrak{A} < \mathfrak{B}$ or $\mathfrak{A} \cong \mathfrak{B}$ or $\mathfrak{A} > \mathfrak{B}$. In order for $<$ to be linear, the second condition should read $\mathfrak{A} = \mathfrak{B}$. We will see how to deal with this problem in the next section.

Theorem 1.11. *Let $\langle A, \leq_A \rangle$ and $\langle B, \leq_B \rangle$ be well-orders. Exactly one of the following statements holds:*

(1) *There exists an isomorphism $f : A \to J$ where $J \subset B$ is a proper initial segment of B.*

(2) *There exists an isomorphism $f : A \to B$.*

(3) *There exists an isomorphism $f : I \to B$ where $I \subset A$ is a proper initial segment of A.*

(*f might be a proper class.*)

Proof. We claim that

$$f := \{ \langle a, b \rangle \in A \times B \mid \text{there is an isomorphism } {\downarrow}a \to {\downarrow}b \}.$$

is the desired isomorphism.

First, we show that $\langle a_0, b_0 \rangle, \langle a_1, b_1 \rangle \in f$ implies

$$a_0 < a_1 \quad \text{iff} \quad b_0 < b_1.$$

For a contradiction, suppose that $a_0 < a_1$ and $b_0 \geq b_1$. We have isomorphisms

$$h_0 : {\downarrow}a_0 \to {\downarrow}b_0 \quad \text{and} \quad h_1 : {\downarrow}a_1 \to {\downarrow}b_1.$$

The restriction of h_1 to $\downarrow a_0$ is an isomorphism

$$h_1 \upharpoonright \downarrow a_0 : \downarrow a_0 \to \downarrow h_1(a_0) .$$

Composing it with h_0^{-1} yields an isomorphism

$$(h_1 \upharpoonright \downarrow a_0) \circ h_0^{-1} : \downarrow b_0 \to \downarrow h_1(a_0) .$$

But this contradicts $h_1(a_0) < b_1 \leq b_0$, by Lemma 1.8.

Therefore, f is the graph of a strictly increasing function. We claim that dom f and rng f are initial segments of, respectively, A and B. Suppose, for a contradiction, that there are elements $a < b$ such that $a \notin \operatorname{dom} f$ and $b \in \operatorname{dom} f$. By definition, there is an isomorphism $h : \downarrow b \to \downarrow f(b)$. Its restriction to $\downarrow a$ yields an isomorphism $h \upharpoonright \downarrow a : \downarrow a \to \downarrow h(a)$ which shows that $a \in \operatorname{dom} f$. Contradiction. Analogously, it follows that rng f is an initial subclass of B.

It remains to show that $\operatorname{dom} f = A$ or $\operatorname{rng} f = B$. Suppose, otherwise. Let a be the minimal element of $A \setminus \operatorname{dom} f$ and b the minimal one of $B \setminus \operatorname{rng} f$. Then $\operatorname{dom} f = \downarrow a$ and $\operatorname{rng} f = \downarrow b$ and f is an isomorphism from $\downarrow a$ to $\downarrow b$. By definition, we therefore have $\langle a, b \rangle \in f$. Contradiction. □

Corollary 1.12. *For all* $\mathfrak{A}, \mathfrak{B} \in \mathrm{Wo}$, *we have either*

$$\mathfrak{A} < \mathfrak{B} \quad \text{or} \quad \mathfrak{A} \cong \mathfrak{B} \quad \text{or} \quad \mathfrak{A} > \mathfrak{B} .$$

We conclude this section with two remarks about continuous mappings between well-orders. The following lemma provides a simple criterion to check whether a mapping between well-orders is continuous.

Lemma 1.13. *Let* $\langle A, \leq \rangle$ *be a well-order and* $\langle B, \leq \rangle$ *an arbitrary partial order. A function* $f : A \to B$ *is continuous if, and only if, it satisfies the following conditions:*

(1) $f(a^+) \geq f(a)$, *for all* $a \in A$,

(2) $f(a) = \sup \{ f(b) \mid b < a \}$, *for every limit* $a \in A$.

Proof. ($\Rightarrow$) By definition, every continuous function satisfies (2). Furthermore, $a^+ = \sup\{a, a^+\}$ implies that $f(a^+) = \sup\{f(a), f(a^+)\}$.

($\Leftarrow$) For the other direction, suppose that f satisfies (1) and (2). First, we show that f is increasing. Suppose otherwise and let $a \in A$ be the minimal element such that $f(b) > f(a)$, for some $b < a$. Note that a is not the minimal element of A since $b < a$. If a were a limit then (2) would imply that

$$f(a) = \sup\{\, f(x) \mid x < a \,\} \geq f(b)\,.$$

Contradiction. Hence, a must be a successor and we have $a = c^+$, for some $c \in A$. By choice of a, we have $f(x) \leq f(c)$, for all $x \leq c$. In particular, $f(c) \geq f(b) > f(a)$. But (1) implies $f(a) = f(c^+) \geq f(c)$. Again a contradiction.

We have shown that f is increasing. But what we really want to prove is that it is continuous. Let $X \subseteq A$ be a nonempty subset of A with supremum $a := \sup X$. If $b \in X$ then $b \leq a$ implies $f(b) \leq f(a)$. Hence, $f(a)$ is an upper bound of $f[X]$. To prove that $f(a)$ is its least upper bound we distinguish two cases.

If $a \in X$ then $f(a) \in f[X]$, which implies $f(a) = \sup f[X]$.

If $a \notin X$ then $a = \sup X$ is a limit and, for every $b < a$, there is some $x \in X$ with $b \leq x$. If c is another upper bound of $f[X]$ then $f(b) \leq f(x) \leq c$. By (2), it follows that

$$f(a) = \sup\{\, f(b) \mid b < a \,\} \leq \sup\{\, f(x) \mid x \in X \,\} \leq c\,.$$

Hence, $f(a)$ is the least upper bound of $f[X]$. □

Lemma 1.14. *Let $\langle A, \leq\rangle$ be a well-order and $f : A \to A$ strictly continuous. If $a \geq f(\bot)$ then*

$$\max\{\, b \in A \mid f(b) \leq a \,\}$$

exists.

Proof. If a is the greatest element of A, we can set $b := a$. Otherwise, we have $f(a^+) > f(a) \geq a$, by Lemma 1.7. Hence, there are elements $x \in A$ with $f(x) > a$. Let c be the least such element. We have $c > \bot$ since $f(c) > a \geq f(\bot)$. If c were a limit then, by choice of c, we would have

$$f(c) = \sup\{ f(x) \mid x < c \} \leq a < f(c).$$

A contradiction. Hence, c is a successor and there exists some $b \in A$ with $c = b^+$. By choice of c, we have $f(b) \leq a$. Furthermore, if $x > b$ then $x \geq c$, which implies that $f(x) \geq f(c) > a$. Therefore, b is the desired element. □

2. *Ordinals*

We have seen that there exists a well-order on Wo if one does not distinguish between isomorphic orders. We would like to define a subclass On $\subseteq$ Wo of *ordinals* such that, for each well-order $\mathfrak{A}$, there exists a unique element $\mathfrak{B} \in$ On that is isomorphic to $\mathfrak{A}$.

We will present two approaches to do so. The usual one – due to von Neumann – has the disadvantage that it requires the Axiom of Replacement. Without it we cannot prove that, for every well-order α, there exists an isomorphic von Neumann ordinal. Therefore, we will adopt a different approach. The relation $\cong$ forms a congruence (see Section B1.4 below) on the class of all well-orders. A first try might thus consist in representing a well-ordering by its congruence class. Unfortunately, the class of all well-orders isomorphic to a given one is not a set. Hence, with this definition one could not form sets of ordinals. Instead of considering *all* isomorphic well-orders we will therefore only take some of them.

Definition 2.1. The *order type* of a well-order $\mathfrak{A}$ is the set

$$\operatorname{ord}(\mathfrak{A}) := [\mathfrak{A}]_{\cong} = \operatorname{cut}\{ \mathfrak{B} \mid \mathfrak{B} \text{ is a well-order isomorphic to } \mathfrak{A} \}.$$

The elements of On := rng(ord) are called *ordinals*.

Instead of a subclass $\mathrm{On} \subseteq \mathrm{Wo}$ the above definition results in a function $\mathrm{ord} : \mathrm{Wo} \to \mathrm{On}$. Below we will see that there exists a canonical way to associate with every ordinal $\alpha \in \mathrm{On}$ a well-order $f(\alpha) \in \mathrm{Wo}$. Using this injection $f : \mathrm{On} \to \mathrm{Wo}$ we can identify the class On with its image $f[\mathrm{On}] \subseteq \mathrm{Wo}$.

First, let us show that the mapping $\mathrm{ord} : \mathrm{Wo} \to \mathrm{On}$ has the desired property of characterising a well-order up to isomorphism.

Lemma 2.2. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be well-orders that are sets. There exists an isomorphism $f : \mathfrak{A} \to \mathfrak{B}$ if, and only if,* $\mathrm{ord}(\mathfrak{A}) = \mathrm{ord}(\mathfrak{B})$.

Proof. If $f : \mathfrak{A} \to \mathfrak{B}$ is an isomorphism then a well-order $\mathfrak{C}$ is isomorphic to $\mathfrak{A}$ if, and only if, it is isomorphic to $\mathfrak{B}$. Therefore $\mathrm{ord}(\mathfrak{A}) = \mathrm{ord}(\mathfrak{B})$. Conversely, suppose $\mathrm{ord}(\mathfrak{A}) = \mathrm{ord}(\mathfrak{B})$. Since $\mathfrak{A}$ is a well-order isomorphic to $\mathfrak{A}$, we have $\mathrm{ord}(\mathfrak{A}) \neq \varnothing$. Fix an arbitrary element $\mathfrak{C} \in \mathrm{ord}(\mathfrak{A})$. By definition, $\mathfrak{C}$ is isomorphic to $\mathfrak{A}$ and to $\mathfrak{B}$. Consequently, $\mathfrak{A}$ and $\mathfrak{B}$ are isomorphic. □

Remark. We will prove in Lemma A4.5.3 with the help of the Axiom of Replacement that any two well-ordered proper classes are isomorphic. In particular, it follows that in the above lemma we can drop the requirement of $\mathfrak{A}$ and $\mathfrak{B}$ being sets.

Definition 2.3. Let $\mathfrak{On} := \langle \mathrm{On}, < \rangle$ where the ordering $<$ is defined by

$$\mathrm{ord}(\mathfrak{A}) < \mathrm{ord}(\mathfrak{B}) \quad :\text{iff} \quad \mathfrak{A} < \mathfrak{B}\,.$$

For $\alpha \in \mathrm{On}$, recall that $\downarrow\alpha = \{\, \beta \in \mathrm{On} \mid \beta < \alpha \,\}$.

Remark. (a) The ordering $<$ is well-defined since $\mathrm{ord}(\mathfrak{A}) = \mathrm{ord}(\mathfrak{A}')$ and $\mathrm{ord}(\mathfrak{B}) = \mathrm{ord}(\mathfrak{B}')$ implies that $\mathfrak{A} < \mathfrak{B}$ iff $\mathfrak{A}' < \mathfrak{B}'$.

(b) In the chapters on set theory we will strictly distinguish between an ordinal α and the set $\downarrow\alpha$. But in the remainder of the book we will usually drop the arrow and write α in both cases.

Combining Corollaries 1.9 and 1.12 and Lemma 2.2 it follows that On is well-ordered.

Theorem 2.4. $\mathfrak{On}$ *is a well-order.*

The notions of a *successor ordinal* and a *limit ordinal* are defined in the same way as for arbitrary well-orders. Recall that we denote the successor of α by α^+. Furthermore, we define

$$0 := \operatorname{ord}\langle\varnothing,\varnothing\rangle\,,\quad 1 := 0^+\,,\quad 2 := 1^+\,,\dots$$

The first limit ordinal is $\omega := \operatorname{ord}\langle\mathbb{N},\leq\rangle$.

Lemma 2.5. *Let* $\alpha,\beta \in \mathrm{On}$. *If* $\alpha \leq \beta$ *then* $S(\alpha) \subseteq S(\beta)$.

Proof. If $\alpha = \beta$, the claim is trivial. Therefore, we assume that $\alpha < \beta$. Let $\mathfrak{A} = \langle A, \leq_A\rangle \in \alpha$ and $\mathfrak{B} = \langle B, \leq_B\rangle \in \beta$. Since $\alpha < \beta$ there exists an isomorphism $f : A \to {\downarrow_B} b$, for some $b \in B$. Set $\mathfrak{B}_0 := \langle {\downarrow_B} b, \leq_B\rangle$. Then $\operatorname{ord}\mathfrak{B}_0 = \alpha$ and $\mathfrak{A} \in \operatorname{ord}\mathfrak{B}_0$ implies that $S(\mathfrak{A}) \subseteq S(\mathfrak{B}_0)$. Since $S(\mathfrak{B}_0) \subseteq S(\mathfrak{B})$ it follows that $S(\mathfrak{A}) \subseteq S(\mathfrak{B})$. We have shown that $S(x) \subseteq S(y)$, for all $x \in \alpha$ and $y \in \beta$. Consequently, we have $S(\alpha) \subseteq S(\beta)$. □

To every ordinal α we can associate a canonical well-order of type α.

Lemma 2.6. $\langle{\downarrow}\alpha,\leq\rangle$ *is a well-order of type* $\operatorname{ord}\langle{\downarrow}\alpha,\leq\rangle = \alpha$.

Proof. Let $\langle A,\leq\rangle$ be a well-order of type $\operatorname{ord}\langle A,\leq\rangle = \alpha$. We claim that the function $f : A \to \mathrm{On}$ with

$$f(a) := \operatorname{ord}\langle {\downarrow_A} a, \leq\rangle$$

is an isomorphism $f : A \to {\downarrow}\alpha$.

f is strictly increasing since, if $a < b$ then ${\downarrow_A} a$ is a proper initial segment of ${\downarrow_A} b$. By Lemma 1.8 and Lemma 2.2, it follows that

$$f(a) = \operatorname{ord}\langle {\downarrow_A} a, \leq\rangle < \operatorname{ord}\langle {\downarrow_A} b, \leq\rangle = f(b)\,.$$

Furthermore, f is surjective since, for every $\beta < \alpha$, there exists some $a \in A$ with

$$\beta = \operatorname{ord}\langle {\downarrow_A} a, \leq\rangle = f(a)\,.$$

□

Lemma 2.7. On *is not a set.*

Proof. Suppose that On is a set. Since On is well-ordered there exists some ordinal $\alpha \in \mathrm{On}$ with $\alpha = \mathrm{ord}\langle \mathrm{On}, \leq \rangle$. We have just seen that $\mathrm{ord}\langle \downarrow\alpha, \leq \rangle = \alpha$. Therefore, there exists an isomorphism $f : \downarrow\alpha \to \mathrm{On}$. But $\downarrow\alpha$ is a proper initial segment of On. This contradicts Lemma 1.8. □

Lemma 2.8. *A subclass $X \subseteq \mathrm{On}$ is a set if, and only if, it has an upper bound.*

Proof. ($\Leftarrow$) If $X \subseteq \mathrm{On}$ has an upper bound α then $X \subseteq \Downarrow\alpha$. Since $\Downarrow\alpha$ is a set the claim follows.

($\Rightarrow$) Suppose that X is a set. Since On is a proper class there exists some ordinal $\alpha \in \mathrm{On} \setminus S(X)$. We claim that α is an upper bound of X. Suppose there exists some $\beta \in X$ with $\beta \nleq \alpha$. Then $\alpha < \beta$ and we have $\alpha \subseteq S(\alpha) \subseteq S(\beta) \in S(X)$, which implies that $\alpha \in S(X)$. This contradicts our choice of α. □

Corollary 2.9. *Every set of ordinals has a supremum.*

Another consequence is the following special case of the Axiom of Replacement which we will introduce in Section A4.5.

Corollary 2.10. *If $F : \mathrm{On} \to \mathrm{On}$ is increasing then $F[\downarrow\alpha]$ is a set, for all $\alpha \in \mathrm{On}$.*

Proof. Suppose that F is increasing. Then we have $F(\beta) \leq F(\alpha)$, for all $\beta < \alpha$. Consequently, $F(\alpha)$ is an upper bound of $F[\downarrow\alpha]$ and, by Lemma 2.8, it follows that $F[\downarrow\alpha]$ is a set. □

Let us give a simpler characterisation of the relation ≤ on well-orders.

Lemma 2.11. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be well-orders. Then $\mathfrak{A} \leq \mathfrak{B}$ if, and only if, there exists a strictly increasing function $f : A \to B$.*

Proof. ($\Rightarrow$) If $\mathfrak{A} \leq \mathfrak{B}$ then, by definition, there exists an isomorphism $f : A \to I$ between A and an initial segment I of B. In particular, $f : A \to B$ is a strictly increasing function.

($\Leftarrow$) Suppose that $f : A \to B$ is a strictly increasing function and let $C := \operatorname{rng} f$. Since $C \subseteq B$ is well-ordered there exists an isomorphism $g : C \to I \subseteq \mathrm{On}$ between C and an initial segment of On. Similarly, there is some isomorphism $h : B \to J \subseteq \mathrm{On}$. We claim that

$$k := h^{-1} \circ g \circ f : A \to B$$

is the desired isomorphism between A and an initial segment of B. Since f, g, and h^{-1} are isomorphisms so is k. What remains to be shown is that k is in fact well-defined, that is, $I = \operatorname{rng} g \subseteq \operatorname{rng} h = J$.

We claim that $g(c) \leq h(c)$, for all $c \in C$. Since I and J are initial segments this implies that $I \subseteq J$. For a contradiction, suppose that there is some $c \in C$ with $g(c) > h(c)$ and let c be the minimal such element. Note that, since g and h are strictly increasing and $\operatorname{rng} g$ and $\operatorname{rng} h$ are initial segments we must have

$$\begin{aligned} & g(c) = \min\left(I \setminus \operatorname{rng}\left(g \restriction {\downarrow_C} c\right)\right) \\ \text{and} \quad & h(c) = \min\left(J \setminus \operatorname{rng}\left(h \restriction {\downarrow_B} c\right)\right). \end{aligned}$$

By choice of c, we have $\operatorname{rng}\left(g \restriction {\downarrow_C} c\right) \subseteq \operatorname{rng}\left(h \restriction {\downarrow_B} c\right)$. But, by the above equations, this implies that $g(c) \leq h(c)$. A contradiction. □

In order to use the theory of ordinals for proofs about arbitrary sets one usually needs to define a well-order on a given set. In general this is only possible if one assumes the Axiom of Choice. Until we introduce this axiom the following theorem will serve as a stopgap. Once we have defined the cardinality of a set in Section A4.2 it will turn out that the ordinal the theorem talks about is $\alpha = |A|^+$.

Theorem 2.12 (Hartogs). *For every set A there exists an ordinal α such that there are no injective functions ${\downarrow}\alpha \to A$.*

Proof. For a contradiction, suppose that there exists a set A such that, for every ordinal α, there is an injective function $f_\alpha : {\downarrow}\alpha \to A$. Let $A_\alpha := \operatorname{rng} f_\alpha \subseteq A$ and set

$$R_\alpha := \left\{ \langle a, b \rangle \in A_\alpha \times A_\alpha \mid f_\alpha^{-1}(a) \leq f_\alpha^{-1}(b) \right\}.$$

By construction, $f_\alpha : \langle{\downarrow}\alpha, \leq\rangle \to \langle A_\alpha, R_\alpha\rangle$ is an isomorphism. Hence, by the definition of an ordinal, we have

$$S(\alpha) \subseteq S(\langle A_\alpha, R_\alpha\rangle)\,.$$

Since $R_\alpha \subseteq A \times A \in \wp^3(A) \subseteq \wp^3(S(A))$ it follows that

$$\langle A_\alpha, R_\alpha\rangle = \big\{\{A_\alpha\},\ \{A_\alpha, R_\alpha\}\big\} \subseteq \wp^4(S(A))\,.$$

We have shown that

$$\alpha \subseteq S(\alpha) \subseteq S(\langle A_\alpha, R_\alpha\rangle) \subseteq \wp^4(S(A))\,, \quad \text{for all } \alpha \in \mathrm{On}\,.$$

Consequently, $\mathrm{On} \subseteq \wp^5(S(A))$, which implies that On is a set. This contradicts Lemma 2.7. □

Von Neumann ordinals

We conclude this section with an alternative definition of ordinals. This definition is simpler and the resulting ordinals have many nice properties such that $\alpha = {\downarrow}\alpha$ and $\sup X = \bigcup X$. The only disadvantage is that one needs an additional axiom in order to prove that every well-order is isomorphic to some ordinal. Intuitively, we define a *von Neumann ordinal* to be the set of all smaller ordinals, that is, $\alpha := {\downarrow}\alpha$. As usual, the actual definition is more technical and we have to verify afterwards that it has the desired effect.

Definition 2.13. A set α is a *von Neumann ordinal* if it is transitive and linearly ordered by the membership relation $\in$. We denote the class of all von Neumann ordinals by On_0 and we set $\mathfrak{On}_0 := \langle \mathrm{On}_0, \in\rangle$.

Example. The set $[n] = \{[0], \dots, [n-1]\}$ is a von Neumann ordinal, for each $n \in \mathbb{N}$.

Lemma 2.14. *If $\alpha \in \mathrm{On}_0$ and $\beta \in \alpha$ then $\beta \in \mathrm{On}_0$.*

Proof. First, note that $\beta \in \alpha$ implies $\beta \subseteq \alpha$. As α is linearly ordered by $\in$ it therefore follows that so is $\beta \subseteq \alpha$.

It remains to prove that β is transitive. Suppose that $\eta \in \gamma \in \beta$. By transitivity of α, we have $\eta, \gamma, \beta \in \alpha$. Since α is linearly ordered by $\in$ we know that the relation $\in$, restricted to α, is transitive. Hence, $\eta \in \gamma$ and $\gamma \in \beta$ implies that $\eta \in \beta$. □

Remark. Note that, for $\alpha \in \mathrm{On}_0$, we have

$$\downarrow\alpha = \{\, \beta \in \mathrm{On}_0 \mid \beta \in \alpha \,\}\,.$$

Hence, $\alpha = \downarrow\alpha$ and our definition of a von Neumann ordinal coincides with the intuitive one.

Exercise 2.1. Suppose that $\alpha = \{\beta_0, \dots, \beta_{n-1}\}$ is a von Neumann ordinal with $n < \omega$ elements. Prove, by induction on n, that $\alpha = [n]$.

Theorem 2.15. $\mathfrak{On}_0$ *is a well-order.*

Proof. $\in$ is irreflexive since we have $a \notin a$, for all sets. Furthermore, $\in$ is transitive on On_0 since, $\alpha \in \beta \in \gamma$ implies $\alpha \in \gamma$, by transitivity of γ. Consequently, $\in$ is a strict partial order on On_0. Since $\in$ is well-founded on any class it remains to prove that it is linear.

Let $\alpha, \beta \in \mathrm{On}_0$. The set $\gamma := \alpha \cap \beta$ is transitive by Lemma A1.2.4. As α is linearly ordered by $\in$ so is $\gamma \subseteq \alpha$. Therefore, $\gamma \in \mathrm{On}_0$. Furthermore, γ is an initial segment of α since $\delta \in \eta \in \gamma$ implies $\delta \in \gamma$, by transitivity. By Lemma 1.8, it follows that $\gamma = \alpha$ or $\gamma = \downarrow\delta = \delta$, for some $\delta \in \alpha$. Hence, we either have $\gamma = \alpha$ or $\gamma \in \alpha$. Similarly, we can prove that either $\gamma = \beta$ or $\gamma \in \beta$. Since $\gamma \notin \gamma = \alpha \cap \beta$ it follows that either $\gamma \notin \alpha$ or $\gamma \notin \beta$. Consequently, we either have $\beta = \gamma \in \alpha$ or $\alpha = \gamma \in \beta$ or $\alpha = \gamma = \beta$. □

Exercise 2.2. Show that $\alpha^+ = \alpha \cup \{\alpha\}$, for every $\alpha \in \mathrm{On}_0$.

Lemma 2.16. On_0 *is not a set.*

Proof. On_0 is transitive and well-ordered by $\in$. If it were a set, it would be an element of itself. □

On_0 is linearly ordered by $\in$. The following sequence of lemmas contains several characterisations of this ordering. In particular, we show that the mapping

$$\mathrm{ord} : \langle \mathrm{On}_0, \in \rangle \to \langle \mathrm{On}, < \rangle$$

is strictly increasing. After we have introduced the Axiom of Replacement in Section A4.5 we will prove that it is actually an isomorphism.

Lemma 2.17. *Let $\alpha, \beta \in \mathrm{On}_0$. We have $\alpha \in \beta$ if, and only if, $\alpha \subset \beta$.*

Proof. $(\Rightarrow)$ was already proved in Lemma A1.2.2. For $(\Leftarrow)$, suppose that $\alpha \notin \beta$. By Lemma 2.15, it follows that $\alpha = \beta$ or $\beta \in \alpha$. Since $\alpha \subset \beta$ we therefore have $\beta \subset \beta$ or $\beta \in \beta$. Contradiction. □

Lemma 2.18. *Let $\alpha, \beta \in \mathrm{On}_0$. If $f : \alpha \to \beta$ is an isomorphism between α and an initial segment of β then $f = \mathrm{id}_\alpha$.*

Proof. Suppose that $f \neq \mathrm{id}_\alpha$ and let $\gamma \in \alpha$ be the minimal element of α such that $\delta := f(\gamma) \neq \gamma$. Since f is an isomorphism we have $\xi = f(\xi) \in f(\gamma) = \delta$, for all $\xi \in \gamma$. Hence, $\gamma \subseteq \delta$. Since $\delta \neq \gamma$ it follows that $\gamma \subset \delta$, which implies, by Lemma 2.17, that $\gamma \in \delta$. But $\gamma \notin \mathrm{rng}\, f$, since $f(\xi) = \xi \in \gamma$, for $\xi \in \gamma$, and $f(\xi) \ni f(\gamma) = \delta$, for $\xi \ni \gamma$. Hence, $\mathrm{rng}\, f$ is not an initial segment of β. Contradiction. □

Lemma 2.19. *Let $\alpha, \beta \in \mathrm{On}_0$. The following statements are equivalent:*

(1) $\alpha \in \beta$.

(2) $\alpha \subset \beta$.

(3) $S(\alpha) \in S(\beta)$.

(4) $\langle \alpha, \in \rangle < \langle \beta, \in \rangle$.

Proof. (1) $\Leftrightarrow$ (2) was already shown in Lemma 2.17.

(1) $\Rightarrow$ (3) $a \in b$ implies $S(a) \in S(b)$, for arbitrary sets a and b.

(3) $\Rightarrow$ (1) If $\alpha \notin \beta$ then, by Lemma 2.15, we either have $\alpha = \beta$ or $\beta \in \alpha$. Consequently, either $S(\alpha) = S(\beta)$ or $S(\beta) \in S(\alpha)$. It follows that $S(\alpha) \notin S(\beta)$.

(2) ⇒ (4) If $\alpha \subseteq \beta$, the identity $\mathrm{id}_\alpha : \alpha \to \alpha \subseteq \beta$ is an isomorphism from α to an initial segment of β. Hence, $\alpha < \beta$.

(4) ⇒ (2) If $\alpha < \beta$, there exists an isomorphism $f : \alpha \to I \subset \beta$ between α and a proper initial subset of β. By the preceding lemma, it follows that $f = \mathrm{id}_\alpha$ and $\alpha = I \subset \beta$. □

It follows that, similarly to On, the von Neumann ordinals are linearly ordered by the relation <. If we could prove that every well-order is isomorphic to some von Neumann ordinal, we could use On_0 as representatives instead of On.

Corollary 2.20. *For all $\alpha, \beta \in \mathrm{On}_0$, we have*

$$\alpha < \beta \quad or \quad \alpha = \beta \quad or \quad \alpha > \beta\,.$$

Infimum and supremum of sets of von Neumann ordinals can be computed especially easily.

Lemma 2.21. *Let $X \subseteq \mathrm{On}_0$.*

(a) *If X is nonempty then* $\inf X = \bigcap X$.

(b) *If X has an upper bound then* $\sup X = \bigcup X$.

Proof. (a) Since X is nonempty it has a minimal element α, which is also the infimum of X. Clearly, $\bigcap X \subseteq \alpha$. Conversely, if $\beta \in \alpha$ then $\beta \in \gamma$, for all $\gamma \in X$, which implies $\beta \in \bigcap X$. It follows that $\inf X = \alpha = \bigcap X$.

(b) Note that we have $\alpha \leq \beta$ iff $\alpha \subseteq \beta$, for all von Neumann ordinals $\alpha, \beta \in \mathrm{On}_0$.

Clearly, we have $\alpha \subseteq \bigcup X$, for all $\alpha \in X$. Hence, $\bigcup X$ is an upper bound of X. Conversely, let β be an upper bound of X. Then $\alpha \subseteq \beta$, for all $\alpha \in X$, which implies that $\bigcup X \subseteq \beta$. □

The reason why there might be less von Neumann ordinals than elements of On is that each von Neumann ordinal is contained in a new stage. That is, we have exactly one von Neumann ordinal for every stage.

Lemma 2.22. *The function* $f : \mathrm{On}_0 \to H(\mathbb{S})$ *defined by* $f(\alpha) := S(\alpha)$ *is an isomorphism between* On_0 *and the class of all stages.*

Proof. By Lemma 2.19 it follows that f is injective and increasing. Suppose that it is not surjective. Let S be the minimal stage such that $S \notin \operatorname{rng} f$, and set

$$X := \{\, \alpha \in \mathrm{On}_0 \mid S(\alpha) \in S \,\} .$$

Since $X \subseteq S$, X is a set and, hence, a proper initial segment of On_0. Therefore, there is some $\alpha \in \mathrm{On}_0$ such that $X = {\downarrow}\alpha$. Since $S(\beta) \in S$, for all $\beta \in \alpha$, it follows that $S(\alpha) \subseteq S$. By choice of S, we have $S(\alpha) \neq S$. Hence, $S(\alpha) \in S$, which implies that $\alpha \in X = {\downarrow}\alpha$. Contradiction. □

Definition 2.23. For $\alpha \in \mathrm{On}_0$, we set $S_\alpha := S(\alpha)$.

Remark. In On_0 we have finally found the indices to enumerate the cumulative hierarchy

$$S_0 \subset S_1 \subset \cdots \subset S_\alpha \subset S_{\alpha+1} \subset \cdots$$

The class of all stages can be written in the form

$$H(\mathbb{S}) = \{\, S_\alpha \mid \alpha \in \mathrm{On}_0 \,\} ,$$

and we have $\mathbb{S} = \bigcup \{\, S_\alpha \mid \alpha \in \mathrm{On}_0 \,\}$.

Definition 2.24. The *rank* $\rho(a)$ of a set a is the von Neumann ordinal α such that $S(a) = S_\alpha$.

Remark. (a) For $\alpha \in \mathrm{On}_0$, we have $\rho(\alpha) = \alpha$.
(b) Note that

$$\operatorname{cut} A = \{\, x \in A \mid \rho(x) \leq \rho(y) \text{ for all } y \in A \,\} .$$

Lemma 2.25. *A class* X *is a set if, and only if,* $\{\, \rho(x) \mid x \in X \,\}$ *is bounded.*

Exercise 2.3. Prove the preceding lemma.

3. Induction and fixed points

The importance of ordinals stems from the fact that they allow proofs and constructions by *induction.* The next theorem follows immediately from Theorem 1.5.

Theorem 3.1 (Principle of Transfinite Induction). *Let* $I \subseteq \mathrm{On}$ *be an initial segment of* On. *If* $X \subseteq I$ *is a class such that, for every* $\alpha \in I$,

$$\downarrow\alpha \subseteq X \quad \text{implies} \quad \alpha \in X$$

then $X = I$.

Usually one applies this theorem in the following way. If one wants to prove that all ordinals satisfy a certain property φ, it is sufficient to prove that

- 0 satisfies φ;
- if α satisfies φ then so does α^+;
- if δ is a limit ordinal and every $\alpha < \delta$ satisfies φ then so does δ.

Transfinite induction is not only useful for proofs but also to define sequences. For a class A, we set

$$A^{<\infty} := \{\, f \mid f : \downarrow\beta \to a \text{ for some } \beta \in \mathrm{On} \text{ and } a \subseteq A \,\}\,.$$

Lemma 3.2. *Let* H *be a partial function* $H : \mathbb{S}^{<\infty} \to \mathbb{S}$. *For each ordinal* $\alpha \in \mathrm{On}$, *there exists at most one function* $f : \downarrow\alpha \to \mathbb{S}$ *such that* f *is a set and*

$$f(\beta) = H(f \upharpoonright \downarrow\beta)\,, \qquad \text{for all } \beta < \alpha\,.$$

Proof. Suppose that f and g both satisfy the above condition. We apply the Principle of Transfinite Induction to prove that $f = g$. Let

$$X := \{\, \beta \in \downarrow\alpha \mid f(\beta) = g(\beta) \,\}\,.$$

If $\beta < \alpha$ is an ordinal such that $\downarrow\beta \subseteq X$ then $f \restriction \downarrow\beta = g \restriction \downarrow\beta$, which implies that

$$f(\beta) = H(f \restriction \downarrow\beta) = H(g \restriction \downarrow\beta) = g(\beta)\,.$$

Consequently, $\beta \in X$. By the Principle of Transfinite Induction, it follows that $X = \downarrow\alpha$, that is, $f = g$. □

Remark. If a function f satisfies the conditions of the preceding lemma then so does $f \restriction I$, for every initial segment $I \subseteq \operatorname{dom} f$. In particular, if $f : \downarrow\alpha \to \mathbb{S}$ and $g : \downarrow\beta \to \mathbb{S}$ are two such functions with $\alpha \leq \beta$ then $f = g \restriction \downarrow\alpha$.

Definition 3.3. Let H be a partial function $H : \mathbb{S}^{<\infty} \to \mathbb{S}$ and let f_α be the unique function $f_\alpha : \downarrow\alpha \to \mathbb{S}$ such that f_α is a set and

$$f_\alpha(\beta) = H(f_\alpha \restriction \downarrow\beta)\,, \qquad \text{for all } \beta < \alpha\,.$$

Let $I \subseteq \mathrm{On}$ be the class of all α such that f_{α^+} is defined. (Note that I is an initial segment since if $\alpha \in I$ and $\beta < \alpha$ then $f_{\beta^+} = f_{\alpha^+} \restriction \Downarrow\beta$.)

We say that H defines the function F by *transfinite recursion* if

$$\operatorname{dom} F = I \qquad \text{and} \qquad F(\alpha) = f_{\alpha^+}(\alpha)\,, \quad \text{for all } \alpha \in \operatorname{dom} F\,.$$

Theorem 3.4 (Principle of Transfinite Recursion). *Every partial function $H : \mathbb{S}^{<\infty} \to \mathbb{S}$ defines a unique function F by transfinite recursion. We have $F \notin \operatorname{dom} H$ and*

$$F(\alpha) = H(F \restriction \downarrow\alpha)\,, \qquad \textit{for all } \alpha \in \operatorname{dom} F\,.$$

Proof. The existence of F follows immediately from the definition. Note that, by the remark after Lemma 3.2, we have $f_\beta(\alpha) = f_\gamma(\alpha)$, for all $\beta, \gamma > \alpha$. Consequently,

$$F(\alpha) = f_{\alpha^+}(\alpha) = f_\beta(\alpha)\,, \quad \text{for all } \beta > \alpha\,,$$

which implies that

$$F \restriction {\downarrow}\alpha = f_\beta \restriction {\downarrow}\alpha\,, \quad \text{for all } \beta \geq \alpha\,.$$

Therefore, it follows that

$$F(\alpha) = f_{\alpha^+}(\alpha) = H(f_{\alpha^+} \restriction {\downarrow}\alpha) = H(F \restriction {\downarrow}\alpha)\,.$$

In particular, if F is a set then $F = f_\alpha$, for some α. Hence, we have $\operatorname{dom} F = \operatorname{dom} f_\alpha = {\downarrow}\alpha$. Since $\alpha \notin \operatorname{dom} F$ it follows that f_{α^+} does not exists. Hence, $H(f_\alpha) = H(F)$ is undefined and $F \notin \operatorname{dom} H$. If F is a proper class then we trivially have $F \notin \operatorname{dom} H$. □

Remark. After we have introduced the Axiom of Replacement we can actually show that, if $H : \mathbb{S}^{<\infty} \to \mathbb{S}$ is a total function then $\operatorname{dom} F = \mathrm{On}$.

At the moment we can prove this statement only for the special case where $\operatorname{rng} H$ is a set.

Lemma 3.5. *Let A be a set. If $H : A^{<\infty} \to A$ is a total function that defines the function F by transfinite recursion then F is a proper class with* $\operatorname{dom} F = \mathrm{On}$.

Proof. Note that $\operatorname{rng} F \subseteq \operatorname{rng} H \subseteq A$ is a set. If $\operatorname{dom} F = {\downarrow}\alpha \subset \mathrm{On}$ then $F \in A^{\downarrow\alpha} \subset A^{<\infty} = \operatorname{dom} H$ in contradiction to Theorem 3.4. □

Usually definitions by transfinite recursion have the following simpler form. Given an element $a \in A$ and two functions $s : A \to A$ and $h : \wp(A) \to A$ one can construct a unique function $f : I \to A$ such that

- $f(0) = a$;
- $f(\beta^+) = s(f(\beta))$; and
- $f(\delta) = h(f[{\downarrow}\delta])$, for limit ordinals δ.

Example. We can define addition and multiplication of ordinals as follows. By transfinite recursions, we first define the function $\beta \mapsto \alpha + \beta$

by

$$\begin{aligned} \alpha + 0 &:= \alpha\,, \\ \alpha + \beta^+ &:= (\alpha + \beta)^+\,, \\ \alpha + \delta &:= \sup\{\, \alpha + \beta \mid \beta < \delta \,\}\,, \qquad \text{for limit ordinals } \delta\,, \end{aligned}$$

and then we define the function $\beta \mapsto \alpha \cdot \beta$ by

$$\begin{aligned} \alpha \cdot 0 &:= 0\,, \\ \alpha \cdot \beta^+ &:= \alpha \cdot \beta + \alpha,\,, \\ \alpha \cdot \delta &:= \sup\{\, \alpha \cdot \beta \mid \beta < \delta \,\}\,, \qquad \text{for limit ordinals } \delta\,. \end{aligned}$$

By the above theorem, we know that these operations are defined on some initial segment of On and that they are uniquely determined by these equations. Below we will give a different, more concrete definition of addition and multiplication.

Definitions by transfinite recursion are special cases of so-called *inductive fixed points*. Consider a partial order $\langle A, \leq \rangle$ and a function $f : A \to A$. If certain conditions on $\langle A, \leq \rangle$ and f are satisfied, one can compute a fixed point of f in the following way. Starting with some element $a \in A$ we construct the sequence $a,\ f(a),\ f(f(a)), \dots$. If it converges, its limit will be a fixed point of f. The next definition formalises this process.

Definition 3.6. Let $\langle A, \leq \rangle$ be a partial order. A function $f : A \to A$ is *inductive* over an element $a \in A$ if there exists an increasing function $F : I \to A$ where $I \subset \mathrm{On}$ is an initial segment of On such that F is a proper class and we have

$$\begin{aligned} F(0) &= a\,, \\ F(\beta^+) &= f(F(\beta))\,, \\ \text{and} \quad F(\delta) &= \sup F[{\downarrow}\delta]\,, \qquad \text{for limits } \delta\,. \end{aligned}$$

We call F the *fixed-point induction* of f over a. The element $F(\alpha)$ is also called the α-th *stage* of the induction.

Remark. (a) Note that, if A is a set then, by the Principle of Transfinite Recursion, there exists a unique function $F : \mathrm{On} \to A$ satisfying the above equations provided we can show that, for every limit δ, the supremum $\sup F[\downarrow\delta]$ exists. If, furthermore, we can prove that $F(\beta^+) \geq F(\beta)$, for all β, then it follows that f is inductive.

(b) Every fixed-point induction F is continuous, by Lemma 1.13.

Example. (a) The function $f : \mathrm{On} \to \mathrm{On} : \alpha \mapsto \alpha^+$ is inductive. Its fixed-point induction over 0 is the identity function $F : \mathrm{On} \to \mathrm{On} : \alpha \mapsto \alpha$.

(b) Let $f : \mathbb{S} \to \mathbb{S}$ be the function with $f(a) := \wp(a)$. The fixed-point induction of f over $\varnothing$ is the function $F : \mathrm{On}_0 \to \mathbb{S}$ with

$$F(\alpha) := S_\alpha .$$

(c) The graph of addition

$$A := \{\, (x, y, z) \in \mathbb{N}^3 \mid x + y = z \,\}$$

is the least fixed point of the function $f : \wp(\mathbb{N}^3) \to \wp(\mathbb{N}^3)$ with

$$\begin{aligned} f(R) := {} & \{\, (x, 0, x) \mid x \in \mathbb{N} \,\} \\ & \cup \{\, (x, y + 1, z + 1) \mid (x, y, z) \in R \,\} . \end{aligned}$$

Its fixed-point induction over $\varnothing$ is the function

$$F(\alpha) := \begin{cases} \{\, (x, y, z) \mid x + y = z \,, y < \alpha \,\} & \text{if } \alpha < \omega \,, \\ A & \text{if } \alpha \geq \omega \,. \end{cases}$$

(d) Let $\langle V, E \rangle$ be a graph. The function

$$f : \wp(V \times V) \to \wp(V \times V)$$

defined by $f(R) := E \cup E \circ R$ is increasing. Let F be the fixed-point induction of f over $\varnothing$. Then

$$\begin{aligned} F(0) &= \varnothing \,, \\ F(1) &= E \,, \\ F(2) &= E \cup E \circ E \,, \\ F(3) &= E \cup E \circ E \cup E \circ E \circ E \,, \end{aligned}$$

and, generally, we have

$$F(n) = \bigcup_{k<n} E^k\,, \qquad \text{for } n < \omega\,,$$
$$\text{and} \quad F(\alpha) = \bigcup_{k<\omega} E^k\,, \qquad \text{for } \alpha \geq \omega\,.$$

Hence, the inductive fixed point of f is the relation $\bigcup_{k<\omega} E^k = \mathrm{TC}(E)$.

(e) We consider the following simple game between two players. It is played on a graph $\langle V, E\rangle$ where the set of vertices $V = V_0 \cup V_1$ is partitioned into vertices V_0 that belong to player 0 and vertices V_1 belonging to player 1. At the start of the game a pebble is placed on the starting position $v_0 \in V$. In every round one of the players moves this pebble along an edge to a new vertex. If the pebble is on a vertex in V_0 then player 0 can choose where to move, if it is on a vertex in V_1 then player 1 may move. Hence, a play of the game determines a path $v_0, \dots, v_n$ through the graph. If at some point the pebble is on a vertex in V_i without outgoing edge then player i loses. If none of the players manage to manœuvre his opponent into such a situation then the game never stops and *both* players lose. The *winning region* W_i for player i is the set of all vertices w such that, if we start the game in w, then player i has a strategy to win the game. We can compute these winning regions by the fixed-point induction F_i of the function

$$\begin{aligned} f_i(X) := \{\, x \in V_i \mid \text{there is some } y \in X \text{ with } \langle x, y\rangle \in E \,\} \\ \cup \{\, x \in V_{1-i} \mid \text{every } y \text{ with } \langle x, y\rangle \in E \text{ is element of } X \,\}\,. \end{aligned}$$

Note that $F_i(1)$ is the set of all vertices $x \in V_{1-i}$ without outgoing edge. Generally, $F_i(n)$ contains all vertices such that player i has a strategy to win the game in at most n rounds.

Exercise 3.1. Let $\langle V, E\rangle$ be a graph. Prove that $\mathrm{TC}(E) = \bigcup_{n<\omega} E^n$.

If the fixed point induction of a function f converges, its limit is a fixed point of f.

Lemma 3.7. *Let F be the fixed-point induction of a function f. If $F(\alpha) = F(\alpha^+)$ then $F(\alpha) \in \operatorname{fix} f$.*

Proof. $F(\alpha)$ is a fixed point of f since $f(F(\alpha)) = F(\alpha^+) = F(\alpha)$. □

Thus, we can use the fixed point induction F of f to compute a fixed point provided F converges.

Lemma 3.8. *Let F be the fixed-point induction of a function f. If $F(\alpha) = F(\alpha^+)$ then $F(\alpha) = F(\beta)$, for all $\beta \geq \alpha$.*

Proof. We prove the claim by induction on β. If $\beta = \alpha$ then the claim is trivial. For the successor step, we have

$$F(\beta^+) = f(F(\beta)) = f(F(\alpha)) = F(\alpha^+) = F(\alpha)\,.$$

Finally, if $\delta > \alpha$ is a limit ordinal, then

$$\begin{aligned} F(\delta) = \sup\{\, F(\beta) \mid \beta < \delta \,\} &= \sup\{\, F(\beta) \mid \alpha \leq \beta < \delta \,\} \\ &= \sup\{F(\alpha)\} = F(\alpha)\,. \end{aligned}$$ □

If the universe A is a set, every fixed-point induction stabilises at some ordinal. Intuitively, the reason is that the size of the universe A is bounded. Therefore, if we repeat the application of f long enough, we will obtain some element $a \in A$ that already appeared in the sequence.

Theorem 3.9. *Let $\langle A, \leq \rangle$ be a partial order where A is a set. Let $f : A \to A$ be inductive over $a \in A$ and $F : \mathrm{On} \to A$ the corresponding fixed-point induction. There exists some ordinal α such that $F(\alpha) = F(\beta)$, for all $\beta \geq \alpha$.*

Proof. By Theorem 2.12, there exists an ordinal γ such that there is no injective function $\downarrow\gamma \to A$. We claim that there is some $\alpha < \gamma$ such that $F(\alpha) = F(\alpha^+)$. By Lemma 3.8, it then follows that $F(\beta) = F(\alpha)$, for all $\beta \geq \alpha$. Suppose that $F(\alpha) \neq F(\alpha^+)$, for all $\alpha < \gamma$. Since F is increasing it follows that $F \restriction \downarrow\gamma : \downarrow\gamma \to A$ is injective. This contradicts our choice of γ. □

Remark. This proof actually shows that $\alpha < |A|^+$ where $|A|$ is the cardinality of A (see Section A4.2).

Definition 3.10. Let $f : A \to A$ be inductive and $F : \mathrm{On} \to A$ the corresponding fixed-point induction. The minimal ordinal α such that $F(\alpha) = F(\alpha^+)$ is called the *closure ordinal* of the induction and the element $F(\infty) := F(\alpha)$ is the *inductive fixed point* of f over a.

Remark. If A is a set, every inductive function $f : A \to A$ has an inductive fixed point.

Example. Let $\langle A, R\rangle$ be a graph. The *well-founded part* of R is the maximal subset $B \subseteq A$ such that $\langle B, R|_B\rangle$ is well-founded and, for all $\langle a, b\rangle \in R$ with $b \in B$, we also have $a \in B$. We can compute B as inductive fixed point over $\varnothing$ of the function

$$f(X) := \{\, x \in A \mid R^{-1}(x) \subseteq X \cup \{x\} \,\}\,.$$

If we want to apply the above machinery to compute fixed points, we need methods to show that a given function f is inductive. Basically, there are two conditions a function f has to satisfy. The sequence obtained by iterating f has to be linearly ordered and its supremum must exists.

Definition 3.11. Let $\mathfrak{A} = \langle A, \leq\rangle$ be a partial order.

(a) $\mathfrak{A}$ is *inductively ordered* if every chain $C \subseteq A$ that is a set has a supremum.

(b) A function $f : A \to A$ is *inflationary* if $f(a) \geq a$, for all $a \in A$.

Remark. (a) Every inductively ordered set has a least element $\perp$ since the set $\varnothing$ is linearly ordered.

(b) Every complete partial order is inductively ordered.

(c) $\langle \mathrm{On}, \leq\rangle$ is inductively ordered.

(d) If $\langle A, \leq\rangle$ is a well-order then according to Lemma 1.7 all strictly continuous functions $f : A \to A$ are inflationary.

Example. (a) The partial order $\langle F, \subseteq\rangle$ where

$$F := \{\, X \subseteq \mathbb{N} \mid X \text{ is finite} \,\}$$

is not inductively ordered since the chain

$$[0] \subset [1] \subset [2] \subset \cdots \subset [n] \subset \cdots$$

has no upper bound.

(b) Let V be a vector space over the field K and set

$$I := \{ B \subseteq V \mid B \text{ is linearly independent} \} .$$

We claim that $\langle I, \subseteq \rangle$ is inductively ordered.

Let $C \subseteq I$ be a chain. We show that $\sup C = \bigcup C$. By Corollary A2.3.10, it is sufficient to prove that $\bigcup C \in I$.

Suppose otherwise. Then $\bigcup C$ is not linearly independent and there are elements $v_0, \dots, v_n \in \bigcup C$ and $\lambda_0, \dots, \lambda_n \in K$ such that $\lambda_i \neq 0$, for all i, and

$$\lambda_0 v_0 + \dots + \lambda_n v_n = 0 .$$

For each v_i, fix some $B_i \in C$ with $v_i \in B_i$. Since C is linearly ordered so is the set $\{B_0, \dots, B_n\}$. This set is finite and, therefore, it has a maximal element B_k, that is, $B_i \subseteq B_k$, for all i. It follows that $v_0, \dots, v_n \in B_k$, which implies that B_k is not linearly independent. Contradiction.

Lemma 3.12. *Let* $\mathfrak{A} = \langle A, \leq \rangle$ *be inductively ordered.*

(a) *If* $f : A \to A$ *is inflationary,* f *is inductive over every element* $a \in A$.

(b) *If* $f : A \to A$ *is increasing,* f *is inductive over every element* a *with* $f(a) \geq a$.

(c) *If* $f : A \to A$ *is continuous,* f *is inductive over every element* a *with* $f(a) \geq a$. *Furthermore, if the inductive fixed point of* f *over* a *exists, its closure ordinal is at most* ω.

Proof. (a) By transfinite recursion, we construct an increasing function $F : I \to A$ satisfying the equations in Definition 3.6. Let $F(0) := a$. For the inductive step, suppose that $F(\alpha)$ is already defined. We set $F(\alpha^+) := f(F(\alpha))$. Since f is inflationary, it follows that $F(\alpha^+) = f(F(\alpha)) \geq F(\alpha)$. Finally, suppose that δ is a limit ordinal. If $F \restriction {\downarrow}\delta$ is a proper class, we are done. Otherwise, $F[\downarrow\delta]$ is a set which, furthermore, is linearly ordered because $F \restriction \downarrow\delta$ is increasing. As $\langle A, \leq \rangle$ is inductively ordered it follows that $F[\downarrow\delta]$ has a supremum and we can set $F(\delta) := \sup F[\downarrow\delta]$.

(b) Again we define an increasing function $F : I \to A$ by transfinite recursion. Let $F(0) := a$. For the inductive step, suppose that $F(\alpha)$ is already defined. We set $F(\alpha^+) := f(F(\alpha))$. To prove that $F(\alpha^+) \geq F(\alpha)$ we consider three cases. For $\alpha = 0$ we have $F(1) = f(a) \geq a = F(0)$. If $\alpha = \beta^+$ is a successor, we know by inductive hypothesis that $F(\beta^+) \geq F(\beta)$. Since f is increasing it follows that

$$F(\alpha^+) = f(F(\beta^+)) \geq f(F(\beta)) = F(\beta^+) = F(\alpha) .$$

If α is a limit then $F(\alpha) = \sup F[\downarrow\alpha]$ and

$$F(\alpha^+) = f(\sup F[\downarrow\alpha]) \geq f(F(\beta)) = F(\beta^+) , \qquad \text{for all } \beta < \alpha .$$

This implies that

$$F(\alpha^+) \geq \sup F[\downarrow\alpha] = F(\alpha) .$$

Finally, let δ be a limit ordinal. Again, if $F \restriction \downarrow\delta$ is a proper class, we are done. Otherwise, $F[\downarrow\delta]$ is a set and, as above, $F(\delta) := \sup F[\downarrow\delta]$ exists.

(c) Since continuous functions are increasing it follows from (b) that f is inductive over a. Let F be the corresponding fixed-point induction. It remains to show that, if $\omega \in \operatorname{dom} F$ then $F(\infty) = F(\omega)$. Since f is continuous we have

$$\begin{aligned} F(\omega^+) &= f(\sup F[\downarrow\omega]) \\ &= \sup \{ f(F(\alpha)) \mid \alpha < \omega \} \\ &= \sup \{ F(\alpha^+) \mid \alpha < \omega \} \\ &= \sup F[\downarrow\omega] = F(\omega) , \end{aligned}$$

as desired. □

Lemma 3.13. *Let* $f : \mathrm{On} \to \mathrm{On}$ *be strictly continuous and let* $\alpha \in \mathrm{On}$.

(a) *f is inductive over α.*

(b) *If F is the fixed-point induction of f over α then $F(\infty)$ exists if, and only if, the set $\{ f^n(\alpha) \mid n < \omega \}$ is bounded. In this case we have $F(\infty) = F(\omega)$.*

Proof. (a) In Lemma 1.7 we have shown that every strictly continuous function on a well-order is inflationary. Therefore, Lemma 3.12 implies that f is inductive over α.

(b) We prove by induction on $n < \omega$ that $n \in \operatorname{dom} F$. By definition we have $0 \in \operatorname{dom} F$. If $n \in \operatorname{dom} F$ then $f(F(n)) \geq F(n)$ since f is inflationary. Hence, $F(n+1) = f(F(n))$ is defined. If

$$\{ f^n(\alpha) \mid n < \omega \} = F[\downarrow\omega]$$

is bounded, it follows that $F(\omega) = \sup F[\downarrow\omega]$ is defined. Consequently, Lemma 3.12 implies that $F(\infty) = F(\omega)$. □

Exercise 3.2. Let $f : \wp(A) \to \wp(A)$ be inflationary and increasing, and let $c : \wp(A) \to \wp(A)$ be the function that maps $X \subseteq A$ to the inductive fixed point of f over X. Prove that c is a closure operator.

We conclude this section with two theorems which can be used to prove the existence of fixed points. The first one is an immediate consequence of the above results.

Theorem 3.14 (Bourbaki). *Let $\langle A, \leq \rangle$ be an inductively ordered graph. If A is a set then every inflationary function $f : A \to A$ has an inductive fixed point.*

Proof. By Lemma 3.12, f is inductive over $\perp$. Consequently, f has an inductive fixed point, by Theorem 3.9. □

Example. The condition of A being a set is necessary. For instance, $\mathfrak{On}$ is inductively ordered since every set of ordinals has a supremum and the function $f : \mathrm{On} \to \mathrm{On} : \alpha \mapsto \alpha^+$ is inflationary. But f has no fixed point.

The second theorem is a version of the Theorem of Knaster and Tarski which shows that we can compute the least fixed point of a function f by a fixed-point induction.

Theorem 3.15. *Let $\langle A, \leq \rangle$ be an inductively ordered graph where A is a set and let $f : A \to A$ be an increasing function. If the least fixed point of f exists then it coincides with its inductive fixed point over $\bot$.*

Proof. Let $F : \mathrm{On} \to A$ be the fixed-point induction of f over $\bot$. Suppose that $a := \operatorname{lfp} f$ exists. We prove by induction on α that $F(\alpha) \leq a$. Then it follows that $F(\infty) \leq a$ and the minimality of a implies that $F(\infty) = a$.

Clearly, $F(0) = \bot \leq a$. For the inductive step, suppose that $F(\alpha) \leq a$. Since f is increasing it follows that

$$F(\alpha^+) = f(F(\alpha)) \leq f(a) = a\,.$$

Finally, if δ is a limit ordinal, the inductive hypothesis implies that

$$F(\delta) = \sup \{ F(\alpha) \mid \alpha < \delta \} \leq a\,. \qquad \square$$

4. Ordinal arithmetic

Many properties of natural numbers can be generalised to ordinals. We have already seen that ordinals allow proofs by induction. In this section we will show how to define addition, multiplication, and exponentiation for such numbers.

We start by defining these operations for arbitrary linear orders. Intuitively, the sum of two linear orders $\mathfrak{A}$ and $\mathfrak{B}$ is the order consisting of a copy of $\mathfrak{A}$ followed by a copy of $\mathfrak{B}$. Similarly, their product is obtained from $\mathfrak{B}$ by replacing every element by a copy of $\mathfrak{A}$.

Definition 4.1. Let $\mathfrak{A} = \langle A, \leq_A \rangle$ and $\mathfrak{B} = \langle B, \leq_B \rangle$ be linear orders.

(a) The *sum* $\mathfrak{A} + \mathfrak{B}$ is the graph $\langle C, \leq_C \rangle$ where

$$C := A \mathbin{\dot\cup} B = (\{0\} \times A) \cup (\{1\} \times B)$$

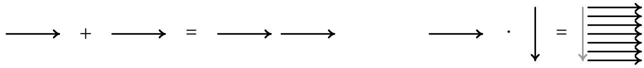


Figure 1.. Sum and product of linear orders

and the order is defined by

$$\begin{aligned}\langle i, a\rangle \leq_C \langle k, b\rangle \quad &:\text{iff} \quad i = k = 0 \text{ and } a \leq_A b \\ &\text{or } i = k = 1 \text{ and } a \leq_B b \\ &\text{or } i = 0 \text{ and } k = 1\,.\end{aligned}$$

(b) The *product* $\mathfrak{A} \cdot \mathfrak{B}$ is the graph $\langle C, \leq_C\rangle$ where $C := A \times B$ and the order is defined by

$$\langle a, b\rangle \leq_C \langle a', b'\rangle \quad :\text{iff} \quad b <_B b' \text{ or } (b = b' \text{ and } a \leq_A a')\,.$$

(This is the reversed lexicographic ordering, see Definition B2.1.1.)

(c) If $\mathfrak{A}$ and $\mathfrak{B}$ are well-orders then we define $\mathfrak{A}^{(\mathfrak{B})} := \langle C, \leq_C\rangle$ where

$$\begin{aligned}C := \{\, f \in A^B \mid{} &\text{there are only finitely many } b \in B \text{ with} \\ &f(b) \neq \bot \,\}\,,\end{aligned}$$

and the order is defined by

$$\begin{aligned}f <_C g \quad :\text{iff} \quad &\text{the set } \{\, b \in B \mid f(b) \neq g(b) \,\} \text{ has a maximal} \\ &\text{element } b_0 \text{ and we have } f(b_0) <_A g(b_0)\,.\end{aligned}$$

For natural numbers, these operations coincide with the usual ones.

Exercise 4.1. Let $\mathfrak{K} := \langle [k], \leq\rangle$ and $\mathfrak{M} := \langle [m], \leq\rangle$ where $k, m < \omega$. Prove that

(a) $\mathfrak{K} + \mathfrak{M} \cong \langle [k + m], \leq\rangle$,

(b) $\mathfrak{K} \cdot \mathfrak{M} \cong \langle [km], \leq\rangle$,

(c) $\mathfrak{K}^{(\mathfrak{M})} \cong \langle [k^m], \leq \rangle$.

Addition of linear orders is associative and the empty order is a neutral element. Below we will give an example showing that, in general, it is not commutative.

Lemma 4.2. *If $\mathfrak{A}$, $\mathfrak{B}$, and $\mathfrak{C}$ are linear orders then*

$$(\mathfrak{A} + \mathfrak{B}) + \mathfrak{C} \cong \mathfrak{A} + (\mathfrak{B} + \mathfrak{C}) .$$

Proof. Let $\mathfrak{A} = \langle A, \leq_A \rangle$, $\mathfrak{B} = \langle B, \leq_B \rangle$, and $\mathfrak{C} = \langle C, \leq_C \rangle$. We can define a bijection $f : (A \cup B) \cup C \to A \cup (B \cup C)$ by

$$\begin{aligned} f\langle 0, \langle 0, a \rangle\rangle &:= \langle 0, a \rangle && \text{for } a \in A , \\ f\langle 0, \langle 1, b \rangle\rangle &:= \langle 1, \langle 0, b \rangle\rangle && \text{for } b \in B , \\ f\langle 1, c \rangle &:= \langle 1, \langle 1, c \rangle\rangle && \text{for } c \in C . \end{aligned}$$

Since this bijection preserves the ordering it is the desired isomorphism. □

As we want to define arithmetic operations on ordinals we have to show that, if we apply the above operations to well-orders, we again obtain a well-order.

Lemma 4.3. *If $\mathfrak{A}$ and $\mathfrak{B}$ are well-orders then so are $\mathfrak{A} + \mathfrak{B}$, $\mathfrak{A} \cdot \mathfrak{B}$, and $\mathfrak{A}^{(\mathfrak{B})}$.*

Proof. Suppose that $\mathfrak{A} = \langle A, \leq_A \rangle$ and $\mathfrak{B} = \langle B, \leq_B \rangle$. We will prove the claim only for $\mathfrak{C} := \mathfrak{A}^{(\mathfrak{B})}$. The other operations are left as an exercise to the reader.

Let $\mathfrak{C} = \langle C, \leq_C \rangle$. The relation $<_C$ is irreflexive since, for each $f \in C$, the set $\{ b \in B \mid f(b) \neq f(b) \}$ is empty and has no maximal element. Furthermore, $<_C$ is linear. For transitivity, let $f, g, h \in C$ be functions such that $f <_C g <_C h$. Let $b_0, b_1 \in B$ be the maximal elements such that, respectively, $f(b_0) \neq g(b_0)$ and $g(b_1) \neq h(b_1)$. By definition, we have $f(b_0) <_A g(b_0)$ and $g(b_1) <_A h(b_1)$. If $b_0 \leq_B b_1$ then

$$f(b_1) \leq g(b_1) <_A h(b_1)$$

$$\text{and} \quad f(b) = g(b) = h(b)\,, \qquad \text{for } b >_B b_1\,,$$

implies that $f <_C h$. Similarly, if $b_1 <_B b_0$ then

$$f(b_0) <_A g(b_0) = h(b_0)$$

$$\text{and} \quad f(b) = g(b) = h(b)\,, \qquad \text{for } b >_B b_0\,.$$

In both cases it follows that $f <_C h$. Consequently, $<_C$ is a strict linear order.

It remains to prove that every nonempty subset $X \subseteq C$ has a minimal element. We prove the claim by induction on $\beta := \operatorname{ord}(\mathfrak{B})$. If $\beta = 0$ then $C = A^{(\varnothing)} = \{\varnothing\}$ and we are done. Suppose that $\beta > 0$ and select an arbitrary element $f \in X$. If $f(b) = \bot$, for all $b \in B$, then f is the minimal element of X and we are done. Hence, we may assume that there is some $b \in B$ with $f(b) \neq \bot$. Since there are only finitely many such elements we may assume that b is the maximal one. Define

$$Y := \{\, g \in X \mid g(c) = \bot \text{ for all } c > b \,\}\,.$$

This set is nonempty since $f \in Y$. Set

$$a := \min\{\, g(b) \mid g \in Y \,\} \quad \text{and} \quad Z := \{\, g \in Y \mid g(b) = a \,\}\,.$$

By construction, we have $g <_C h$ whenever $g \in Z$ and $h \in X \setminus Z$. Consequently, if we can find a minimal element of Z, we also have the minimal element of X. Let

$$U := \{\, g \restriction {\downarrow} b \mid g \in Z \,\} \subseteq A^{({\downarrow} b)}.$$

Since $\operatorname{ord}({\downarrow} b) < \beta$ we can apply the inductive hypothesis and there exists a minimal element $h \in U$. Note that the restriction map

$$\rho : Z \to U : g \mapsto g \restriction {\downarrow} b$$

is a bijection since we have

$$g(c) = g'(c) \quad \text{for all } g, g' \in Z \text{ and every } c \geq b\,.$$

Furthermore, ρ preserves the ordering, that is, it is an isomorphism. It follows that $\rho^{-1}(h)$ is the minimal element of Z and of X. □

Exercise 4.2. Show that, if $\mathfrak{A}$ and $\mathfrak{B}$ are well-orders then so are $\mathfrak{A} + \mathfrak{B}$ and $\mathfrak{A} \cdot \mathfrak{B}$.

It is easy to see that $\mathfrak{A} \cong \mathfrak{A}'$ and $\mathfrak{B} \cong \mathfrak{B}'$ implies that the sums, products, and powers are also isomorphic. Therefore, we can define the corresponding operations on ordinals by taking representatives.

Definition 4.4. For $\alpha = \operatorname{ord}(\mathfrak{A})$ and $\beta = \operatorname{ord}(\mathfrak{B})$ we define

$$\begin{aligned} \alpha + \beta &:= \operatorname{ord}(\mathfrak{A} + \mathfrak{B})\,, \\ \alpha \cdot \beta &:= \operatorname{ord}(\mathfrak{A} \cdot \mathfrak{B})\,, \\ \alpha^{(\beta)} &:= \operatorname{ord}\left(\mathfrak{A}^{(\mathfrak{B})}\right). \end{aligned}$$

Example. The following equations can be proved easily by the lemmas below. We encourage the reader to derive them directly from the definitions.

$$\begin{aligned} 1 + 1 &= 2 & (3 + 6)\omega &= 9\omega = \omega < \omega 2 = 3\omega + 6\omega \\ \omega + \omega &= \omega 2 & (\omega 6 + 17)\omega &= \omega\omega = \omega^{(2)} \\ 1 + \omega &= \omega < \omega + 1 & 2^{(\omega)} &= \omega \\ 2\omega &= \omega < \omega 2 \end{aligned}$$

Exercise 4.3. Show that $\alpha + \beta$, $\alpha \cdot \beta$, and $\alpha^{(\beta)}$ are well-defined, for all $\alpha, \beta \in \mathrm{On}$.

Exercise 4.4. Show that $\alpha^+ = \alpha + 1$.

Ordinal addition

The properties of ordinal addition, multiplication, and exponentiation are similar to, but not quite the same as those for integers. The following sequence of lemmas summarises them. We start with addition.

Lemma 4.5. *Let $\alpha, \beta, \gamma \in \mathrm{On}$. If $\beta < \gamma$ then $\alpha + \beta < \alpha + \gamma$.*

Proof. Fix representatives $\alpha = \mathrm{ord}(\mathfrak{A})$, $\beta = \mathrm{ord}(\mathfrak{B})$, and $\gamma = \mathrm{ord}(\mathfrak{C})$. There exists an isomorphism $f : B \to I \subset C$ between B and some proper initial segment I of C. We define an isomorphism $g : A \uplus B \to A \uplus I$ by

$$g(\langle 0, a\rangle) := \langle 0, a\rangle\,, \qquad \text{for } a \in A\,,$$
and $$g(\langle 1, b\rangle) := \langle 1, f(b)\rangle\,, \qquad \text{for } b \in B\,.$$

Hence, $\mathfrak{A} + \mathfrak{B} < \mathfrak{A} + \mathfrak{C}$. □

In the last section we gave an inductive definition of addition. The next lemma shows that it is equivalent to the official definition above.

Lemma 4.6. *Let $\alpha, \beta \in \mathrm{On}$.*

(a) $\alpha + 0 = \alpha$.

(b) $\alpha + \beta^+ = (\alpha + \beta)^+$.

(c) $\alpha + \delta = \sup\{\, \alpha + \beta \mid \beta < \delta \,\}$, *for limit ordinals δ.*

Proof. Fix representatives $\alpha = \mathrm{ord}(\mathfrak{A})$ and $\beta = \mathrm{ord}(\mathfrak{B})$.

(a) follows immediately since $\mathfrak{A} + \langle \emptyset, \leq\rangle \cong \mathfrak{A}$.

(b) By Lemma 4.2, we have

$$(\mathfrak{A} + \mathfrak{B}) + \mathfrak{C} \cong \mathfrak{A} + (\mathfrak{B} + \mathfrak{C})\,, \qquad \text{for all linear orders } \mathfrak{A}, \mathfrak{B}, \mathfrak{C}\,.$$

Since $\beta^+ = \mathrm{ord}(\mathfrak{B} + \langle [1], \leq\rangle)$ the result follows.

(c) Let $X := \{\, \alpha + \beta \mid \beta < \delta \,\}$ and set $\gamma := \sup X$. By Lemma 4.5, we have $\alpha + \beta < \alpha + \delta$, for all $\beta < \delta$, which implies that $\gamma \leq \alpha + \delta$.

For a contradiction suppose that $\gamma < \alpha + \delta$. Fix representatives $\gamma = \operatorname{ord}(\mathfrak{C})$ and $\delta = \operatorname{ord}(\mathfrak{D})$. Since $\alpha + 0 < \gamma < \alpha + \delta$ there exists an isomorphism $f : C \to A \uplus I$, for some proper initial segment $\varnothing \subset I \subset D$. Let $C_0 := f^{-1}[A]$ and $C_1 := f^{-1}[I]$. Since f is an isomorphism we have

$$\mathfrak{A} \cong \langle C_0, \leq \rangle \quad \text{and} \quad \mathfrak{C} \cong \langle C_0, \leq \rangle + \langle C_1, \leq \rangle .$$

Set $\beta := \operatorname{ord}(\langle C_1, \leq \rangle)$. It follows that $\gamma = \alpha + \beta$. Furthermore, because of the inclusion map $I \to D$ we have $\beta < \delta$. By (b) it follows that

$$\gamma < (\alpha + \beta)^+ = \alpha + \beta^+ \leq \sup X .$$

Contradiction. □

Corollary 4.7. *The function* $f_\alpha : \mathrm{On} \to \mathrm{On}$ *with* $f_\alpha(\beta) := \alpha + \beta$ *is strictly continuous, for every* $\alpha \in \mathrm{On}$.

Proof. The claim follows immediately from the preceding lemma and Lemma 1.13. □

Since ordinal addition is not commutative there are two possible ways to subtract ordinals. Given $\alpha \geq \beta$ we can ask for some ordinal γ such that $\alpha = \beta + \gamma$, or we can ask for some γ with $\alpha = \gamma + \beta$. The next lemma shows that the first operation is well-defined. The second one is not since, for example, $1 + \omega = \omega = 2 + \omega$.

Lemma 4.8. *For all ordinals* $\beta \leq \alpha$, *there exists a unique ordinal* γ *such that* $\alpha = \beta + \gamma$.

Proof. By Corollary 4.7 and Lemma 1.14, there exists a greatest ordinal γ such that $\beta + \gamma \leq \alpha$. If $\beta + \gamma < \alpha$ then we would have

$$(\beta + \gamma)^+ = \beta + \gamma^+ \leq \alpha$$

in contradiction to the choice of γ. Hence, $\beta + \gamma = \alpha$. The uniqueness of γ follows from the fact that the function $\gamma \mapsto \beta + \gamma$ is injective. □

The next lemma summarises the laws of ordinal addition.

Lemma 4.9. *Let* $\alpha, \beta, \gamma \in \mathrm{On}$.

(a) $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$.

(b) $\alpha + \beta = \alpha + \gamma$ *implies* $\beta = \gamma$.

(c) $\alpha \leq \beta$ *implies* $\alpha + \gamma \leq \beta + \gamma$.

(d) *If* $X \subseteq \mathrm{On}$ *is nonempty and bounded then*

$$\alpha + \sup X = \sup \{ \alpha + \beta \mid \beta \in X \} .$$

(e) $\beta \leq \alpha$ *if, and only if,* $\alpha = \beta + \gamma$, *for some* $\gamma \in \mathrm{On}$.

(f) $\beta < \alpha$ *if, and only if,* $\alpha = \beta + \gamma$, *for some* $\gamma \in \mathrm{On} \setminus \{0\}$.

Proof. Fix representatives $\alpha = \mathrm{ord}(\mathfrak{A})$, $\beta = \mathrm{ord}(\mathfrak{B})$ and $\gamma = \mathrm{ord}(\mathfrak{C})$.

(a) follows from Lemma 4.2; (b) follows from Lemma 4.8; and (d) follows from Corollary 4.7.

(c) We prove the claim by induction on γ. For $\gamma = 0$, we have

$$\alpha + 0 = \alpha \leq \beta = \beta + 0 .$$

For the successor step, note that $\alpha \leq \beta$ implies $\alpha^+ \leq \beta^+$. Hence, it follows that

$$\alpha + \gamma^+ = (\alpha + \gamma)^+ \leq (\beta + \gamma)^+ = \beta + \gamma^+ .$$

It remains to consider the limit step. For every $\eta < \gamma$, the inductive hypothesis yields

$$\alpha + \eta \leq \beta + \eta < \beta + \gamma .$$

Therefore, Lemma 4.6 (c) implies that

$$\alpha + \gamma = \sup \{ \alpha + \eta \mid \eta < \gamma \} \leq \beta + \gamma .$$

(e) If $\beta < \alpha$, we obtain by Lemma 4.8 some $\gamma \in \mathrm{On}$ with $\alpha = \beta + \gamma$. Conversely, if $\beta + \gamma = \alpha$ then there exists an isomorphism

$$f : B \uplus C \to A\,.$$

We can define an isomorphism $g : B \to I \subseteq A$ by

$$g(b) := f(\langle 0, b\rangle)\,.$$

This implies that $\mathfrak{B} \leq \mathfrak{A}$.

(f) follows immediately from (e). ☐

Ordinal multiplication

After addition we turn to ordinal multiplication. The development is analogous to the one above. First, we show that the function $\beta \mapsto \alpha\beta$ is strictly increasing.

Lemma 4.10. *Let $\alpha, \beta, \gamma \in \mathrm{On}$. If $\alpha \neq 0$ and $\beta < \gamma$ then $\alpha\beta < \alpha\gamma$.*

Proof. Fix representatives $\alpha = \mathrm{ord}(\mathfrak{A})$, $\beta = \mathrm{ord}(\mathfrak{B})$, and $\gamma = \mathrm{ord}(\mathfrak{C})$. By assumption, there exists an isomorphism $f : B \to I \subset C$ between B and a proper initial segment of C. We can define an isomorphism $g : A \times B \to A \times I$ by

$$g(\langle a, b\rangle) := \langle a, f(b)\rangle\,.$$

Since $A \times I$ is a proper initial segment of $A \times C$ it follows that $\alpha\beta < \alpha\gamma$. ☐

Again the inductive definition coincides with the official one.

Lemma 4.11. *Let $\alpha, \beta \in \mathrm{On}$.*

(a) $\alpha \cdot 0 = 0\,.$

(b) $\alpha\beta^+ = \alpha\beta + \alpha\,.$

(c) $\alpha\delta = \sup\{\, \alpha\beta \mid \beta < \delta \,\}\,,$ *for limit ordinals δ.*

Proof. Fix representatives $\alpha = \mathrm{ord}(\mathfrak{A})$ and $\beta = \mathrm{ord}(\mathfrak{B})$.

(a) follows immediately from the fact that $\mathfrak{A} \cdot \langle \varnothing, \varnothing \rangle = \langle \varnothing, \varnothing \rangle$.

(b) The canonical bijection

$$A \times (B \cup [1]) \to (A \times B) \cup A$$

given by

$$\langle a, \langle 0, b \rangle \rangle \mapsto \langle 0, \langle a, b \rangle \rangle ,$$
$$\langle a, \langle 1, 0 \rangle \rangle \mapsto \langle 1, a \rangle ,$$

induces an isomorphism

$$\mathfrak{A} \cdot (\mathfrak{B} + \langle [1], \leq \rangle) \to \mathfrak{A} \cdot \mathfrak{B} + \mathfrak{A} .$$

(c) Let $X := \{ \alpha\beta \mid \beta < \delta \}$ and set $\gamma := \sup X$. By Lemma 4.10, we have $\alpha\beta < \alpha\delta$, for all $\beta < \delta$. Hence, $\gamma = \sup X \leq \alpha\delta$.

For a contradiction suppose that $\gamma < \alpha\delta$. Fix representatives $\gamma = \mathrm{ord}(\mathfrak{C})$ and $\delta = \mathrm{ord}(\mathfrak{D})$. Since $\gamma < \alpha\delta$ there exists an isomorphism $f : C \to I$, for some proper initial segment $\varnothing \subset I \subset A \times D$. Let $\langle a, d \rangle$ be the minimal element of $A \times D \setminus I$. Then $I = (A \times {\downarrow}d) \cup ({\downarrow}a \times \{d\})$, which implies that

$$\gamma = \alpha \cdot \mathrm{ord}({\downarrow}d) + \mathrm{ord}({\downarrow}a) .$$

Since $\mathrm{ord}({\downarrow}a) < \alpha$ and $\beta := \mathrm{ord}({\downarrow}d) < \delta$ it follows that

$$\gamma < \alpha\beta + \alpha = \alpha\beta^+ \leq \sup X .$$

Contradiction. ☐

Corollary 4.12. *The function* $f_\alpha : \mathrm{On} \to \mathrm{On}$ *with* $f_\alpha(\beta) := \alpha\beta$ *is strictly continuous, for every* $\alpha > 0$.

Proof. The claim follows immediately from the preceding lemma and Lemma 1.13. ☐

We can also show that ordinals allow a limited form of division.

Lemma 4.13. *For all ordinals $\alpha, \beta \in \mathrm{On}$ with $\beta \neq 0$, there exist unique ordinals γ and $\rho < \beta$ such that $\alpha = \beta\gamma + \rho$.*

Proof. By Corollary 4.12 and Lemma 1.14, there exists a greatest ordinal γ such that $\beta\gamma \leq \alpha$, and, by Lemma 4.8, there exists some ordinal ρ such that $\beta\gamma + \rho = \alpha$. By choice of γ, we have

$$\beta\gamma + \beta = \beta(\gamma + 1) > \alpha = \beta\gamma + \rho\,,$$

which implies that $\rho < \beta$.

Suppose there exist ordinals $\delta \neq \gamma$ and $\sigma < \beta$ such that $\beta\delta + \sigma = \alpha$. Since $\beta\delta \leq \alpha$ we have $\delta < \gamma$, which implies that

$$\alpha = \beta\gamma + \rho \geq \beta\delta^+ = \beta\delta + \beta > \beta\delta + \sigma = \alpha\,.$$

A contradiction. It follows that γ is unique. Hence, the uniqueness of ρ follows from Lemma 4.8. □

Lemma 4.14. *α is a limit ordinal if, and only if, $\alpha = \omega\beta$, for some $\beta > 0$.*

Proof. ($\Rightarrow$) By Lemma 4.13, we have $\alpha = \omega\beta + n$ for some $\beta \in \mathrm{On}$ and $n < \omega$. Suppose that $n \neq 0$. Then $n = m + 1$, for some $m < \omega$, and

$$\alpha = \omega\beta + (m + 1) = (\omega\beta + m) + 1\,.$$

Consequently, α is a successor ordinal. Contradiction.

($\Leftarrow$) Suppose that $\omega\beta$ is a successor ordinal. That is, $\omega\beta = \gamma + 1$, for some γ. By Lemma 4.13, we can write γ as $\gamma = \omega\eta + n$, for some $n < \omega$. Hence,

$$\omega\beta = \gamma + 1 = \omega\eta + (n + 1)\,.$$

By Lemma 4.13, it follows that $\beta = \eta$ and $0 = n + 1$. Contradiction. □

The laws of ordinal multiplication are summarised in the following lemma.

Lemma 4.15. *Let* $\alpha, \beta, \gamma \in \mathrm{On}$.

(a) $\alpha(\beta\gamma) = (\alpha\beta)\gamma$.

(b) $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.

(c) *If* $\alpha \neq 0$ *and* $\alpha\beta = \alpha\gamma$ *then* $\beta = \gamma$.

(d) $\alpha \leq \beta$ *implies* $\alpha\gamma \leq \beta\gamma$.

(e) *If* $X \subseteq \mathrm{On}$ *is nonempty and bounded then*

$$\alpha \cdot \sup X = \sup \{ \alpha\beta \mid \beta \in X \} .$$

Proof. (b) We prove the claim by induction on γ. For $\gamma = 0$, we have

$$\alpha(\beta + 0) = \alpha\beta = \alpha\beta + 0 = \alpha\beta + \alpha 0 .$$

For the successor step, we have

$$\begin{aligned} \alpha(\beta + \gamma^+) &= \alpha(\beta + \gamma)^+ \\ &= \alpha(\beta + \gamma) + \alpha \\ &= \alpha\beta + \alpha\gamma + \alpha \\ &= \alpha\beta + \alpha\gamma^+ . \end{aligned}$$

Finally, if γ is a limit ordinal then

$$\begin{aligned} \alpha(\beta + \gamma) &= \alpha \cdot \sup \{ \beta + \rho \mid \rho < \gamma \} \\ &= \sup \{ \alpha(\beta + \rho) \mid \rho < \gamma \} \\ &= \sup \{ \alpha\beta + \alpha\rho \mid \rho < \gamma \} \\ &= \alpha\beta + \sup \{ \alpha\rho \mid \rho < \gamma \} \\ &= \alpha\beta + \alpha\gamma . \end{aligned}$$

(a) and (d) can also be proved by induction on γ. We leave the details as an exercise to the reader.

(c) and (e) follow immediately from Corollary 4.12. □

Ordinal exponentiation

Finally, we consider ordinal exponentiation. Again, the basic steps are the same as for addition and multiplication.

Lemma 4.16. *Let $\alpha, \beta, \gamma \in \mathrm{On}$. If $\alpha > 1$ and $\beta < \gamma$ then $\alpha^{(\beta)} < \alpha^{(\gamma)}$.*

Proof. Fix representatives $\alpha = \mathrm{ord}(\mathfrak{A})$, $\beta = \mathrm{ord}(\mathfrak{B})$, and $\gamma = \mathrm{ord}(\mathfrak{C})$. There exists an isomorphism $f : B \to I \subset C$ between B and a proper initial segment I of C. The desired isomorphism

$$A^{(B)} \to A^{(I)} \subset A^{(C)}$$

is given by the mapping $g \mapsto g \circ f^{-1}$. □

Ordinal exponentiation can also be defined inductively.

Lemma 4.17. *Let $\alpha, \beta \in \mathrm{On}$.*

(a) $\alpha^{(0)} = 1$.

(b) $\alpha^{(\beta^+)} = \alpha^{(\beta)} \alpha$.

(c) $\alpha^{(\delta)} = \sup \{ \alpha^{(\beta)} \mid \beta < \delta \}$, *for limit ordinals δ.*

Proof. Fix representatives $\alpha = \mathrm{ord}(\mathfrak{A})$ and $\beta = \mathrm{ord}(\mathfrak{B})$.

(a) Since $\varnothing$ is the only function with empty domain we have $A^{(\varnothing)} = A^{\varnothing} = \{\varnothing\}$.

(b) There is a canonical bijection $A^{(B \cup [1])} \to A^{(B)} \times A$ given by

$$f \mapsto \langle f', f(\langle 1, 0 \rangle) \rangle$$

where the function $f' : B \to A$ is defined by $f'(b) := f(\langle 0, b \rangle)$. This bijection induces the desired isomorphism

$$\mathfrak{A}^{(\mathfrak{B} + \langle [1], \leq \rangle)} \to \mathfrak{A}^{(\mathfrak{B})} \cdot \mathfrak{A}.$$

(c) If $\alpha < 2$, the claim is trivial. Hence, we may assume that $\alpha > 1$. Let $X := \{ \alpha^{(\beta)} \mid \beta < \delta \}$ and set $\gamma := \sup X$. By Lemma 4.16, we have $\alpha^{(\beta)} < \alpha^{(\delta)}$, for all $\beta < \delta$. Hence, $\gamma = \sup X \leq \alpha^{(\delta)}$.

For a contradiction suppose that $\gamma < \alpha^{(\delta)}$. Fix representatives $\gamma = \operatorname{ord}(\mathfrak{C})$ and $\delta = \operatorname{ord}(\mathfrak{D})$. Since $\gamma < \alpha^{(\delta)}$, there exists an isomorphism $f : C \to I$, for some proper initial segment $I \subset A^{(D)}$. Let g be the minimal element of $A^{(D)} \setminus I$ and let $d_0 < \dots < d_n$ be the enumeration of the set $\{ d \in D \mid g(d) \neq 0 \}$. We can decompose I as $I = I_n \cup \dots \cup I_0$ where, for each $i \leq n$,

$$I_i := \{ h \in A^D \mid h(d_i) < g(d_i) \text{ and } h(x) = g(x) \,,\ \text{for } x > d_i \} \,.$$

Set $\beta_i := \operatorname{ord}(\downarrow d_i) < \delta$ and $\eta_i := \operatorname{ord}(\downarrow g(d_i))$. It follows that

$$\begin{aligned} \gamma &= \alpha^{(\beta_n)} \cdot \eta_n + \dots + \alpha^{(\beta_0)} \cdot \eta_0 \\ &< \alpha^{(\beta_n)} \alpha + \dots + \alpha^{(\beta_0)} \alpha \\ &\leq \alpha^{(\beta_n)} \alpha + \dots + \alpha^{(\beta_n)} \alpha \\ &= \alpha^{(\beta_n + 1)} (n + 1) \,. \end{aligned}$$

Since $\alpha > 1$ there is some finite ordinal m such that $\alpha^{(m)} \geq n+1$. Therefore, it follows by (b) that

$$\gamma < \alpha^{(\beta_n + 1)} \alpha^{(m)} = \alpha^{(\beta_n + m + 1)} \leq \sup X \,.$$

Contradiction. □

Corollary 4.18. *The function $f_\alpha : \mathrm{On} \to \mathrm{On}$ with $f_\alpha(\beta) := \alpha^{(\beta)}$ is strictly continuous, for every $\alpha > 1$.*

Proof. The claim follows immediately from the preceding lemma and Lemma 1.13. □

Besides subtraction and division we can also take a limited form of logarithms.

Lemma 4.19. *For all ordinals $\alpha, \beta \in \mathrm{On}$ with $\alpha > 0$ and $\beta > 1$, there exist unique ordinals γ, η, and ρ with $0 < \gamma < \beta$ and $\rho < \beta^{(\eta)}$ such that $\alpha = \beta^{(\eta)} \gamma + \rho$.*

Proof. By Corollary 4.18 and Lemma 1.14, there exists a greatest ordinal η such that $\beta^{(\eta)} \leq \alpha$, and, by Lemma 4.13, there exist ordinals γ and $\rho < \beta^{(\eta)}$ such that $\beta^{(\eta)}\gamma+\rho = \alpha$. If $\gamma = 0$, we would have $\rho = \alpha \geq \beta^{(\eta)} > \rho$. A contradiction. And, if $\gamma \geq \beta$, we would have

$$\alpha < \beta^{(\eta+1)} = \beta^{(\eta)}\beta \leq \beta^{(\eta)}\gamma \leq \beta^{(\eta)}\gamma + \rho = \alpha\,.$$

Again a contradiction. Therefore, $0 < \gamma < \beta$.

Suppose there exist ordinals $\mu \neq \eta$, δ, and σ such that $\beta^{(\mu)}\delta + \sigma = \alpha$. Since $\beta^{(\mu)} \leq \alpha$ we have $\mu < \eta$, which implies that

$$\begin{aligned}\alpha = \beta^{(\eta)}\gamma + \rho \geq \beta^{(\mu^+)} = \beta^{(\mu)}\beta \geq \beta^{(\mu)}(\delta+1) = \beta^{(\mu)}\delta + \beta^{(\mu)} \\ > \beta^{(\mu)}\delta + \sigma = \alpha\,.\end{aligned}$$

A contradiction. It follows that η is unique. Hence, the uniqueness of γ and ρ follows from Lemma 4.8. □

Let us summarise the laws of ordinal exponentiation.

Lemma 4.20. *Let* $\alpha, \beta, \gamma \in \mathrm{On}$.

(a) $\alpha^{(\beta+\gamma)} = \alpha^{(\beta)}\alpha^{(\gamma)}$.

(b) $\alpha^{(\beta\gamma)} = (\alpha^{(\beta)})^{(\gamma)}$.

(c) $\alpha > 1$ *implies* $\beta \leq \alpha^{(\beta)}$.

(d) *If* $\alpha > 1$ *and* $\alpha^{(\beta)} = \alpha^{(\gamma)}$ *then* $\beta = \gamma$.

(e) $\alpha \leq \beta$ *implies* $\alpha^{(\gamma)} \leq \beta^{(\gamma)}$.

(f) *If* $\alpha > 1$ *then we have* $\beta < \gamma$ *if, and only if,* $\alpha^{(\beta)} < \alpha^{(\gamma)}$.

(g) *If* $X \subseteq \mathrm{On}$ *is nonempty and bounded then we have*

$$\alpha^{(\sup X)} = \sup\{\, \alpha^{(\beta)} \mid \beta \in X \,\}\,.$$

Proof. (a), (b) and (e) can be proved by a simple induction on γ. (c) follows from Lemma 1.7, while (d), (f) and (g) are immediate consequences of Corollary 4.18. □

Cantor normal form

We can apply the logarithm to decompose every ordinal in a canonical way.

Theorem 4.21. *For all ordinals* $\alpha, \beta \in \mathrm{On}$ *with* $\beta > 1$*, there are unique finite sequences* $(\gamma_i)_{i<n}$ *and* $(\eta_i)_{i<n}$ *of ordinal numbers such that*

$$\alpha = \beta^{(\eta_0)}\gamma_0 + \dots + \beta^{(\eta_{n-1})}\gamma_{n-1}\,,$$

$$\eta_0 > \dots > \eta_{n-1}\,, \quad \textit{and} \quad 0 < \gamma_i < \beta\,, \quad \textit{for } i < n\,.$$

Proof. We decompose α successively with the help of Lemma 4.19. We start by writing $\alpha = \beta^{(\eta_0)}\gamma_0 + \rho_0$. Applying the lemma to ρ_0 we get $\rho_0 = \beta^{(\eta_1)}\gamma_1 + \rho_1$. By induction on i, we obtain $\rho_i = \beta^{(\eta_{i+1})}\gamma_{i+1} + \rho_{i+1}$. If this process did not terminate then we would get an infinite decreasing sequence $\alpha > \rho_0 > \rho_1 > \dots$ of ordinals which is impossible. Consequently, there is some number n such that $\rho_n = 0$ and we have

$$\alpha = \beta^{(\eta_0)}\gamma_0 + \dots + \beta^{(\eta_{n-1})}\gamma_{n-1}\,.$$

□

Definition 4.22. Let α be an ordinal. The unique decomposition

$$\alpha = \omega^{(\eta_0)}\gamma_0 + \dots + \omega^{(\eta_n)}\gamma_n\,,$$

with $\eta_0 > \dots > \eta_n$ and $0 < \gamma_i < \omega$, for $i \leq n$.

is called the *Cantor normal form* of α.

The Cantor normal form is very convenient for ordinal calculations. Let us see how this is done. We start with addition.

Lemma 4.23. $\alpha < \beta$ *implies* $\omega^{(\alpha)} + \omega^{(\beta)} = \omega^{(\beta)}$.

Proof. Suppose that $\beta = \alpha + \gamma$, for $\gamma > 0$. We have

$$\begin{aligned}\omega^{(\alpha)} + \omega^{(\beta)} &= \omega^{(\alpha)} + \omega^{(\alpha+\gamma)} \\ &= \omega^{(\alpha)} + \omega^{(\alpha)}\omega^{(\gamma)} \\ &= \omega^{(\alpha)}\big(1 + \omega^{(\gamma)}\big) \\ &= \omega^{(\alpha)}\omega^{(\gamma)} \\ &= \omega^{(\alpha+\gamma)} = \omega^{(\beta)}.\end{aligned}$$ □

Corollary 4.24. *Let $\alpha, \beta \in \mathrm{On}$ be ordinals with Cantor normal form*

$$\begin{aligned}\alpha &= \omega^{(\eta_0)}k_0 + \dots + \omega^{(\eta_{m-1})}k_{m-1}\,, \\ \beta &= \omega^{(\gamma_0)}l_0 + \dots + \omega^{(\gamma_{n-1})}l_{n-1}\,.\end{aligned}$$

If i is the maximal index such that $\eta_i \geq \gamma_0$ then we have

$$\alpha + \beta = \omega^{(\eta_0)}k_0 + \dots + \omega^{(\eta_i)}k_i + \omega^{(\gamma_0)}l_0 + \dots + \omega^{(\gamma_{n-1})}l_{n-1}\,.$$

Lemma 4.25. *An ordinal $\alpha > 0$ is of the form $\alpha = \omega^{(\eta)}$, for some η, if, and only if, $\beta + \gamma < \alpha$, for all $\beta, \gamma < \alpha$.*

Proof. ($\Rightarrow$) Let

$$\beta = \omega^{(\rho_m)}k_m + \dots + \omega^{(\rho_0)}k_0 \quad \text{and} \quad \gamma = \omega^{(\sigma_n)}l_n + \dots + \omega^{(\sigma_0)}l_0$$

be the Cantor normal forms of β and γ. If $\beta, \gamma < \omega^{(\eta)}$ then $\rho_m, \sigma_n < \eta$. By symmetry, we may assume that $\gamma \leq \beta$. Thus,

$$\begin{aligned}\beta + \gamma &\leq \beta + \beta \\ &= \omega^{(\rho_n)}(k_m + k_m) + \omega^{(\rho_{m-1})}k_{m-1} + \dots + \omega^{(\rho_0)}k_0 \\ &< \omega^{(\eta)}.\end{aligned}$$

($\Leftarrow$) Suppose that $\alpha = \omega^{(\eta)}k + \rho$ where $k < \omega$ and $\rho < \omega^{(\eta)}$. We have to show that $k = 1$ and $\rho = 0$.

If $k > 1$, we set $\beta := \omega^{(\eta)}(k-1) + \rho < \alpha$. It follows that

$$\beta + \beta = \omega^{(\eta)}(k + (k-2)) + \rho \geq \omega^{(\eta)}k + \rho = \alpha\,.$$

Contradiction.

Suppose that $k = 1$ but $\rho > 0$. In this case we can set $\beta := \omega^{(\eta)}$ and we have

$$\beta + \beta = \omega^{(\eta)} + \omega^{(\eta)} > \omega^{(\eta)} + \rho = \alpha\,.$$

Again a contradiction. □

The next two lemmas provide the laws of multiplication and exponentiation of ordinals in Cantor normal form.

Lemma 4.26. *If* $\gamma > 0$, $0 \leq \rho < \omega^{(\eta)}$, *and* $0 < k < \omega$ *then*

$$(\omega^{(\eta)}k + \rho)\omega^{(\gamma)} = \omega^{(\eta+\gamma)}.$$

Proof. We have

$$\begin{aligned}\omega^{(\eta)}\omega^{(\gamma)} &\leq \big(\omega^{(\eta)}k + \rho\big)\omega^{(\gamma)} \\ &\leq \big(\omega^{(\eta)}(k+1)\big)\omega^{(\gamma)} \\ &= \omega^{(\eta)}\big((k+1)\omega^{(\gamma)}\big) = \omega^{(\eta)}\omega^{(\gamma)}.\end{aligned}$$

□

Lemma 4.27. *If* $\gamma, \eta > 0$, $0 \leq \rho < \omega^{(\eta)}$, *and* $0 < k < \omega$ *then*

$$(\omega^{(\eta)}k + \rho)^{(\omega^{(\gamma)})} = \omega^{(\eta\omega^{(\gamma)})}.$$

Proof. We have

$$\begin{aligned}\omega^{(\eta\omega^{(\gamma)})} &= (\omega^{(\eta)})^{(\omega^{(\gamma)})} \\ &\leq (\omega^{(\eta)}k + \rho)^{(\omega^{(\gamma)})} \\ &\leq (\omega^{(\eta+1)})^{(\omega^{(\gamma)})} \\ &= \omega^{((\eta+1)\omega^{(\gamma)})} = \omega^{(\eta\omega^{(\gamma)})}.\end{aligned}$$

□

Example. By the above lemmas we have

$$
\begin{aligned}
&\left(\omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(5)}\right)^{(\omega^{(2)}2+\omega+1)} \\
&= \left(\omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(5)}\right)^{(\omega^{(2)}2)} \cdot \left(\omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(5)}\right)^{(\omega)} \cdot \\
&\quad \cdot \left(\omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(5)}\right) \\
&= \left(\omega^{((\omega^{(5)}+\omega 4+2)\omega^{(2)})}\right)^{(2)} \cdot \omega^{((\omega^{(5)}+\omega 4+2)\omega)} \cdot \left(\omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(5)}\right) \\
&= \left(\omega^{(\omega^{(7)})}\right)^{(2)} \cdot \omega^{(\omega^{(6)})} \cdot \left(\omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(5)}\right) \\
&= \omega^{(\omega^{(7)}2)} \cdot \omega^{(\omega^{(6)})} \cdot \left(\omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(5)}\right) \\
&= \omega^{(\omega^{(7)}2+\omega^{(6)})} \cdot \left(\omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(5)}\right) \\
&= \omega^{(\omega^{(7)}2+\omega^{(6)})} \cdot \omega^{(\omega^{(5)}+\omega 4+2)} + \omega^{(\omega^{(7)}2+\omega^{(6)})} \cdot \omega^{(5)} \\
&= \omega^{(\omega^{(7)}2+\omega^{(6)}+\omega^{(5)}+\omega 4+2)} + \omega^{(\omega^{(7)}2+\omega^{(6)}+5)}.
\end{aligned}
$$

Exercise 4.5. Compute the cantor normal form of

$$\left(\omega^{(\omega^{(2)}7+\omega 3+4)}3 + \omega^{(\omega 6+3)}4 + \omega^{(4)}3 + 1\right)^{(\omega^{(2)}5+\omega 7+2)}$$

Remark. We will prove in Lemma A4.5.6 that we can find, for every β, arbitrarily large ordinals $\alpha_0, \alpha_1, \alpha_2$ such that

$$\alpha_0 = \beta + \alpha_0\,, \quad \alpha_1 = \beta\alpha_1\,, \quad \text{and} \quad \alpha_2 = \beta^{(\alpha_2)}.$$

In particular, there are ordinals ε such that $\varepsilon = \omega^{(\varepsilon)}$. By ε_α we denote the α-th ordinal such that $\beta^{(\varepsilon_\alpha)} = \varepsilon_\alpha$, for all $\beta < \varepsilon_\alpha$. Note that the Cantor normal form of ε_α is $\varepsilon_\alpha = \omega^{(\varepsilon_\alpha)}$.

Let us summarise the picture of On that we have obtained. The first ordinals are

$$\begin{aligned}
&0, 1, 2, 3, \dots\\
&\dots, \omega, \omega + 1, \omega + 2, \dots\\
&\dots, \omega 2, \omega 2 + 1, \omega 2 + 2, \dots\\
&\dots, \omega 3, \dots, \omega 4, \dots, \omega^{(2)}, \dots, \omega^{(3)}, \dots\\
&\dots, \omega^{(\omega)}, \dots, \omega^{(\omega^{(\omega)})}, \dots\\
&\dots, \varepsilon_0, \dots, \varepsilon_0^{(\varepsilon_0)}, \dots, \varepsilon_1, \dots, \varepsilon_2, \dots, \varepsilon_\omega, \dots\\
&\dots, \omega_1, \dots, \omega_2, \dots, \omega_\omega, \dots
\end{aligned}$$

The ordinals ω_α will be defined in Section A4.2.

A4. Zermelo-Fraenkel set theory

1. The Axiom of Choice

We have seen that induction is a powerful technique to prove statements and to construct objects. But in order to use this tool we have to relate the sets we are interested in to ordinals. In basic set theory this is not always possible. Therefore, we will introduce a new axiom which states that, for every set A, there is a well-order over A. Before doing so, let us present several statements that are equivalent to this axiom. We need two new notions.

Definition 1.1. A set $F \subseteq \wp(A)$ has *finite character* if, for all sets $x \subseteq A$, we have

$$x \in F \quad \text{iff} \quad x_0 \in F\,, \text{ for every finite set } x_0 \subseteq x\,.$$

Lemma 1.2. *Suppose that $F \subseteq \wp(A)$ has finite character.*

(a) *F is an initial segment of $\wp(A)$.*

(b) *If $X \subseteq F$ is nonempty then $\bigcap X \in F$.*

(c) *If $C \subseteq F$ is a chain and $\bigcup C$ is a set then $\bigcup C \in F$.*

Proof. (a) follows immediately from the definition and (b) is a consequence of (a). For (c), let $C \subseteq F$ be a chain such that $X := \bigcup C$ is a set. If $X_0 \subseteq X$ is finite, there exists some element $Z \in C$ with $X_0 \subseteq Z \in F$. Hence, $X_0 \in F$, for all finite subsets $X_0 \subseteq X$. This implies that $X \in F$. □

Lemma 1.3. *If F has finite character then $\langle F, \subseteq \rangle$ is inductively ordered.*

Proof. Let $C \subseteq F$ be a linearly ordered subset of F. By Corollary A2.3.10 and Lemma 1.2 (c), it follows that $\sup C = \bigcup C \in F$. □

Example. Let V be a vector space over the field K. The set

$$F := \{ B \subseteq V \mid B \text{ is linearly independent} \}$$

has finite character.

The second notion we need is that of a choice function. Intuitively, a choice function is a function that, given some set A, selects an element of A.

Definition 1.4. A function f is a *choice function* if $f(a) \in a$, for all $a \in \operatorname{dom} f$.

Exercise 1.1. Let $\mathcal{I}$ be the set of all open intervals (a, b) of real numbers $a, b \in \mathbb{R}$ with $a < b$. Define a choice function $\mathcal{I} \to \mathbb{R}$.

Lemma 1.5. *Let A be a set and C the set of all choice functions f with* $\operatorname{dom} f \subseteq \wp(A)$.

(a) *C has finite character.*

(b) *If f is a $\subseteq$-maximal element of C then* $\operatorname{dom} f = \wp(A) \setminus \{\varnothing\}$.

Proof. (a) Suppose that f is a binary relation such that every finite $f_0 \subseteq f$ is a choice function. If $\langle a, b\rangle, \langle a, c\rangle \in f$ then $\{\langle a, b\rangle, \langle a, c\rangle\} \in C$ implies that $b = c$. Hence, f is a partial function. Furthermore, if $\langle a, b\rangle \in f$ then $\{\langle a, b\rangle\} \in C$ implies that $b \in a$. Consequently, f is a choice function.

(b) Let $f \in C$ be $\subseteq$-maximal. Since f is a choice function we have $\varnothing \notin \operatorname{dom} f$. Therefore, $\operatorname{dom} f \subseteq \wp(A) \setminus \{\varnothing\}$. Suppose that there is some element $B \in (\wp(A) \setminus \{\varnothing\}) \setminus \operatorname{dom} f$. Since $B \neq \varnothing$ we can choose some element $b \in B$. The relation $f \cup \{\langle B, b\rangle\} \supset f$ is again a choice function in contradiction to the maximality of f. □

Lemma 1.6. *Let A be a set. Given a choice function $f : \wp(A) \setminus \{\varnothing\} \to A$ we can define a well-order R on A.*

Proof. Let $f : \wp(A) \setminus \{\varnothing\} \to A$ be a choice function. We define a function $g : \wp(A) \to \wp(A)$ by

$$g(X) := \begin{cases} A & \text{if } X = A\,, \\ X \cup \{f(A \setminus X)\} & \text{if } X \neq A\,. \end{cases}$$

Since $g(X) \supseteq X$ this function is inflationary. Furthermore, the partial order $\langle \wp(A), \subseteq \rangle$ is complete. By Theorem A3.3.14, g has an inductive fixed point. Since $g(X) \neq X$, for $X \neq A$, it follows that this fixed point is A. Let $G : \mathrm{On} \to \wp(A)$ be the fixed-point induction of g over $\varnothing$ and let α be the closure ordinal. For every $\beta < \alpha$, there exists a unique element a_β such that $G(\beta + 1) \setminus G(\beta) = \{a_\beta\}$. We define a function, $h : {\downarrow}\alpha \to A$ by $h(\beta) := a_\beta$. Since $G(0) = \varnothing$ it follows that $\operatorname{rng} h = G(\infty) = A$. Hence, $h : {\downarrow}\alpha \to A$ is bijective and we can define the desired well-order R over A by

$$R := \{ \langle a, b \rangle \mid h^{-1}(a) \leq h^{-1}(b) \}\,. \qquad \square$$

Each of the following statements cannot be proved in basic set theory.

Theorem 1.7. *The following statements are equivalent:*

(1) *For every set A, there exists a well-order R over A.*

(2) *For every set A, there exists a choice function* $f : \wp(A) \setminus \{\varnothing\} \to A$.

(3) *If* $(A_i)_{i \in I}$ *is a sequence of nonempty sets then* $\prod_{i \in I} A_i \neq \varnothing$.

(4) *If* $(A_i)_{i \in I}$ *is a sequence of disjoint nonempty sets then* $\prod_{i \in I} A_i \neq \varnothing$.

(5) *Every inductively ordered partial order has a maximal element.*

(6) *If F is a set of finite character and* $A \in F$*, there exists a maximal element* $B \in F$ *with* $A \subseteq B$.

(7) *For all sets A and B, there exists an injective function* $f : A \to B$ *or an injective function* $f : B \to A$.

(8) *For every surjective function* $f : A \to B$ *where A is a set, there exists a function* $g : B \to A$ *such that* $f \circ g = \mathrm{id}_B$.

Proof. (2) ⇒ (3) If $\prod_{i \in I} A_i$ is a proper class, it is nonempty and we are done. Hence, we may assume that it is a set. Then $A := \bigcup \{ A_i \mid i \in I \}$ is also a set. By (2) there exists a choice function $f : \wp(A) \setminus \{\emptyset\} \to A$. Let $g : I \to A$ be the function defined by $g(i) := f(A_i)$. Since $g(i) \in A_i$ it follows that $g \in \prod_{i \in I} A_i \neq \emptyset$.

(3) ⇒ (4) is trivial.

(4) ⇒ (2) Let $I := \wp(A) \setminus \{\emptyset\}$ and set $A_X := X \times \{X\}$, for $X \in I$. Since $\prod_{X \in I} A_X \neq \emptyset$ there exists some element $f \in \prod_{X \in I} A_X$. We can define the desired choice function $g : \wp(A) \setminus \{\emptyset\} \to A$ by

$$g(X) = a \quad :\text{iff} \quad f(X) = \langle a, X \rangle \,.$$

(2) ⇒ (1) was proved in Lemma 1.6.

(1) ⇒ (5) Suppose that $\langle A, \leq \rangle$ is inductively ordered, but A has no maximal element. For every $a \in A$, we can find some $b \in A$ with $b > a$. By assumption, there exists a well-order R over A. Let $f : A \to A$ be the function such that $f(a)$ is the R-minimal element $b \in A$ with $b > a$. By definition, we have $f(a) > a$, for all $a \in A$. Hence, f is inflationary and, by Theorem A3.3.14, f has a fixed point a. But $f(a) = a$ contradicts the definition of f.

(5) ⇒ (6) Let F be a set of finite character and $A \in F$. It is sufficient to prove that the subset $F_0 := \{ X \in F \mid A \subseteq X \}$ is inductively ordered by $\subseteq$. By Lemma 1.3, we know that $\langle F, \subseteq \rangle$ is inductively ordered. Let C be a chain in F_0. Then $C \subseteq F_0 \subseteq F$ and C is also a chain in F. Consequently, it has a least upper bound $B \in F$. Since $A \subseteq X$, for all $X \in C$, it follows that $A \subseteq B$, that is, $B \in F_0$ and B is also the least upper bound of C in F_0.

(6) ⇒ (2) Let A be a set. By Lemma 1.5 (a), the set C of choice functions f with $\operatorname{dom} f \subseteq \wp(A) \setminus \{\emptyset\}$ has finite character and, therefore, there is a maximal element $f \in C$. By Lemma 1.5 (b), it follows that f is the desired choice function.

(1) ⇒ (7) Fix well-orders R and S on, respectively, A and B. By Corollary A3.1.12, exactly one of the following conditions is satisfied:

$$\langle A, R \rangle < \langle B, S \rangle \quad \text{or} \quad \langle A, R \rangle \cong \langle B, S \rangle \quad \text{or} \quad \langle A, R \rangle > \langle B, S \rangle.$$

In the first two cases there exists an injection $A \to B$ and in the second and third case there exists an injection $B \to A$ in the other direction.

(7) ⇒ (1) Let A be a set. By Theorem A3.2.12, there exists an ordinal α such that there is no injective function $\downarrow\alpha \to A$. Consequently, there exists an injective function $f : A \to \downarrow\alpha$. We define a relation R on A by

$$R := \{ \langle a, b \rangle \mid f(a) < f(b) \} .$$

Since f is injective and $\operatorname{rng} f \subseteq \downarrow\alpha$ is well-ordered it follows that R is the desired well-order on A.

(2) ⇒ (8) Let $h : \wp(A) \setminus \{\emptyset\} \to A$ be a choice function. We can define $g : B \to A$ by

$$g(b) := h(f^{-1}(b)) .$$

(8) ⇒ (4) Let $(A_i)_{i \in I}$ be a family of disjoint nonempty sets. We define a function $f : \bigcup \{ A_i \mid i \in I \} \to I$ by

$$f(a) = i \quad :\text{iff} \quad a \in A_i .$$

Since the A_i are disjoint and nonempty it follows that f is well-defined and surjective. Hence, there exists a function $g : I \to \bigcup \{ A_i \mid i \in I \}$ such that $f(g(i)) = i$, for all $i \in I$. By definition of f, this implies that $g(i) \in A_i$. Hence, $g \in \prod_{i \in I} A_i \neq \emptyset$. □

Axiom of Choice. *For every set A there exists a well-order R over A.*

Lemma 1.8. *A left-narrow partial order $(A, \leq)$ is well-founded if, and only if, there exists no infinite strictly decreasing sequence $a_0 > a_1 > \dots$.*

Proof. One direction was already proved in Lemma A3.1.3. For the other one, fix a choice function $f : \wp(A) \setminus \emptyset \to A$. Suppose that there exists a nonempty set $A_0 \subseteq A$ without minimal element. We can define a descending chain $a_0 > a_1 > \dots$ by induction. Let $a_0 := f(A_0)$ and, for $k > 0$, set

$$a_k := f(\{ b \in A_0 \mid b < a_{k-1} \}) .$$

Note that a_k is well-defined since a_{k-1} is not a minimal element of A_0. □

Exercise 1.2. We call a set *a countable* if there exists a bijection $\downarrow\omega \to a$. Prove that a left-narrow partial order $\langle A, \leq \rangle$ is well-founded if, and only if, every countable nonempty subset $X \subseteq A$ has a minimal element.

Exercise 1.3. Let $\langle A, R \rangle$ be a well-founded partial order that is a set. Prove that there exists a well-order $\leq$ on A with $R \subseteq \leq$.

The following variant of the Axiom of Choice (statement (5) in the above theorem) is known as 'Zorn's Lemma'.

Lemma 1.9 (Kuratowski, Zorn). *Every inductively ordered partial order has a maximal element.*

Example. We have seen that the system of all linearly independent subsets of a vector space V is inductively ordered. It follows that every vector space contains a maximal linearly independent subset, that is, a basis.

This example can be generalised to a certain kind of closure operators.

Definition 1.10. Let c be a closure operator on A.

(a) c has the *exchange property* if

$$b \in c(X \cup \{a\}) \setminus c(X) \quad \text{implies} \quad a \in c(X \cup \{b\}) .$$

(b) A set $I \subseteq A$ is *c-independent* if

$$a \notin c(I \setminus \{a\}) , \qquad \text{for all } a \in I .$$

We call $D \subseteq A$ *c-dependent* if it is not *c-independent.*

(c) Let $X \subseteq A$. A set $I \subseteq X$ is a *c-basis* of X if I is c-independent and $c(I) = c(X)$.

Lemma 1.11. *Let c be a closure operator on A and let $F \subseteq \mathcal{P}(A)$ be the class of all c-independent sets. If c has finite character then F has finite character.*

Proof. Let $I \in F$ and $I_0 \subseteq I$. For every $a \in I_0$, we have

$$a \notin c(I \setminus \{a\}) \supseteq c(I_0 \setminus \{a\}) .$$

Hence, I_0 is c-independent. Conversely, suppose that $I \notin F$. Then there is some $a \in I$ with

$$a \in c(I \setminus \{a\}) .$$

Since c has finite character we can find a finite subset $I_0 \subseteq I \setminus \{a\}$ with $a \in c(I_0)$. Thus, $I_0 \cup \{a\}$ is a finite subset of I that is not c-independent. □

Before proving the converse let us show with the help of the Axiom of Choice that there is always a c-basis. We start with an alternative description of the exchange property.

Lemma 1.12. *Let c be a closure operator on A with the exchange property. If $D \subseteq A$ is a minimal c-dependent set then*

$$a \in c(D \setminus \{a\}) , \quad \textit{for all } a \in D .$$

Proof. Let $a \in D$. Since D is c-dependent there exists some element $b \in D$ with $b \in c(D \setminus \{b\})$. If $b = a$ then we are done. Hence, suppose that $b \neq a$ and let $D_0 := D \setminus \{a, b\}$. By minimality of D we have $b \notin c(D_0)$. Hence, $b \in c(D_0 \cup \{a\}) \setminus c(D_0)$ and the exchange property implies that $a \in c(D_0 \cup \{b\})$. □

Proposition 1.13. *Let c be a closure operator on A that has finite character and the exchange property. Every set $X \subseteq A$ has a c-basis.*

Proof. The family F of all c-independent subsets of X has finite character. By the Axiom of Choice, there exists a maximal c-independent set $I \subseteq X$. We claim that $c(I) = c(X)$, that is, I is a c-basis of X.

Clearly, $c(I) \subseteq c(X)$. If $X \subseteq c(I)$, it follows that

$$c(X) \subseteq c(c(I)) = c(I)$$

and we are done. Hence, it remains to consider the case that there is some element $a \in X \smallsetminus c(I)$. We derive a contradiction to the maximality of I by showing that $I \cup \{a\}$ is c-independent.

Suppose that $I \cup \{a\}$ is not c-independent. Since F has finite character there exists a finite c-dependent subset $D \subseteq I \cup \{a\}$ with $a \in D$. Suppose that D is chosen minimal. By Lemma 1.12, it follows that $a \in c(D \smallsetminus \{a\}) \subseteq c(I)$. A contradiction. □

Proposition 1.14. *Let c be a closure operator on A with the exchange property and let $F \subseteq \wp(A)$ be the class of all c-independent sets. Then c has finite character if, and only if, F has finite character.*

Proof. (⇒) has already been proved in Lemma 1.11.

(⇐) For a contradiction, suppose that there is a set $X \subseteq A$ such that

$$Z := \bigcup \{\, c(X_0) \mid X_0 \subseteq X \text{ is finite} \,\}$$

is a proper subset of $c(X)$. Fix some element $a \in c(X) \smallsetminus Z$. By Proposition 1.13 there exists a c-basis I for X. It follows that $a \in c(X) = c(I)$. Since F has finite character we can find a finite subset $I_0 \subseteq I$ such that $I_0 \cup \{a\}$ is c-dependent. By Lemma 1.12, it follows that $a \in c(I_0) \subseteq Z$. A contradiction. □

A more extensive treatment of closure operators with the exchange property will be given in Section F1.1.

2. *Cardinals*

The notion of the cardinality of a set is a very natural one. It is based on the same idea which led to the definition of the order type of a well-order. But instead of well-orders we consider just sets without any relation. Although conceptually simpler than ordinals we introduce cardinals quite late in the development of our theory since most of their properties cannot be proved without resorting to ordinals and the Axiom of Choice.

Intuitively, the cardinality of a set A measures its size, that is, the number of its elements. So, how do we count the elements of a set? We can say that 'A has α elements' if there exists an enumeration of A of length α, that is, a bijection $\downarrow\alpha \to A$. For infinite sets, such an enumeration is not unique. We can find several sequences $\downarrow\alpha \to A$ with different values of α. To get a well-defined number we therefore pick the least one.

Definition 2.1. The *cardinality* $|A|$ of a class A is the least ordinal α such that there exists a bijection $\downarrow\alpha \to A$. If there exists no such ordinal then we write $|A| := \infty$. Let $\mathrm{Cn} := \mathrm{rng}\,|\cdot| \subseteq \mathrm{On}$ be the range of this mapping. (We do not consider ∞ to be an element of the range.) We set $\mathfrak{Cn} := \langle \mathrm{Cn}, \leq \rangle$. The elements of Cn are called *cardinals.*

Remark. Clearly, if $|A|, |B| < \infty$ then we have $|A| = |B|$ iff there exists a bijection $A \to B$.

Lemma 2.2. *Every set A has a cardinality and we have $|A| < \infty$.*

Proof. Let A be a set. By the Axiom of Choice, we can find a well-order R over A. Set $\alpha := \mathrm{ord}\,\langle A, R \rangle$. By definition of an ordinal, there exists a bijection $\downarrow\alpha \to A$. In particular, the class of all ordinals β such that there exists a bijection $\downarrow\beta \to A$ is nonempty and, therefore, there exists a least such ordinal. □

Lemma 2.3. *Let A and B be nonempty sets. The following statements are equivalent:*

(1) $|A| \leq |B|$

(2) *There exists an injective function $A \to B$.*

(3) *There exists an surjective function $B \to A$.*

Proof. Set $\kappa := |A|$ and $\lambda := |B|$ and let $g : \downarrow\kappa \to A$ and $h : \downarrow\lambda \to B$ be the corresponding bijections.

(1) ⇒ (2) Since $\kappa \leq \lambda$ there exists an isomorphism $f : \downarrow\kappa \to I$ between $\downarrow\kappa$ and an initial segment $I \subseteq \downarrow\lambda$. In particular, f is injective. The composition $h \circ f \circ g^{-1} : A \to B$ is the desired injective function.

(2) ⇒ (1) For a contradiction, suppose that there exists an injective function $A \to B$ but we have $|A| > |B|$. By (1) ⇒ (2), the latter implies that there is an injective function $B \to A$. Hence, applying Theorem A2.1.12 we find a bijection $A \to B$. It follows that $|A| = |B|$. Contradiction.

(2) ⇒ (3) Let $f : A \to B$ be injective. By Lemma A2.1.10 (b), there exists a function $g : B \to A$ such that $g \circ f = \mathrm{id}_A$. Furthermore, it follows by Lemma A2.1.10 (d) that g is surjective.

(3) ⇒ (2) As above, given a surjective function $f : B \to A$ we can apply Lemma A2.1.10 (and the Axiom of Choice) to obtain an injective function $g : A \to B$ with $f \circ g = \mathrm{id}_B$. □

For every cardinal, there is a canonical set with this cardinality.

Lemma 2.4. *For every cardinal* $\kappa \in \mathrm{Cn}$, *we have* $\kappa = |{\downarrow}\kappa|$. *It follows that* $\mathrm{Cn} = \{\, \alpha \in \mathrm{On} \mid |{\downarrow}\alpha| = \alpha \,\}$.

Exercise 2.1. Let α and β be ordinals such that $|\alpha| \leq \beta \leq \alpha$. Show that $|\alpha| = |\beta|$.

Exercise 2.2. Prove that $\alpha \in \mathrm{Cn}$, for every ordinal $\alpha \leq \omega$. *Hint.* Show, by induction on α, that there is no surjective function ${\downarrow}\alpha \to {\downarrow}\beta$ with $\alpha < \beta \leq \omega$.

Using the notion of cardinality we can restate Theorem A2.1.13 in the following way.

Theorem 2.5. *We have* $|A| < |\wp(A)|$, *for every set A.*

Proof. By Theorem A2.1.13, there exists an injective function $A \to \wp(A)$ but no surjective one. By Lemma 2.3, it follows that $|A| \leq |\wp(A)|$ and $|\wp(A)| \nleq |A|$. □

Cn is a proper class since it is an unbounded subclass of On.

Lemma 2.6. Cn *is a proper class.*

Proof. For a contradiction, suppose otherwise. By Lemma A3.2.8, it follows that there is some $\alpha \in \mathrm{On}$ such that $\kappa < \alpha$, for all cardinals κ. But, by Theorem A3.2.12, there exists some ordinal β such that $\lambda := |{\downarrow}\beta| > |{\downarrow}\alpha|$, which implies that $\lambda > \alpha$. A contradiction. □

Lemma 2.7. $\mathfrak{On}_0 \leq \mathfrak{Cn} \leq \mathfrak{On}$.

Proof. Since $\mathrm{Cn} \subseteq \mathrm{On}$ it follows that $\mathfrak{Cn}$ is a well-order. Therefore, there exists an isomorphism $h : \mathrm{Cn} \to I$, for some initial segment $I \subseteq \mathrm{On}$.

By Theorem 2.5 we know that the function $f : \mathrm{On}_0 \to \mathrm{Cn}$ with $f(\alpha) := |S_\alpha|$ is strictly increasing. Consequently, we have $\mathfrak{On}_0 \leq \mathfrak{Cn}$, by Lemma A3.2.11. □

Remark. With the Axiom of Replacement which we will introduce in Section 5 we can actually prove that $\langle \mathrm{On}_0, \in \rangle \cong \langle \mathrm{On}, < \rangle$. Therefore, all three orders are isomorphic.

Definition 2.8. (a) By the preceding lemma and Lemma A3.1.10, there exists a unique isomorphism $h : I \to \mathrm{Cn}$ where I is an initial segment of On. We define $\aleph_\alpha := h(\omega + \alpha)$ (*'aleph alpha'*), for all α such that $\omega + \alpha \in I$. Furthermore, we denote by ω_α the minimal ordinal such that $|\omega_\alpha| = \aleph_\alpha$.

(b) A set A is *finite* if $|A| < \aleph_0$. Otherwise, A is called *infinite*. Similarly, we say that A is *countable* if $|A| \leq \aleph_0$, and A is *uncountable*, if $|A| > \aleph_0$. A countable set that is not finite is called *countably infinite*.

(c) For cardinals κ, we will denote by κ^+ the minimal *infinite* cardinal greater than κ.

Note that, by our definition of a cardinal, we have $\omega_\alpha = \aleph_\alpha$ and $\aleph_0 = \omega_0 = \omega$. Furthermore, $\aleph_\alpha^+ = \aleph_{\alpha+1}$. Since we have defined the operation κ^+ differently for cardinals and ordinals we will use this notation only for cardinals in the remainder of this book. If we consider the successor of an ordinal α we will write $\alpha + 1$.

3. Cardinal arithmetic

Similarly to ordinals we can define arithmetic operations on cardinals. Note that, except for finite cardinals, these operations are different from the ordinal operations. Therefore, we have chosen different symbols to denote them.

Definition 3.1. Let $\kappa, \lambda \in \mathrm{Cn}$ be cardinals. We define

$$\kappa \oplus \lambda := |{\downarrow}\kappa \cup {\downarrow}\lambda|\,, \quad \kappa \otimes \lambda := |{\downarrow}\kappa \times {\downarrow}\lambda|\,, \quad \kappa^\lambda := |{\downarrow}\kappa^{{\downarrow}\lambda}|\,.$$

The following lemmas follows immediately from the definition if one recalls that, for $\kappa := |A|$ and $\lambda := |B|$, there exist bijections $A \to {\downarrow}\kappa$ and $B \to {\downarrow}\lambda$.

Lemma 3.2. *Let A and B be sets.*

$$|A \cup B| = |A| \oplus |B|\,, \quad |A \times B| = |A| \otimes |B|\,, \quad |A^B| = |A|^{|B|}.$$

Corollary 3.3. *For all $\alpha, \beta \in \mathrm{On}$, we have*

$$|{\downarrow}(\alpha + \beta)| = |{\downarrow}\alpha| \oplus |{\downarrow}\beta| \quad \textit{and} \quad |{\downarrow}(\alpha\beta)| = |{\downarrow}\alpha| \otimes |{\downarrow}\beta|\,.$$

The corresponding equation for ordinal exponentiation will be delayed until Lemma 4.4.

Exercise 3.1. Prove that, if A is a set then $|\wp(A)| = 2^{|A|}$. *Hint.* Take the obvious bijection $\wp(A) \to 2^A$.

For finite cardinals these operations coincide with the usual ones.

Lemma 3.4. *For $m, n < \omega$, we have*

$$m \oplus n = m + n\,, \quad m \otimes n = mn\,, \quad m^n = m^n\,,$$

where the operations on the left are the ones defined above while those on the right are the usual arithmetic operations.

Let us summarise the basic properties of cardinal arithmetic. The proofs are similar to, but much simpler than, the corresponding ones for ordinal arithmetic.

Lemma 3.5. *Let* $\kappa, \lambda, \mu \in \mathrm{Cn}$.

(a) $(\kappa \oplus \lambda) \oplus \mu = \kappa \oplus (\lambda \oplus \mu)$

(b) $\kappa \oplus \lambda = \lambda \oplus \kappa$

(c) $\kappa \oplus 0 = \kappa$

(d) $\kappa \leq \lambda$ *if, and only if, there is some* μ *with* $\lambda = \kappa \oplus \mu$.

(e) $\lambda \leq \mu$ *implies* $\kappa \oplus \lambda \leq \kappa \oplus \mu$.

(f) $\kappa \geq \aleph_0$ *if, and only if,* $\kappa \oplus 1 = \kappa$

Proof. (a) There is a canonical bijection $(A \uplus B) \uplus C \to A \uplus (B \uplus C)$ with

$$\begin{aligned} \langle 0, \langle 0, a\rangle\rangle &\mapsto \langle 0, a\rangle\,, \\ \langle 0, \langle 1, b\rangle\rangle &\mapsto \langle 1, \langle 0, b\rangle\rangle\,, \\ \langle 1, c\rangle &\mapsto \langle 1, \langle 1, c\rangle\rangle\,. \end{aligned}$$

(b) There is a canonical bijection $A \uplus B \to B \uplus A$ with $\langle 0, a\rangle \mapsto \langle 1, a\rangle$ and $\langle 1, b\rangle \mapsto \langle 0, b\rangle$.

(c) $A \uplus \varnothing = \{0\} \times A$. We can define a bijection $A \to \{0\} \times A$ by $a \mapsto \langle 0, a\rangle$.

(d) If $\kappa \leq \lambda$, there exists an injective function $f : {\downarrow}\kappa \to {\downarrow}\lambda$. Let $X := {\downarrow}\lambda \smallsetminus \operatorname{rng} f$ and $\mu := |X|$. We can define a bijection ${\downarrow}\kappa \uplus X \to {\downarrow}\lambda$ by

$$\langle 0, a\rangle \mapsto f(a) \quad \text{and} \quad \langle 1, a\rangle \mapsto a\,.$$

(e) If there is an injective function $f : B \to C$, we can define an injective function $A \uplus B \to A \uplus C$ by

$$\langle 0, a\rangle \mapsto \langle 0, a\rangle \quad \text{and} \quad \langle 1, b\rangle \mapsto \langle 1, f(b)\rangle\,.$$

(f) If $\kappa \geq \aleph_0 = \omega$ then $\kappa = \omega + \alpha$, for some $\alpha \in \mathrm{On}$. We can define a bijection ${\downarrow}\omega \to {\downarrow}(\omega + 1)$ by $0 \mapsto \omega$ and $n \mapsto n - 1$, for $n > 0$. This function can be extended to a bijection ${\downarrow}\omega \uplus {\downarrow}\alpha \to {\downarrow}\omega \uplus {\downarrow}\alpha \uplus [1]$. Conversely, if $\kappa < \omega$ then $\kappa \oplus 1 = \kappa + 1 > \kappa$. ☐

Lemma 3.6. *Let* $\kappa, \lambda, \mu \in \mathrm{Cn}$.

(a) $(\kappa \otimes \lambda) \otimes \mu = \kappa \otimes (\lambda \otimes \mu)$

(b) $\kappa \otimes \lambda = \lambda \otimes \kappa$

(c) $\kappa \otimes 0 = 0$, $\kappa \otimes 1 = \kappa$, $\kappa \otimes 2 = \kappa \oplus \kappa$.

(d) $\kappa \otimes (\lambda \oplus \mu) = (\kappa \otimes \lambda) \oplus (\kappa \otimes \mu)$

(e) $\lambda \leq \mu$ *implies* $\kappa \otimes \lambda \leq \kappa \otimes \mu$.

Proof. (a) There is a canonical bijection $(A \times B) \times C \to A \times (B \times C)$ with $\langle\langle a, b\rangle, c\rangle \mapsto \langle a, \langle b, c\rangle\rangle$.

(b) There is a canonical bijection $A \times B \to B \times A$ with $\langle a, b\rangle \mapsto \langle b, a\rangle$.

(c) $A \times \varnothing = \varnothing$. There are canonical bijections

$$A \times \{0\} \to A \quad \text{and} \quad A \cup A = [2] \times A \to A \times [2] .$$

(d) There exists a bijection $A \times (B \cup C) \to (A \times B) \cup (A \times C)$ with

$$\langle a, \langle 0, b\rangle\rangle \mapsto \langle 0, \langle a, b\rangle\rangle \quad \text{and} \quad \langle a, \langle 1, c\rangle\rangle \mapsto \langle 1, \langle a, c\rangle\rangle .$$

(e) Given an injective function $f : B \to C$ we define an injective function $A \times B \to A \times C$ by $\langle a, b\rangle \mapsto \langle a, f(b)\rangle$. □

Lemma 3.7. *Let* $\kappa, \lambda, \mu, \nu \in \mathrm{Cn}$.

(a) $(\kappa^\lambda)^\mu = \kappa^{\lambda \otimes \mu}$

(b) $(\kappa \otimes \lambda)^\mu = \kappa^\mu \otimes \lambda^\mu$

(c) $\kappa^{\lambda \oplus \mu} = \kappa^\lambda \otimes \kappa^\mu$

(d) $\kappa^0 = 1$, $\kappa^1 = \kappa$, $\kappa^2 = \kappa \otimes \kappa$.

(e) *If* $\kappa \leq \lambda$ *and* $\mu \leq \nu$ *then* $\kappa^\mu \leq \lambda^\nu$.

(f) $\kappa < 2^\kappa$

Proof. (a) There is a canonical bijection $(A^B)^C \to A^{B \times C}$ given by $f \mapsto g$ where $g(b, c) := f(c)(b)$.

(b) We define a bijection $A^C \times B^C \to (A \times B)^C$ by

$$\langle g, h \rangle \mapsto f \quad \text{where} \quad f(c) := \langle g(c), h(c) \rangle .$$

(c) We define a bijection $A^{B \cup C} \to A^B \times A^C$ by $f \mapsto \langle g, h \rangle$ where

$$g(b) := f(\langle 0, b \rangle) \quad \text{and} \quad h(c) := f(\langle 1, c \rangle) .$$

(d) $A^\varnothing = \{\varnothing\}$. A bijection $A^{[1]} \to A$ is given by $f \mapsto f(0)$, and a bijection $A^{[2]} \to A \times A$ by $f \mapsto \langle f(0), f(1) \rangle$.

(e) Suppose that $f : A \to B$ and $g : C \to D$ are injective. According to Lemma A2.1.10 (b), there exists a surjective function $g' : D \to C$ such that $g' \circ g = \mathrm{id}_C$. We define an injection $A^C \to B^D$ by $h \mapsto f \circ h \circ g'$. To show that this mapping is injective consider functions $h, h' \in A^C$ with $h \neq h'$. Fix some some $c \in C$ with $h(c) \neq h'(c)$ and set $d := g(c)$. Then $g'(d) = g'(g(c)) = \mathrm{id}_C(c) = c$. Since f is injective it follows that

$$(f \circ h \circ g')(d) = f(h(c)) \neq f(h'(c)) = (f \circ h' \circ g')(d) .$$

Consequently, $f \circ h \circ g' \neq f \circ h' \circ g'$.

(f) follows immediately from Theorem 2.5. □

We will show that addition and multiplication of infinite cardinals is especially simple since they just consist of taking the maximum of the operands. In particular, we have $\kappa \oplus \lambda = \kappa \otimes \lambda$ if at least one operand is infinite.

Exercise 3.2. Prove that $\aleph_0 \otimes \aleph_0 = \aleph_0$ by showing that the function

$$\downarrow\omega \times \downarrow\omega \to \downarrow\omega : \langle i, k \rangle \mapsto \tfrac{1}{2}(i + k)(i + k + 1) + k$$

is bijective.

We start by computing $\kappa \otimes \kappa$ by induction on $\kappa \geq \aleph_0$.

Theorem 3.8. *If $\kappa \geq \aleph_0$ then $\kappa \otimes \kappa = \kappa$.*

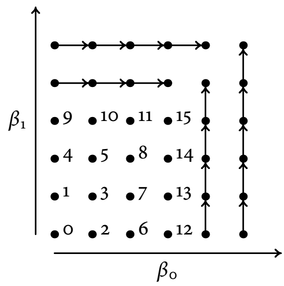


Figure 1.. Ordering on $\downarrow\kappa \times \downarrow\kappa$

Proof. We have $\kappa = \kappa \otimes 1 \leq \kappa \otimes \kappa$. For the converse, we prove that $\kappa \otimes \kappa \leq \kappa$ by induction on κ.

Note that, since κ is a cardinal we have $\alpha < \kappa$ if, and only if, $|\downarrow\alpha| < \kappa$, for all ordinals α. We define an order on $K := \downarrow\kappa \times \downarrow\kappa$ by

$$\begin{aligned} &\langle \beta_0, \beta_1 \rangle < \langle \gamma_0, \gamma_1 \rangle \\ :\text{iff}\quad &\max\{\beta_0, \beta_1\} < \max\{\gamma_0, \gamma_1\}, \text{ or} \\ &\max\{\beta_0, \beta_1\} = \max\{\gamma_0, \gamma_1\} \text{ and } \beta_0 < \gamma_0, \text{ or} \\ &\max\{\beta_0, \beta_1\} = \max\{\gamma_0, \gamma_1\} \text{ and } \beta_0 = \gamma_0 \text{ and } \beta_1 < \gamma_1 . \end{aligned}$$

One can check easily that this order is a well-order. For every ordinal $\alpha \leq \kappa$, the set

$$I(\alpha) := \downarrow\alpha \times \downarrow\alpha$$

is an initial subset of K. If $\omega \leq \alpha < \kappa$, it follows by inductive hypothesis that

$$|I(\alpha)| = |\downarrow\alpha \times \downarrow\alpha| = |\downarrow\alpha| \otimes |\downarrow\alpha| = |\downarrow\alpha| < \kappa .$$

Similarly, if $\alpha < \omega$ then we have

$$|I(\alpha)| = |\downarrow\alpha| \otimes |\downarrow\alpha| = |\downarrow\alpha|^2 = \alpha^2 < \aleph_0 \leq \kappa .$$

Hence, we have $\operatorname{ord} I(\alpha) < \kappa$, for all ordinals $\alpha < \kappa$.

We claim that $K = \bigcup \{ I(\alpha) \mid \alpha < \kappa \}$. Let $\langle \alpha, \beta \rangle \in K$. Since $\alpha, \beta < \kappa$ and κ is a limit ordinal we have $\gamma := \max \{\alpha + 1, \beta + 1\} < \kappa$ and $\langle \alpha, \beta \rangle \in I(\gamma)$. It follows that

$$\operatorname{ord} \langle K, \leq \rangle = \sup \{ \operatorname{ord} \langle I(\alpha), \leq \rangle \mid \alpha < \kappa \} \leq \kappa .$$

In particular, there exists an isomorphism between K and some initial segment of κ. This implies that $\kappa \otimes \kappa = |K| \leq \kappa$. □

The general case now follows easily.

Lemma 3.9. *If $\kappa > 0$ and $\lambda \geq \aleph_0$ then $\kappa \oplus \lambda = \kappa \otimes \lambda = \max \{\kappa, \lambda\}$.*

Proof. By symmetry, we may assume that $\kappa \leq \lambda$. For $\kappa = 1$, the claim follows from Lemmas 3.5 and 3.6. Suppose that $\kappa > 1$. Then

$$\lambda \leq \kappa \oplus \lambda \leq \lambda \oplus \lambda = 2 \otimes \lambda \leq \kappa \otimes \lambda \leq \lambda \otimes \lambda = \lambda . \qquad \square$$

Corollary 3.10. *If $\kappa \geq \aleph_0$ then $\kappa^n = \kappa$, for all $n < \omega$.*

Example. We have

$$\begin{aligned} \aleph_4^{\aleph_3} \otimes (\aleph_5 \oplus \aleph_4^{\aleph_7})^{\aleph_2} &= \aleph_4^{\aleph_3} \otimes (\aleph_4^{\aleph_7})^{\aleph_2} = \aleph_4^{\aleph_3} \otimes \aleph_4^{\aleph_7 \otimes \aleph_2} \\ &= \aleph_4^{\aleph_3} \otimes \aleph_4^{\aleph_7} = \aleph_4^{\aleph_3 \oplus \aleph_7} = \aleph_4^{\aleph_7} . \end{aligned}$$

4. *Cofinality*

Frequently, we will construct objects as the union of an increasing sequence $A_0 \subseteq A_1 \subseteq \dots$ of sets. In this section we will study the cardinality of such unions.

Definition 4.1. For a sequence $(\kappa_i)_{i<\alpha}$ of cardinals, we define

$$\sum_{i<\alpha} \kappa_i := \Big| \dot{\bigcup_{i<\alpha}} \downarrow\kappa_i \Big| \quad \text{and} \quad \prod_{i<\alpha} \kappa_i := \Big| \prod_{i<\alpha} \downarrow\kappa_i \Big| .$$

Lemma 4.2. *If $\kappa \geq \aleph_0$ and $\lambda_i \geq 1$, for $i < \kappa$, then*

$$\sum_{i<\kappa} \lambda_i = \kappa \otimes \sup \{ \lambda_i \mid i < \kappa \} .$$

Proof. Let $\mu := \sup \{ \lambda_i \mid i < \kappa \}$. Note that

$$\kappa = \textstyle\sum_{i<\kappa} 1 \leq \sum_{i<\kappa} \lambda_i \quad \text{and} \quad \mu = \sup \{ \lambda_i \mid i < \kappa \} \leq \sum_{i<\kappa} \lambda_i$$

implies $\kappa \otimes \mu = \max \{\mu, \kappa\} \leq \sum_{i<\kappa} \lambda_i \leq \sum_{i<\kappa} \mu = \kappa \otimes \mu$. □

Corollary 4.3. *If $\kappa \geq \aleph_0$ and $\lambda_i \leq \kappa$, for $i < \kappa$, then $\sum_{i<\kappa} \lambda_i \leq \kappa$.*

We have seen in Lemma 3.7 (f) that $\kappa^\lambda > \kappa$, for infinite λ. Ordinal exponentiation, on the other hand, does not increase the cardinality.

Lemma 4.4. *If α and $\beta > 0$ are ordinals and at least one of them is infinite then*

$$\left|{\downarrow}(\alpha^{(\beta)})\right| = |{\downarrow}\alpha| \otimes |{\downarrow}\beta| .$$

Proof. If $\alpha = 0$ then $|{\downarrow}(\alpha^{(\beta)})| = 0 = |{\downarrow}\alpha| \otimes |{\downarrow}\beta|$. Otherwise, we obviously have $|{\downarrow}\alpha| \leq |{\downarrow}(\alpha^{(\beta)})|$ and $|{\downarrow}\beta| \leq |{\downarrow}(\alpha^{(\beta)})|$. Conversely,

$${\downarrow}(\alpha^{(\beta)}) = \bigcup_{n<\omega} \bigcup \{ ({\downarrow}\alpha)^X \mid X \subseteq {\downarrow}\beta , |X| = n \} .$$

Since $|({\downarrow}\alpha)^n| \leq |{\downarrow}\alpha| \oplus \aleph_0$, for $n < \omega$, it follows from Corollary 4.3 that

$$\begin{aligned}
\left|{\downarrow}(\alpha^{(\beta)})\right| &\leq \sum_{n<\omega} \sum_{i<|({\downarrow}\beta)^n|} |({\downarrow}\alpha)^n| \\
&\leq \sum_{n<\omega} |({\downarrow}\beta)^n| \otimes |{\downarrow}\alpha| \otimes \aleph_0 \\
&= \sum_{n<\omega} |{\downarrow}\alpha| \otimes |{\downarrow}\beta| \otimes \aleph_0 \\
&= \aleph_0 \otimes |{\downarrow}\alpha| \otimes |{\downarrow}\beta| \otimes \aleph_0 \\
&= |{\downarrow}\alpha| \otimes |{\downarrow}\beta| .
\end{aligned}$$

□

Corollary 4.5. *Let A and $B \neq \varnothing$ be sets, at least one of them infinite. There are exactly $|A| \oplus |B|$ functions $p : A_0 \to B$ with finite domain $A_0 \subseteq A$.*

Theorem 4.6 (Kőnig). *If $\kappa_i < \lambda_i$, for $i < \alpha$, then*

$$\sum_{i<\alpha} \kappa_i < \prod_{i<\alpha} \lambda_i \,.$$

Proof. We show that there is no surjective function

$$f : \dot{\bigcup}_{i<\alpha} {\downarrow}\kappa_i \to \prod_{i<\alpha} {\downarrow}\lambda_i \,.$$

For a contradiction, suppose such a function exists and define

$$Z_k := \{\, \beta_k < \lambda_k \mid (\beta_i)_i = f\langle k, \gamma\rangle \text{ for some } \gamma < \kappa_k \,\} \,.$$

Then $|Z_k| \leq \kappa_k < \lambda_k$. Hence, ${\downarrow}\lambda_k \setminus Z_k \neq \varnothing$ and there is some sequence $(\beta_i)_i \in \prod_{i<\alpha}({\downarrow}\lambda_i \setminus Z_i)$. As f is surjective there must be some element $\langle k, \gamma\rangle$ with $f\langle k, \gamma\rangle = (\beta_i)_i$. But this implies that $\beta_k \in Z_k$. A contradiction. □

Consider some set A of cardinality $|A| = \kappa$. What is the shortest sequence of sets $(B_\alpha)_{\alpha<\lambda}$ of cardinality $|B_\alpha| < \kappa$ such that $A = \bigcup_{\alpha<\lambda} B_\alpha$? This question leads to the notion of cofinality.

Definition 4.7. (a) Let $\langle A, \leq\rangle$ be a linear order. A subset $X \subseteq A$ is *cofinal* in A if, for every $a \in A$, there is some element $x \in X$ with $a \leq x$.

We call a function $f : B \to A$ cofinal if $\operatorname{rng} f$ is cofinal in A.

(b) The *cofinality* $\operatorname{cf} \alpha$ of an ordinal α is the minimal ordinal β such that there exists a cofinal function $f : {\downarrow}\beta \to {\downarrow}\alpha$.

Exercise 4.1. Prove that every linear order $\langle A, \leq\rangle$ contains a cofinal subset $X \subseteq A$ such that $\langle X, \leq\rangle$ is well-ordered.

Lemma 4.8. *Let $\langle A, \leq\rangle$ be a linear order. If X is cofinal in A and Y is cofinal in X then Y is cofinal in A.*

We can restate the definition of the cofinality in a more useful form as follows.

Lemma 4.9. *If* $(\alpha_i)_{i<\lambda}$ *is a sequence of ordinals* $\alpha_i < \kappa$ *of length* $\lambda < \operatorname{cf} \kappa$ *then*

$$\sup \{ \alpha_i \mid i < \lambda \} < \kappa .$$

Exercise 4.2. Prove that $\operatorname{cf} \omega = \omega$ and $\operatorname{cf} \aleph_\omega = \omega$.

The following lemmas provide tools to compute the cofinality of an ordinal.

Lemma 4.10. *For every ordinal* α*, we have*

$$\operatorname{cf} \alpha \leq \alpha \quad \textit{and} \quad \operatorname{cf}(\alpha + 1) = 1 .$$

Proof. For the first inequality, it is sufficient to note that the identity function $\mathrm{id}_{\downarrow\alpha} : {\downarrow}\alpha \to {\downarrow}\alpha$ is cofinal. The second claim follows from the fact that the function $f : [1] \to {\downarrow}(\alpha + 1)$ with $f(0) := \alpha$ is cofinal. □

Lemma 4.11. *If there exists a cofinal function* $f : {\downarrow}\beta \to {\downarrow}\alpha$*, we can construct such a function that is strictly increasing.*

Proof. The function $g : {\downarrow}\beta \to {\downarrow}\alpha$ with

$$g(\gamma) = \max\big\{ f(\gamma),\ \sup \{ g(\eta) + 1 \mid \eta < \gamma \} \big\}$$

is cofinal and increasing. □

Lemma 4.12. *If* $f : {\downarrow}\alpha \to {\downarrow}\beta$ *is strictly increasing and cofinal then* $\operatorname{cf} \alpha = \operatorname{cf} \beta$.

Proof. Let $g : {\downarrow} \operatorname{cf} \alpha \to {\downarrow}\alpha$ and $h : {\downarrow} \operatorname{cf} \beta \to {\downarrow}\beta$ be strictly increasing cofinal maps. Since the composition $f \circ g : {\downarrow} \operatorname{cf} \alpha \to {\downarrow}\beta$ is cofinal we have $\operatorname{cf} \alpha \leq \operatorname{cf} \beta$.

For the converse, we distinguish two cases. If $\alpha = \alpha_0 + 1$ is a successor, then $\operatorname{cf} \alpha = 1$ and $\{f(0)\}$ is cofinal in ${\downarrow}\beta$. Hence, $\beta = f(0) + 1$ is a successor and $\operatorname{cf} \beta = 1$. If α is a limit ordinal, we define a function $k : {\downarrow} \operatorname{cf} \beta \to {\downarrow}\alpha$ by

$$k(\gamma) := \min \{ \eta \mid f(\eta) > h(\gamma) \} .$$

This function is cofinal since, given $\eta < \alpha$, there is some $\gamma < \operatorname{cf} \beta$ with $h(\gamma) \geq f(\eta)$. It follows that $k(\gamma) \geq \eta$ since $f(k(\gamma)) > h(\gamma) \geq f(\eta)$ and f is strictly increasing. □

Corollary 4.13. $\operatorname{cf}(\operatorname{cf} \alpha) = \operatorname{cf} \alpha$, *for every* $\alpha \in \mathrm{On}$.

We will see many examples showing that cardinals κ with $\operatorname{cf} \kappa = \kappa$ behave in a sane way while, for other cardinals, we might have to deal with pathological cases. Cardinals of the first kind are therefore called *regular,* the other ones are *singular.*

Definition 4.14. An ordinal α is called *regular* if α is a limit ordinal and $\operatorname{cf} \alpha = \alpha$. A cardinal which is not regular is called *singular.*

Remark. In Corollary 4.13 we have shown that every ordinal of the form $\operatorname{cf} \alpha$ is regular. It follows that the class of all regular ordinals is precisely the range $\operatorname{rng}(\operatorname{cf})$ of the function cf.

Example. ω and $\aleph_1$ are regular while $\aleph_\omega$ is singular.

The next lemma indicates that the notion of cofinality is mainly interesting for cardinals.

Lemma 4.15. *Every regular ordinal is a cardinal.*

Proof. Let $\alpha \in \mathrm{On} \setminus \mathrm{Cn}$ be an ordinal that is not a cardinal and set $\kappa := |\alpha| < \alpha$. By definition, there exists a bijection $f : {\downarrow}\kappa \to {\downarrow}\alpha$. This function is surjective and, hence, cofinal. Consequently, we have $\operatorname{cf} \alpha \leq \kappa < \alpha$. □

It turns out that all successor cardinals are regular while most limit cardinals are singular.

Lemma 4.16. *Every successor cardinal is regular.*

Proof. Suppose there exists a cardinal $\kappa \in \mathrm{Cn}$ such that $\alpha := \operatorname{cf} \kappa^+ < \kappa^+$. Let $f : {\downarrow}\alpha \to {\downarrow}\kappa^+$ be a cofinal map. Since κ^+ is a limit ordinal we have

$$ {\downarrow}\kappa^+ = \bigcup \{ {\downarrow} f(\beta) \mid \beta < \alpha \} . $$

By Corollary 4.3, it follows that

$$\kappa^+ = |{\downarrow}\kappa^+| = \Big|\bigcup \{ {\downarrow} f(\beta) \mid \beta < \alpha \}\Big| \leq \sum_{\beta<\alpha} \kappa = \kappa \,.$$

A contradiction. □

Lemma 4.17. *If δ is a limit ordinal then* $\operatorname{cf} \aleph_\delta = \operatorname{cf} \delta$.

Proof. We can define a strictly increasing cofinal function $f : {\downarrow}\delta \to {\downarrow}\aleph_\delta$ by $f(\alpha) := \aleph_\alpha$. Hence, the claim follows from Lemma 4.12. □

It follows that regular limit cardinals are quite rare.

Corollary 4.18. *If δ is a limit ordinal such that $\aleph_\delta$ is regular then $\aleph_\delta = \delta$.*

Cardinal exponentiation is the least understood operation of those introduced so far. There are many open questions that the usual axioms of set theory are not strong enough to answer. For example, we do not know what the value of $2^{\aleph_0}$ is. Given an arbitrary model of set theory we can construct a new model where $2^{\aleph_0} = \aleph_1$, but we can also find models where $2^{\aleph_0}$ equals $\aleph_2$ or $\aleph_3$.

In the remainder of this section we present some elementary results that *can* be proved. The notion of cofinality appears at several places in these proofs. First, let us compute the cardinality of all stages S_α, by a simple induction.

Definition 4.19. We define the cardinal $\beth_\alpha(\kappa)$ (*'beth alpha'*), for $\alpha \in \mathrm{On}$ and $\kappa \in \mathrm{Cn}$, recursively by

$$\beth_0(\kappa) := \kappa \,,$$

$$\beth_{\alpha+1}(\kappa) := 2^{\beth_\alpha(\kappa)} \,,$$

and $$\beth_\delta(\kappa) := \sup \{ \beth_\alpha(\kappa) \mid \alpha < \delta \} \,, \quad \text{for limit ordinals } \delta \,.$$

Further, let $\beth_\alpha := \beth_\alpha(\aleph_0)$.

Lemma 4.20. *For $\alpha \in \mathrm{On}_0$, we have*

$$|S_\alpha| = \beth_\alpha(0) \quad \textit{and} \quad |S_{\omega+\alpha}| = \beth_\alpha .$$

The next lemma shows that most questions about cardinal exponentiation can be reduced to the computation of the cardinality of power sets.

Lemma 4.21. *If $2 \leq \kappa \leq 2^\lambda$ and $\lambda \geq \aleph_0$ then $\kappa^\lambda = 2^\lambda$.*

Proof. $2^\lambda \leq \kappa^\lambda \leq (2^\lambda)^\lambda = 2^{\lambda \otimes \lambda} = 2^\lambda$. □

What is the value of κ^λ, for $\lambda < \kappa$? We can give only partial answers.

Lemma 4.22. *If $\kappa \geq \aleph_0$ and $\lambda \geq \operatorname{cf} \kappa$ then $\kappa^\lambda > \kappa$. In particular, $\kappa^{\operatorname{cf} \kappa} > \kappa$.*

Proof. Fix a cofinal function $f : {\downarrow}\lambda \to {\downarrow}\kappa$. By Theorem 4.6, we have

$$\kappa^\lambda = \left|({\downarrow}\kappa)^{{\downarrow}\lambda}\right| = \left|\prod_{\alpha<\lambda} {\downarrow}\kappa\right| > \left|\dot{\bigcup_{\alpha<\lambda}} {\downarrow} f(\alpha)\right| \geq |{\downarrow}\kappa| = \kappa .$$ □

Corollary 4.23. $\operatorname{cf} 2^\kappa > \kappa$.

Proof. $\operatorname{cf} 2^\kappa \leq \kappa$ would imply $(2^\kappa)^{\operatorname{cf} 2^\kappa} \leq (2^\kappa)^\kappa = 2^{\kappa \otimes \kappa} = 2^\kappa < (2^\kappa)^{\operatorname{cf} 2^\kappa}$. Contradiction. □

The next theorem summarises the extend of our knowledge about cardinal exponentiation. First, we introduce some abbreviations.

Definition 4.24. For cardinals κ and λ we set

$$({<}\kappa)^\lambda := \sup \{ \mu^\lambda \mid \mu < \kappa \} \quad \text{and} \quad \kappa^{<\lambda} := \sup \{ \kappa^\mu \mid \mu < \lambda \} .$$

Lemma 4.25. $\operatorname{cf} ({<}\kappa)^\lambda \leq \operatorname{cf} \kappa$ *and* $\operatorname{cf} \kappa^{<\lambda} \leq \operatorname{cf} \lambda$.

Theorem 4.26. *Let $\kappa \geq 2$ and $\lambda \geq \aleph_0$.*

(a) *If $2 < \kappa \leq \lambda$ then $\kappa^\lambda = 2^\lambda = ({<}\kappa)^\lambda$.*

(b) *If $\operatorname{cf} \kappa \leq \lambda < \kappa$ then $\kappa < \kappa^\lambda = (({<}\kappa)^\lambda)^{\operatorname{cf} \kappa} \leq 2^\kappa$.*

(c) *If* $\lambda < \operatorname{cf}\kappa$ *then* $\kappa^\lambda = \kappa \oplus (<\kappa)^\lambda$.

Proof. (a) The first equality was proved in Lemma 4.21. For the second one, note that $\kappa > 2$ implies $2^\lambda \leq (<\kappa)^\lambda \leq \kappa^\lambda$.

(b) By (a) and Corollary 4.22, it follows that $\kappa < \kappa^\lambda \leq 2^\kappa$. Further, $(<\kappa)^\lambda \leq \kappa^\lambda$ implies that

$$\big((<\kappa)^\lambda\big)^{\operatorname{cf}\kappa} \leq \big(\kappa^\lambda\big)^{\operatorname{cf}\kappa} = \kappa^{\lambda \otimes \operatorname{cf}\kappa} = \kappa^\lambda .$$

For the converse, fix a cofinal function $f : {\downarrow}\operatorname{cf}\kappa \to {\downarrow}\kappa$. We have

$$\begin{aligned} \kappa^\lambda \leq \Big|\dot{\bigcup_{\alpha<\operatorname{cf}\kappa}} {\downarrow} f(\alpha)\Big|^\lambda &\leq \Big|\prod_{\alpha<\operatorname{cf}\kappa} {\downarrow} f(\alpha)\Big|^\lambda \\ &= \Big|\prod_{\alpha<\operatorname{cf}\kappa} {\downarrow} f(\alpha)^{\downarrow\lambda}\Big| \\ &\leq \Big|\prod_{\alpha<\operatorname{cf}\kappa} {\downarrow}(<\kappa)^\lambda\Big| \leq \big((<\kappa)^\lambda\big)^{\operatorname{cf}\kappa} . \end{aligned}$$

(c) If $\lambda < \operatorname{cf}\kappa$ then

$$({\downarrow}\kappa)^{\downarrow\lambda} = \bigcup \{\, ({\downarrow}\mu)^{\downarrow\lambda} \mid \mu < \kappa \,\} ,$$

since the range of every function ${\downarrow}\lambda \to {\downarrow}\kappa$ is bounded by some $\mu < \kappa$. Hence,

$$\kappa^\lambda \leq \sum_{\mu<\kappa} \mu^\lambda \leq \sum_{\mu<\kappa} (<\kappa)^\lambda = \kappa \otimes (<\kappa)^\lambda .$$

If $\kappa = \mu^+$ then $(<\kappa)^\lambda = \mu^\lambda$ and

$$\kappa^\lambda \leq \kappa \otimes (<\kappa)^\lambda = \kappa \otimes \mu^\lambda \leq \kappa^\lambda .$$

Otherwise, κ is a limit and $(<\kappa)^\lambda \geq \sup\{\, \mu \mid \mu < \kappa \,\} = \kappa$, which implies that

$$\kappa^\lambda \leq \kappa \otimes (<\kappa)^\lambda = (<\kappa)^\lambda \leq \kappa^\lambda . \qquad \square$$

Corollary 4.27. *If κ and λ are cardinals such that $2^\mu = \mu^+$, for all $\mu \leq \kappa$, then*

$$\kappa^\lambda = \begin{cases} 2^\lambda & \text{if } \kappa \leq \lambda\,, \\ \kappa^+ & \text{if } \operatorname{cf}\kappa \leq \lambda < \kappa\,, \\ \kappa & \text{if } \lambda < \operatorname{cf}\kappa\,. \end{cases}$$

Lemma 4.28. *Let κ be a cardinal. We have $\kappa = \beth_\delta$, for some limit ordinal δ, if and only if $\kappa > \aleph_0$ and $2^\lambda < \kappa$, for all $\lambda < \kappa$.*

Proof. ($\Rightarrow$) We have $\beth_\delta > \beth_0 = \aleph_0$. If $\lambda < \beth_\delta$ then $\lambda \leq \beth_\alpha$, for some $\alpha < \delta$. Hence, $2^\lambda \leq 2^{\beth_\alpha} = \beth_{\alpha+1} < \beth_\delta$.

($\Leftarrow$) Let $A := \{\, \alpha + 1 \mid \beth_\alpha < \kappa \,\}$ and $\delta := \sup A$. By definition of A, it follows that $\beth_\delta \geq \kappa$. On the other hand,

$$\begin{aligned} \kappa &= \sup\{\, 2^\lambda \mid \lambda < \kappa \,\} \\ &\geq \sup\{\, 2^{\beth_\alpha} \mid \beth_\alpha < \kappa \,\} = \sup\{\, \beth_\alpha \mid \alpha \in A \,\} = \beth_\delta\,. \end{aligned}$$

Hence, $\kappa = \beth_\delta$. Since $\beth_\delta = \kappa > \aleph_0$ we have $\delta > 0$. To show that δ is a limit suppose that $\delta = \alpha + 1$. Then $\beth_\alpha < \kappa$ implies $\beth_\delta = 2^{\beth_\alpha} < \kappa$. Contradiction. □

We conclude this section with some results about sets of sequences indexed by ordinals. As we will see in Section B2.1, such a set forms the domain of a *tree*. Recall that a sequence indexed by an ordinal α is just a function $\downarrow\alpha \to A$.

Definition 4.29. If A is a set and $\alpha \in \mathrm{On}$, we define

$$A^\alpha := A^{\downarrow\alpha} \quad \text{and} \quad A^{<\alpha} := \bigcup_{\beta<\alpha} A^\beta\,.$$

Let us compute the cardinality of $A^{<\alpha}$. We are especially interested in the case where $\alpha = \omega$, i.e., in the set of all finite sequences.

Lemma 4.30. *If $|A| > 1$ then $|A^{<\alpha}| = |A|^{<|\alpha|}$.*

Lemma 4.31. *If $\kappa > 0$ then $\kappa^{<\aleph_0} = \kappa \oplus \aleph_0$.*

Proof. If $\kappa \geq \aleph_0$ then

$$\kappa^{<\aleph_0} = \sup\{\, \kappa^n \mid n < \aleph_0 \,\} = \sup\{\kappa\} = \kappa = \kappa \oplus \aleph_0\,.$$

For $\kappa = 1$, we can define a bijection $[1]^{<\omega} \to \downarrow\omega$ by

$$\underbrace{\langle 0, \dots, 0 \rangle}_{n \text{ times}} \mapsto n\,.$$

Hence, $1^{<\aleph_0} = \aleph_0$. If $1 < \kappa < \aleph_0$, it follows that

$$\aleph_0 = 1^{<\aleph_0} \leq \kappa^{<\aleph_0} \leq \aleph_0^{<\aleph_0} = \aleph_0\,. \qquad \square$$

Corollary 4.32. *$\kappa^{<\kappa} \geq \kappa$, for all $\kappa > 0$. If $\kappa \geq \aleph_0$ then $\kappa \leq 2^{<\kappa} \leq \kappa^{<\kappa}$.*

Proof. If $\kappa \geq \aleph_0$ then $2^{<\kappa} = \sup\{\, 2^\lambda \mid \lambda < \kappa \,\} \geq \sup\{\, \lambda^+ \mid \lambda < \kappa \,\} \geq \kappa\,.$ $\square$

Lemma 4.33. *If κ is an infinite regular cardinal then $\kappa^{<\kappa} = 2^{<\kappa}$.*

Proof. For $\aleph_0 \leq \lambda, \mu < \kappa$ we have

$$\lambda^\mu \leq (\lambda \oplus \mu)^{\lambda \oplus \mu} = 2^{\lambda \oplus \mu} \leq 2^{<\kappa}.$$

If $\operatorname{cf}\kappa = \kappa$, it follows by Theorem 4.26 and Corollary 4.32 that

$$\kappa^\mu = \kappa \oplus (<\kappa)^\mu = \kappa \oplus \sup\{\, \lambda^\mu \mid \lambda < \kappa \,\} \leq 2^{<\kappa}, \quad \text{for all } \mu < \kappa\,.$$

Consequently, $\kappa^{<\kappa} \leq 2^{<\kappa}$. $\square$

Corollary 4.34. *Let κ be an infinite cardinal. We have $\kappa^{<\kappa} = \kappa$ if, and only if, κ is regular and $2^{<\kappa} = \kappa$.*

Proof. One direction follows from the preceding lemma. For the other one, note that $\operatorname{cf}\kappa < \kappa$ implies $\kappa^{<\kappa} \geq \kappa^{\operatorname{cf}\kappa} > \kappa$, and $2^{<\kappa} > \kappa$ implies $\kappa^{<\kappa} \geq 2^{<\kappa} > \kappa$. $\square$

5. The Axiom of Replacement

At several times when mappings between classes were concerned we remarked that we need an additional axiom to prove the desired statement. This axiom is the generalisation of the following lemma to functions that are proper classes.

Lemma 5.1. *Let f be a function. If f is a set then so is $f[A]$, for all $A \subseteq \operatorname{dom} f$.*

Proof. Since f is a set so is $\operatorname{rng} f$. Therefore,

$$f[A] = \{\, y \in \operatorname{rng} f \mid y = f(x) \text{ for some } x \in A \,\}$$

is a set. □

Before stating the axiom let us collect several equivalent formulations of it.

Theorem 5.2. *The following statements are equivalent:*

(1) *If F is a function and $A \subseteq \operatorname{dom} F$ is a set then $F[A]$ is also a set.*

(2) *If F is a function and $\operatorname{dom} F$ is a set then so is $\operatorname{rng} F$.*

(3) *A function F is a set if, and only if, $\operatorname{dom} F$ is a set.*

(4) *There exists no bijection $F : a \to B$ between a set a and a proper class B.*

(5) *A class A is a set if, and only if, $|A| < \infty$.*

(6) *If $\alpha \in \mathrm{On}$ is an ordinal and $(A_i)_{i<\alpha}$ a sequence of sets then the class $\bigcup_{i<\alpha} A_i$ is also a set.*

Proof. (3) ⇒ (2) Let F be a function and suppose that $\operatorname{dom} F$ is a set. Then F is a set and so is $\operatorname{rng} F$.

(2) ⇒ (3) Clearly, if F is a set then so is $\operatorname{dom} F$. For the converse, let F be a function such that $\operatorname{dom} F$ is a set. By assumption, then $\operatorname{rng} F$ is also a set. Since $F \subseteq \operatorname{dom} F \times \operatorname{rng} F$ it follows that F is a set.

(2) ⇒ (1) Let F be a function and $A \subseteq \operatorname{dom} F$ a set. Let $G := F \upharpoonright A$ be the restriction of F to A. We apply the assumption to G. Since $\operatorname{dom} G = A$ is a set so is $\operatorname{rng} G = F[A]$.

(1) ⇒ (6) Let $F : \downarrow\alpha \to \mathbb{S}$ be the function with $F(i) = A_i$, for $i < \alpha$. By assumption, $B := F[\downarrow\alpha]$ is a set. Hence, so is

$$\bigcup B = \bigcup_{i<\alpha} A_i .$$

(6) ⇒ (2) Let $F : A \to B$ be a function and $A = \operatorname{dom} F$ a set. Let $\kappa := |A|$ and fix a bijection $g : \downarrow\kappa \to A$. We define a sequence $(B_i)_{i<\kappa}$ of sets by $B_i := S(F(g(i)))$. By assumption, $C := \bigcup_{i<\kappa} B_i$ is a set. For every $a \in A$, we have $S(F(a)) \subseteq C$ or, equivalently, $S(F(a)) \in \wp(C)$. It follows that $S(\operatorname{rng} F) = S(F[A]) \subseteq \wp(C)$. In particular, $\operatorname{rng} F$ is a set.

(2) ⇒ (5) If A is a set then $|A| < \infty$, by Lemma 2.2. For the converse, suppose that $\kappa := |A| < \infty$ and let $F : \downarrow\kappa \to A$ be a bijection. Since κ is a set it follows by assumption that $A = \operatorname{rng} F$ is also a set.

(5) ⇒ (4) Let $F : a \to B$ be a bijection where a is a set. Then $|B| = |a| < \infty$. Hence, B is also a set.

(4) ⇒ (2) Let $F : A \to B$ be a function where $A = \operatorname{dom} F$ is a set. Let $B_0 := \operatorname{rng} F$. Since the function $F : a \to B_0$ is surjective there exists a function $G : B_0 \to a$ such that $F \circ G = \mathrm{id}_{B_0}$. Let $A_0 := \operatorname{rng} G$. The restriction $F : A_0 \to B_0$ is a bijection. Since $A_0 \subseteq A$ is a set so is $B_0 = \operatorname{rng} F$. □

Axiom of Replacement. *If F is a function and* $\operatorname{dom} F$ *is a set then so is* $\operatorname{rng} F$.

Let us finally prove the results we promised in the preceding sections. First, up to isomorphism, $\mathfrak{On}$ is the only well-order that is a proper class.

Lemma 5.3. *Let $\mathfrak{A} = \langle A, \leq_A \rangle$ and $\mathfrak{B} = \langle B, \leq_B \rangle$ be well-orders. If A and B are proper classes then $\mathfrak{A} \cong \mathfrak{B}$.*

Proof. Suppose that $\mathfrak{A} \ncong \mathfrak{B}$. By Theorem A3.1.11, there either exists an isomorphism $f : A \to \downarrow b$, for some $b \in B$, or some isomorphism $g :$

$\downarrow a \to B$, for some $a \in A$. By symmetry, we may assume w.l.o.g. the latter. $\downarrow a$ is a set since $\leq_A$ is left-narrow. Hence, by the Axiom of Replacement, $B = g[\downarrow a]$ is also a set. Contradiction. □

It follows that it does not matter which of the two definitions of an ordinal we adopt.

Corollary 5.4. $\mathfrak{On}_0 \cong \mathfrak{Cn} \cong \mathfrak{On}$.

Finally, we state the general form of the Principle of Transfinite Recursion.

Theorem 5.5 (Principle of Transfinite Recursion). *If $H : A^{<\infty} \to A$ is a total function that defines the function F by transfinite recursion then* $\operatorname{dom} F = \mathrm{On}$.

Proof. For a contradiction, suppose that $\operatorname{dom} F = \downarrow\alpha \subset \mathrm{On}$. In particular, $\operatorname{dom} F$ is a set. By the Axiom of Replacement, it follows that $\operatorname{rng} F$ is also a set. Since $\operatorname{rng} F \subseteq A$ we therefore have $F \in A^{<\infty} = \operatorname{dom} H$ in contradiction to Theorem A3.3.4. □

Lemma 5.6. *Every strictly continuous function $f : \mathrm{On} \to \mathrm{On}$ has arbitrarily large fixed points.*

Proof. For every $\alpha \in \mathrm{On}$ we have to find a fixed point $\gamma \geq \alpha$. If F is the fixed-point induction of f over α then $F[\downarrow\omega]$ exists. By Lemma A3.3.13 it follows that $\gamma := F(\infty) = F(\omega) \geq \alpha$ is a fixed point of f. □

Corollary 5.7. *There are arbitrarily large cardinals κ such that* $\operatorname{cf} \kappa = \aleph_0$ *and either $\aleph_\kappa = \kappa$ or $\beth_\kappa = \kappa$.*

Proof. The functions $f : \alpha \mapsto \aleph_\alpha$ and $g : \alpha \mapsto \beth_\alpha$ are strictly continuous. Furthermore, they are defined by transfinite recursion. Therefore, Theorem 5.5 implies that their domain is all of On. By Lemma A3.3.13 and Lemma 5.6, it follows that f and g have arbitrarily large inductive fixed points κ, and these fixed points are of the form

$$\kappa = \sup \{ f^n(\alpha) \mid n < \omega \}, \qquad \text{for some } \alpha .$$

In particular, $\operatorname{cf} \kappa = \aleph_0$. □

Exercise 5.1. Prove that $S_{\omega 2}$ satisfies all axioms of set theory except for the Axiom of Replacement.

6. *Stationary sets*

There are many places in mathematics where one wants to argue that there are 'many' objects with a certain property. This has lead to several notions of 'large' and 'small' sets, for instance, being dense, being cofinite, having measure 1, or belonging to a given ultrafilter.

Example. Let κ be a regular cardinal and A a set of size $|A| = \kappa$. We call a subset $X \subseteq A$ *large* if it has size κ. A subset $X \subseteq A$ is *very large* if its complement $A \setminus X$ is not large. It is straightforward to check that the classes of large and very large sets have the following properties:

(a) Every very large set is large.

(b) A set X is large if, and only if, it has a non-empty intersection with every very large set.

(c) The intersection of less than κ very large sets is very large.

(d) The intersection of a very large set and a large one is large.

(e) Every large set can be partitioned into κ disjoint large subsets.

(f) If $f : X \to Y$ is a function from a large set X into a set Y that is not large, there is some element $y \in Y$ such that the fibre $f^{-1}(y)$ is large.

In this section we introduce two notions of 'largeness' for sets of ordinals which exhibit the same properties as the large and very large sets of the above example: *closed unbounded sets* correspond to the very large sets and *stationary sets* correspond to the large one. We will prove analogues to all of the above properties. We start with closed unbounded sets.

Definition 6.1. Let κ be a cardinal. A subset $C \subseteq \kappa$ is *closed unbounded* if it is cofinal in κ and, for every non-empty subset $X \subseteq C$ with $\sup X < \kappa$, we have $\sup X \in C$.

Example. For every ordinal $\alpha < \kappa$, the set $\Uparrow\alpha$ is obviously closed unbounded. Another example of a closed unbounded set is the set of all limit ordinals $\alpha < \kappa$.

Before verifying the above properties let us present two ways to construct closed unbounded subsets of a given closed unbounded set.

Lemma 6.2. *Let κ be an uncountable regular cardinal and $C \subseteq \kappa$ closed unbounded.*

(a) *The set $C' := \{\, \alpha \in C \mid C \cap \alpha$ is cofinal in $\alpha \,\}$ is closed unbounded.*

(b) *For every cardinal λ such that $C \cap \lambda$ is cofinal in λ, the set $C \cap \lambda$ is closed unbounded in λ.*

Proof. (a) To show that C' is cofinal, let $\alpha < \kappa$. Since C is cofinal, we can construct an increasing sequence $\alpha < \beta_0 < \beta_1 < \dots$ of elements $\beta_n \in C$, for $n < \omega$. Since C is closed and κ is regular, it follows that $\delta := \sup_{n<\omega} \beta_n \in C$. Furthermore, the fact that all β_n belong to $C \cap \delta$ implies that $C \cap \delta$ is cofinal in δ. Hence, $\delta \in C'$.

It remains to show that C' is closed. Consider a set $X \subseteq C'$ such that $\delta := \sup X < \kappa$. If $\delta \in X \subseteq C'$, we are done. Hence, we may assume that $\delta \notin X$. Note that $X \subseteq C$ implies that $\delta \in C$. Furthermore, $X \subseteq C \cap \delta$ implies that $C \cap \delta$ is cofinal in δ. Consequently, $\delta \in C'$.

(b) By assumption, $C \cap \lambda$ is cofinal in λ. To show that it is also closed, let $X \subseteq C \cap \lambda$ be a set with $\sup X < \lambda$. Then $X \subseteq C$ implies that $\sup X \in C$. Hence, $\sup X \in C \cap \lambda$. □

The first property we check is that closed unbounded sets are closed under intersections. We consider two variants: ordinary intersections and so-called diagonal intersections.

Lemma 6.3. *Let κ be an uncountable regular cardinal. If $C, D \subseteq \kappa$ are closed unbounded then so is $C \cap D$.*

Proof. If $X \subseteq C \cap D$ and $\sup X < \kappa$ then $X \subseteq C$ implies $\sup X \in C$ and $X \subseteq D$ implies $\sup X \in D$. Consequently, we have $\sup X \in C \cap D$.

To show that $C \cap D$ is cofinal let $\alpha < \kappa$. Then there is some element $\beta_0 \in C$ with $\alpha \leq \beta_0$. Similarly, there is some element $\gamma_0 \in D$ with $\beta_0 \leq \gamma_0$. Continuing in this way we obtain an increasing sequence

$$\alpha \leq \beta_0 \leq \gamma_0 \leq \beta_1 \leq \gamma_1 \leq \dots$$

where $\beta_i \in C$ and $\gamma_i \in D$. Since $\operatorname{cf} \kappa > \omega$ it follows that

$$\delta := \sup_i \beta_i = \sup_i \gamma_i < \kappa .$$

As C and D are closed unbounded we have $\delta \in C$ and $\delta \in D$. Thus, we have found an element $\delta \in C \cap D$ with $\alpha \leq \delta$. ☐

Exercise 6.1. Show that this lemma fails for closed unbounded subsets of $\aleph_0$.

Proposition 6.4. *Let κ be an uncountable regular cardinal. If $\mathcal{C} \subseteq \wp(\kappa)$ is a family of closed unbounded sets with $|\mathcal{C}| < \kappa$ then $\bigcap \mathcal{C}$ is closed unbounded.*

Proof. Let $(C_i)_{i<\alpha}$ be a sequence of closed unbounded subsets of κ with $\alpha < \kappa$. By induction on α, we prove that $\bigcap_{i<\alpha} C_i$ is closed unbounded.

For $\alpha = 1$ there is nothing to do and the successor step follows immediately from the preceding lemma. Hence, we may assume that α is a limit ordinal. Furthermore, we know by inductive hypothesis that the sets $\bigcap_{i<\beta} C_i$, for $\beta < \alpha$ are closed unbounded. Therefore, replacing C_β by $\bigcap_{i\leq\beta} C_i$ we may assume that $C_0 \supseteq C_1 \supseteq \dots$.

Let $C := \bigcap_{i<\alpha} C_i$. If $X \subseteq C$ is a set with $\sup X < \kappa$, then $X \subseteq C_i$ implies that $\sup X \in C_i$, for all i. Consequently, we have $\sup X \in C$.

To show that C is cofinal let $\beta < \kappa$. We construct an increasing sequence $(\gamma_i)_{i<\alpha}$ as follows. Choose some $\gamma_0 \in C_0$ with $\beta \leq \gamma_0$. For $0 < i < \alpha$, let $\gamma_i \in C_i$ be some element with $\gamma_i \geq \sup \{ \gamma_k \mid k < i \}$. Since κ is regular it follows that $\delta := \sup_i \gamma_i < \kappa$. For $i < \alpha$, let

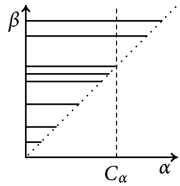
$$X_i := \{ \gamma_k \mid i \leq k < \alpha \} .$$

Then $X_i \subseteq C_i$. Since C_i is closed unbounded it follows that $\delta = \sup X_i \in C_i$. Consequently, we have found an element $\delta \in C$ with $\beta \leq \delta$. □

The second variant of intersection we consider has no correspondence in the above example since it relies on the presence of a linear order.

Definition 6.5. The *diagonal intersection* of a sequence $(C_\alpha)_{\alpha<\kappa}$ of subsets $C_\alpha \subseteq \kappa$ is the set

$$D := \{\, \beta < \kappa \mid \beta \in C_\alpha \text{ for all } \alpha < \beta \,\}.$$



Remark. Note that, if D is the diagonal intersection of $(C_\alpha)_{\alpha<\kappa}$, then $D \setminus (\alpha + 1) \subseteq C_\alpha$, for all α.

Proposition 6.6. *Let κ be an uncountable regular cardinal. The diagonal intersection of a sequence $(C_\alpha)_{\alpha<\kappa}$ of closed unbounded sets is closed unbounded.*

Proof. Let $(C_\alpha)_{\alpha<\kappa}$ be a sequence of closed unbounded sets and let D be their diagonal intersection. By Proposition 6.4, the intersections $C'_\alpha := \bigcap_{\beta<\alpha} C_\beta$ are closed unbounded. Furthermore, the diagonal intersection of $(C'_\alpha)_{\alpha<\kappa}$ is also equal to D. Replacing C_α by C'_α, we may therefore assume that the sequence $(C_\alpha)_{\alpha<\kappa}$ is decreasing.

To show that D is closed, let $X \subseteq D$ be a set with $\delta := \sup X < \kappa$. For $\alpha < \delta$, consider the set $Y_\alpha := \{\, \beta \in D \mid \alpha < \beta < \delta \,\}$. By the definition of the diagonal intersection, we have $Y_\alpha \subseteq D \setminus (\alpha+1) \subseteq C_\alpha$. As C_α is closed, it follows that $\delta = \sup Y_\alpha \in C_\alpha$, for all $\alpha < \delta$. Consequently, $\delta \in D$.

To show that D is unbounded, let $\alpha < \kappa$. To find a bound $\delta \in D$ with $\alpha < \delta$, we construct an increasing sequence $(\beta_n)_{n<\omega}$ of ordinals as follows. Choose some element $\beta_0 \in C_0$ with $\beta_0 > \alpha$. If β_n is already defined, we choose an element $\beta_{n+1} \in C_{\beta_n}$ with $\beta_{n+1} > \beta_n$. We claim that

$\delta := \sup_{n<\omega} \beta_n \in D$. Hence, let $\gamma < \delta$. Then there is some $n < \omega$ with $\gamma < \beta_n$. Since $\beta_k \in C_{\beta_{k-1}} \subseteq C_{\beta_n}$, for $k > n$, it follows that $\delta = \sup_{k>n} \beta_k \in C_{\beta_n} \subseteq C_\gamma$. Hence, $\delta \in C_\gamma$, for all $\gamma < \delta$. This implies that $\delta \in D$. □

Our second notion of a large set is that of a stationary one. As definition we use the analogue of Property (b) from the above example.

Definition 6.7. Let κ be a cardinal. A set $S \subseteq \kappa$ is *stationary* if $S \cap C \neq \emptyset$, for every closed unbounded set $C \subseteq \kappa$.

We start by constructing several kinds of stationary sets.

Lemma 6.8. *Let κ be an uncountable regular cardinal.*

(a) *The set $\{ \alpha < \kappa \mid \operatorname{cf} \alpha = \lambda \}$ is stationary, for every regular $\lambda < \kappa$.*

(b) *Every closed unbounded set is stationary.*

(c) *If S is stationary and C closed unbounded, then $S \cap C$ is stationary.*

Proof. (a) Let $C \subseteq \kappa$ be closed unbounded. We have to find some element $\gamma \in C$ with cofinality λ. Let $f : \langle \kappa, \leq \rangle \to \langle C, \leq \rangle$ be an order isomorphism and set $\gamma := \sup f[\lambda]$. Since C is closed unbounded, we have $\gamma \in C$. As the function $f \restriction \lambda : \lambda \to \gamma$ is a strictly increasing and cofinal, it follows by Lemma 4.12 that $\operatorname{cf} \gamma = \operatorname{cf} \lambda = \lambda$.

(b) Let C be closed unbounded. For every closed unbounded set D, it follows by Lemma 6.3 that the intersection $C \cap D$ is also closed unbounded. In particular, $C \cap D \neq \emptyset$.

(c) If there were a closed unbounded set D with $(S \cap C) \cap D = \emptyset$, then S would not be stationary since $C \cap D$ is closed unbounded, by Lemma 6.3. □

Note that it follows from Lemma 6.8 (a) that there are disjoint stationary sets. Hence, the intersection of two stationary sets is not necessarily stationary.

The next theorem is a very strong version of Property (f) from the example.

Theorem 6.9 (Fodor). *Let κ be an uncountable regular cardinal, $S \subseteq \kappa$ stationary, and $f : S \to \kappa$ a function with $f(\alpha) < \alpha$, for all $\alpha \in S$. Then there exists an ordinal $\gamma < \kappa$ such that $f^{-1}(\gamma)$ is stationary.*

Proof. For a contradiction, suppose that $f^{-1}(\gamma)$ is non-stationary, for every $\gamma < \kappa$. For each $\gamma < \kappa$, choose a closed unbounded set $C_\gamma \subseteq \kappa$ such that $C_\gamma \cap f^{-1}(\gamma) = \varnothing$. By Proposition 6.6, the diagonal intersection D of $(C_\gamma)_{\gamma<\kappa}$ is closed unbounded. Consequently, Lemma 6.8 (c) implies that $S \cap D$ is stationary. Fix an element $\alpha \in S \cap D$. Then $\alpha \in C_\gamma$, for all $\gamma < \alpha$. Since $C_\gamma \cap f^{-1}(\gamma) = \varnothing$, it follows that $\alpha \notin f^{-1}(\gamma)$. Thus, $f(\alpha) \neq \gamma$, for all $\gamma < \alpha$, which implies that $f(\alpha) \geq \alpha$. A contradiction. □

Corollary 6.10. *Let κ be an uncountable regular cardinal, $S \subseteq \kappa$ stationary, and $f : S \to \lambda$ a function with $\lambda < \kappa$. Then there exists an ordinal $\gamma < \lambda$ such that $f^{-1}(\gamma)$ is stationary.*

Proof. By Lemma 6.8 (c), the set $S' := S \setminus \lambda$ is stationary. Since $f(\alpha) < \alpha$, for $\alpha \in S'$, we can apply the Theorem of Fodor to $f \restriction S'$ to find the desired ordinal γ. □

As an application, we prove the existence of so-called *sunflowers*.

Lemma 6.11 (Sunflower lemma). *Let κ be a regular cardinal and λ a cardinal such that $\mu^{<\lambda} < \kappa$, for all $\mu < \kappa$.*

For every family $(S_\alpha)_{\alpha<\kappa}$ of sets of size $|S_\alpha| < \lambda$, there exists a set U and a subset $I \subseteq \kappa$ of size $|I| = \kappa$ such that

$$S_\alpha \cap S_\beta = U\,, \quad \textit{for all distinct } \alpha, \beta \in I\,.$$

Proof. First, we consider the case where $\kappa = \aleph_0$. Then λ is finite and we can prove the claim by induction on λ. We distinguish two cases. If there is no element a that belongs to infinitely many sets S_α, we can choose a set $I \subseteq \kappa$ such that

$$S_\alpha \cap S_\beta = \varnothing\,, \quad \text{for all distinct } \alpha, \beta \in I\,.$$

Otherwise, choose such an element a and set $K := \{\, \alpha < \kappa \mid a \in S_\alpha \,\}$. Applying the inductive hypothesis to the family $(S_\alpha \setminus \{a\})_{\alpha \in K}$, we obtain an infinite set $I \subseteq K$ and some set U' such that

$$(S_\alpha \setminus \{a\}) \cap (S_\beta \setminus \{a\}) = U', \quad \text{for all distinct } \alpha, \beta \in I .$$

Consequently, the sets I and $U := U' \cup \{a\}$ have the desired properties.

It remains to consider the case where κ is uncountable. Note that $\lambda \leq \kappa$. Hence, by choosing some injective function $\bigcup_{\alpha<\kappa} S_\alpha \to \kappa$ we may assume that $S_\alpha \subseteq \kappa$, for every α. According to Lemma 6.8 (a), the set

$$E := \{\, \alpha < \kappa \mid \operatorname{cf} \alpha \geq \lambda \,\}$$

is stationary. We define a function $f : E \to \kappa$ by

$$f(\alpha) = \sup\,(S_\alpha \cap \alpha) .$$

Note that $\operatorname{cf} \alpha \geq \lambda \geq |S_\alpha|$ implies that

$$f(\alpha) = \sup\,(S_\alpha \cap \alpha) < \alpha , \quad \text{for all } \alpha \in E .$$

Consequently, we can use the Theorem of Fodor to find a stationary subset $W \subseteq E$ and an ordinal γ such that

$$f(\alpha) = \gamma , \quad \text{for all } \alpha \in W .$$

Since there are at most $|\gamma|^{<\lambda} < \kappa$ sets of the form $S_\alpha \cap \gamma$, we can use Corollary 6.10 to find a stationary subset $W' \subseteq W$ and some set $U \subseteq \gamma$ such that

$$S_\alpha \cap \gamma = U , \quad \text{for all } \alpha \in W' .$$

We construct a strictly increasing sequence $(\xi_\alpha)_{\alpha<\kappa}$ of ordinals $\xi_\alpha \in W'$ as follows. Let ξ_0 be the minimal element of W'. For the inductive step, suppose that we have already defined ξ_α for all $\alpha < \beta$. Then we chose some element $\xi_\beta \in W'$ such that

$$\xi_\beta > \xi_\alpha \quad \text{and} \quad \xi_\beta > \sup S_{\xi_\alpha} , \quad \text{for all } \alpha < \beta .$$

Note that such an element exists since κ is regular.

Having constructed $(\xi_\alpha)_{\alpha<\kappa}$, it follows for $\alpha < \beta < \kappa$ that

$$S_{\xi_\alpha} \cap S_{\xi_\beta} = (S_{\xi_\alpha} \cap \xi_\beta) \cap S_{\xi_\beta} = S_{\xi_\alpha} \cap (S_{\xi_\beta} \cap \gamma) = U\,.$$

Consequently, the set $I := \{\, \xi_\alpha \mid \alpha < \kappa \,\}$ has the desired properties. □

Exercise 6.2. Let $k, m, n < \omega$ be finite numbers with $n > k!(m-1)^{k+1}$. Prove that, for every family $(S_i)_{i<n}$ of sets of size $|S_i| = k$, there exists a subset $I \subseteq [n]$ of size $|I| = m$ and some set U such that

$$S_i \cap S_j = U\,, \quad \text{for all distinct } i, j \in I\,.$$

We conclude this section by proving that every stationary set can be partitioned into κ disjoint stationary subsets. We start with two technical lemmas.

Lemma 6.12. *Let κ be an uncountable regular cardinal and $S \subseteq \kappa$ a stationary set every element of which is an uncountable regular cardinal. Then the set*

$$W := \{\, \lambda \in S \mid S \cap \lambda \text{ is not stationary in } \lambda \,\}$$

is stationary.

Proof. To show that W is stationary, let $C \subseteq \kappa$ be closed unbounded. By Lemma 6.2 (a), the set

$$C' := \{\, \alpha \in C \mid C \cap \alpha \text{ is cofinal in } \alpha \,\}\,.$$

is closed unbounded. Hence, $S \cap C' \neq \emptyset$. Let λ be the minimal element of $S \cap C'$. Then λ is a regular cardinal and $C \cap \lambda$ is cofinal in λ. Consequently, it follows by Lemma 6.2 (b) that $C \cap \lambda$ is a closed unbounded subset of λ. Hence, Lemma 6.2 (a) implies that $C' \cap \lambda$ is also closed unbounded. Since, by choice of λ, the sets $C' \cap \lambda$ and $S \cap \lambda$ are disjoint, it follows that $S \cap \lambda$ is not stationary. Consequently, $\lambda \in W \cap C$, as desired. □

Lemma 6.13. *Let κ be an uncountable regular cardinal, $S \subseteq \kappa$ stationary, and, for every $\alpha \in S$, let $\gamma_\alpha : \operatorname{cf} \alpha \to \alpha$ be a cofinal and strictly increasing function. If either*

(i) *there is an infinite cardinal λ such that $\operatorname{cf} \alpha = \lambda$, for all $\alpha \in S$, or*

(ii) *every $\alpha \in S$ is a regular cardinal, the functions γ_α are continuous, and $S \cap \operatorname{rng} \gamma_\alpha = \varnothing$,*

then there exists an ordinal $\beta < \kappa$ such that, for every $\xi < \kappa$, the set

$$U_\xi := \{\, \alpha \in S \mid \operatorname{cf} \alpha > \beta \text{ and } \gamma_\alpha(\beta) \geq \xi \,\}$$

is stationary.

Proof. For a contradiction, suppose otherwise. Then we can find, for every $\beta < \kappa$, an ordinal ξ_β and a closed unbounded set C_β such that $U_{\xi_\beta} \cap C_\beta = \varnothing$, that is,

$$\gamma_\alpha(\beta) < \xi_\beta\,, \quad \text{for all } \alpha \in S \cap C_\beta \text{ such that } \operatorname{cf} \alpha > \beta\,.$$

In Case (i) we set $\zeta := \sup_{\beta<\lambda} \xi_\beta$ and $D := \bigcap_{\beta<\lambda} C_\beta$. Then $\gamma_\alpha(\beta) < \zeta$, for all $\beta < \lambda$ and $\alpha \in S \cap D$. Choosing $\alpha \in S \cap D$ with $\alpha > \zeta$ it follows that $\sup_{\beta<\lambda} \gamma_\alpha(\beta) \leq \zeta < \alpha$. A contradiction to the cofinality of γ_α.

It remains to consider Case (ii). Let D be the diagonal intersection of $(C_\beta)_{\beta<\kappa}$. Then $\alpha \in S \cap D$ implies that $\alpha \in S \cap C_\beta$, for all $\beta < \alpha$. Hence,

$$\gamma_\alpha(\beta) < \xi_\beta\,, \quad \text{for } \beta < \alpha\,.$$

The set

$$E := \{\, \alpha \in D \mid \xi_\beta < \alpha \text{ for all } \beta < \alpha \,\}$$

is closed unbounded since it can be written as the intersection of D and the diagonal intersection of the sets $\uparrow\xi_\beta$, $\beta < \kappa$, which are clearly closed unbounded. Hence, it follows by Lemma 6.8 (c) that $S \cap E$ is stationary. Let $\delta < \varepsilon$ be two elements of $S \cap E$. Then

$$\beta < \delta \quad \text{implies} \quad \gamma_\varepsilon(\beta) < \xi_\beta < \delta\,,$$

where the first inequality follows since $\varepsilon \in S \cap D$ and the second one follows since $\delta \in E$. By continuity of γ_ε,

$$\gamma_\varepsilon(\delta) = \sup_{\beta<\delta} \gamma_\varepsilon(\beta) \leq \delta\,.$$

Since γ_ε is strictly increasing, it therefore follows by Lemma A3.1.7 that $\gamma_\varepsilon(\delta) = \delta$. But $\delta \in S$ and $\gamma_\varepsilon(\delta) \in \operatorname{rng} \gamma_\varepsilon \subseteq \kappa \setminus S$. A contradiction. □

The first step in partitioning a stationary set into κ many stationary subsets consists in finding a decreasing chain of stationary subsets.

Lemma 6.14. *Let κ be an uncountable regular cardinal. For every stationary set $S \subseteq \kappa$, there exists a stationary subset $U \subseteq S$ and a function $f : U \to \kappa$ such that $f(\alpha) < \alpha$, for all $\alpha \in U$, and*

$$f^{-1}[\Uparrow\xi] \text{ is stationary,} \quad \text{for all } \xi < \kappa\,.$$

Proof. Consider the function

$$g : S \setminus \{0\} \to \kappa : \alpha \mapsto \begin{cases} \operatorname{cf} \alpha & \text{if } \operatorname{cf} \alpha < \alpha\,, \\ 0 & \text{if } \operatorname{cf} \alpha = \alpha\,. \end{cases}$$

Then $g(\alpha) < \alpha$, for all $\alpha \in S \setminus \{0\}$, and we can use the Theorem of Fodor to obtain a cardinal $\lambda < \kappa$ such that $T := g^{-1}(\lambda)$ is stationary. We distinguish two cases.

First, suppose that $\lambda > 0$. Note that the set T contains a limit ordinal, as the set of all limit ordinals is closed unbounded. This implies that λ is infinite. Therefore, for every $\alpha \in T$, we can choose by Lemma 4.11, a cofinal, strictly increasing function $\gamma_\alpha : \lambda \to \alpha$. By Lemma 6.13 , there exists an ordinal $\beta < \lambda$ such that, for every $\xi < \kappa$, the set

$$U_\xi := \{\, \alpha \in T \mid \gamma_\alpha(\beta) \geq \xi \,\}$$

is stationary. Hence, we can set $U := T$ and define $f : T \to \kappa$ by

$$f(\alpha) := \gamma_\alpha(\beta)\,.$$

If $\lambda = 0$, the set T consists of regular cardinals and Lemma 6.12 implies that the set

$$W := \{ \alpha \in T \mid T \cap \alpha \text{ is not stationary in } \alpha \}$$

is stationary. For every $\alpha \in W$, we fix a closed unbounded set $C_\alpha \subseteq \alpha$ with $(T \cap \alpha) \cap C_\alpha = \emptyset$. Since C_α is well-ordered, there exists an order-isomorphism $\gamma_\alpha : \beta \to C_\alpha$, for some ordinal β. Note that β cannot be smaller than α, because γ_α is cofinal in α and α is regular. Therefore, $\gamma_\alpha : \alpha \to C_\alpha$. Furthermore, $\sup_{\beta<\delta} \gamma_\alpha(\beta) \in C_\alpha$, for each limit ordinal $\delta < \alpha$, since C_α is closed unbounded. Consequently, $\sup_{\beta<\delta} \gamma_\alpha(\beta)$ is the least element of C_α that is larger than every $\gamma_\alpha(\beta)$ with $\beta < \delta$. As this element is $\gamma_\alpha(\delta)$, we obtain

$$\sup_{\beta<\delta} \gamma_\alpha(\beta) = \gamma_\alpha(\delta)\,.$$

Hence, each γ_α is a strictly continuous function with $W \cap \operatorname{rng} \gamma_\alpha = \emptyset$. We can therefore use Lemma 6.13 to find an ordinal $\beta < \kappa$ such that, for every $\xi < \kappa$, the set

$$U_\xi := \{ \alpha \in W \mid \alpha > \beta \text{ and } \gamma_\alpha(\beta) \geq \xi \}$$

is stationary. Thus, we can set $U := W \cap {\uparrow}\beta$ and define $f : U \to \kappa$ by $f(\alpha) := \gamma_\alpha(\beta)$. □

Theorem 6.15 (Solovay). *Let κ be an uncountable regular cardinal. Every stationary set $S \subseteq \kappa$ can be written as a disjoint union of κ stationary subsets of κ.*

Proof. By Lemma 6.14, there exists a stationary subset $U \subseteq S$ and a function $f : U \to \kappa$ such that $f(\alpha) < \alpha$ and the sets $U_\xi := f^{-1}[\Uparrow\xi]$ are stationary, for all $\xi < \kappa$. Applying the Theorem of Fodor to each restriction $f \restriction U_\xi$, we obtain ordinals $\alpha_\xi < \kappa$ such that the sets $W_\xi := (f \restriction U_\xi)^{-1}(\alpha_\xi)$ are stationary, for all $\xi < \kappa$. Note that $W_\xi \cap W_\zeta = \emptyset$, if

$\alpha_\xi \neq \alpha_\zeta$. Furthermore, $W_\xi \neq \varnothing$ implies that $\alpha_\xi \geq \xi$. Hence, $\sup_{\xi<\kappa} \alpha_\xi = \kappa$ and it follows by regularity of κ that

$$\left|\{ W_\xi \mid \xi < \kappa \}\right| = \left|\{ \alpha_\xi \mid \xi < \kappa \}\right| = \kappa \,.$$

Thus, we have found a family of κ disjoint stationary subsets of S. Since every superset of a stationary set is also stationary, we can enlarge these subsets to obtain the desired partition of S. □

7. Conclusion

With the Axiom of Replacement we have introduced our final axiom. The theory consisting of the six axioms

- Extensionality
- Separation
- Infinity
- Creation
- Choice
- Replacement

is called *Zermelo-Fraenkel set theory,* ZFC for short.

We can classify these axioms into three parts. The Axioms of Extensionality and Creation specify what we mean by a set. They postulate that every set is uniquely determined by its elements and that the membership relation is well-founded. The remaining axioms speak about the existence of certain sets. Infinity and Replacement ensure that the cumulative hierarchy is long enough. There are as many stages as there are ordinals. The Axioms of Separation and Choice on the other hand make the hierarchy wide by ensuring that the power-set operation yields enough subsets. In particular, every definable subset exists and on every set there exists a well-ordering.

Finally, let us note that the usual definition of ZFC is based on a different axiomatisation where the Axiom of Creation is replaced by four other axioms and the Axiom of Infinity is stated in a slightly different way. Nevertheless, we are justified in calling the above theory ZFC since the two variants are equivalent: every model satisfying one of the axiom systems also satisfies the other one, and vice versa.

Part B.

General Algebra

B1. Structures and homomorphisms

1. Structures

We have seen how to define graphs and partial orders in set theory. By a straightforward generalisation, we obtain other such structures like groups, fields, or vector spaces. A graph is a set equipped with one binary relation. In general, we allow arbitrary many relations and functions of arbitrary arities. To keep track of which relations and functions are present in a given structure we assign a name to each of them. These names are called *symbols,* the set of all symbols is called a *signature.*

Definition 1.1. A *signature* Σ is a set of relation symbols and function symbols each of which has a fixed (finite) *arity.* We call Σ *relational* if it contains only relation symbols and it is *functional* or *algebraic* if all of its elements are function symbols. A function symbol of arity 0 is also called a *constant symbol.*

Definition 1.2. Let Σ be a signature. A *Σ-structure* $\mathfrak{A}$ consists of

- a set A called the *universe* of $\mathfrak{A}$,
- an n-ary relation $R^{\mathfrak{A}} \subseteq A^n$, for each relation symbol $R \in \Sigma$ of arity n, and
- an n-ary function $f^{\mathfrak{A}} : A^n \to A$, for each function symbol $f \in \Sigma$ of arity n.

Formally, we can define a structure to be a pair $\langle A, \sigma \rangle$ where A is the universe and σ a function $\xi \mapsto \xi^{\mathfrak{A}}$ mapping each symbol $\xi \in \Sigma$ to the relation or function it denotes. But usually, in particular if the signature is finite, we will write structures simply as tuples

$$\mathfrak{A} = \langle A, R_0^{\mathfrak{A}}, R_1^{\mathfrak{A}}, \dots, f_0^{\mathfrak{A}}, f_1^{\mathfrak{A}}, \dots \rangle .$$

We will denote structures by fraktur letters $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \dots$ and their universes by the corresponding roman letters A, B, C, ….

Example. (a) A group G can be seen as structure $\langle G, \cdot \rangle$ where the binary function $\cdot : G \times G \to G$ denotes the group multiplication. Another possibility would be to take the richer structure $\langle G, \cdot, {}^{-1}, e \rangle$ where e is the unit of G and ${}^{-1} : G \to G$ the inverse.

(b) Similarly, a field K corresponds to a structure $\langle K, +, \cdot, 0, 1 \rangle$ with two binary functions and two constants.

The above definition of a structure is still not quite general enough. For instance, vector spaces fit only with some acrobatics into this framework.

Example. When we want to model a K-vector space V as a structure we face the problem of which set should be taken for the universe. One possibility is to define the structure $\langle V, +, (\lambda_a)_{a \in K} \rangle$ where the universe just consists of the vectors and, for each field element $a \in K$, we add a function $\lambda_a : V \to V : v \mapsto av$ for scalar multiplication with a. This formalism is mainly suited if one is interested in K-vector spaces for a fixed field K.

Another way of encoding vector spaces that treats K and V equally is to choose the structure $\langle V \cup K, V, K, A, M \rangle$ where the universe consists of the union of K and V, we have two unary predicates V and K that can be used to determine which elements are vectors and which are field elements, and there are two ternary relations $A \subseteq V \times V \times V$ and $M \subseteq K \times V \times V$ for vector addition and scalar multiplication. Note that we cannot use functions in this case since those would have to be defined for all elements of $(V \cup K) \times (V \cup K)$.

To make such codings unnecessary we extend the definition to allow structures that contain elements of different *sorts* like vectors and scalars.

Definition 1.3. Let S be a set and suppose that, for each $s \in S$, we are given some set A_s such that A_s and A_t are disjoint, for $s \neq t$. The elements of S will be called *sorts*.

(a) For $\bar{s} \subseteq S$, we write $A^{\bar{s}} := \prod_i A_{s_i}$.

(b) The *type* of an n-ary relation $R \subseteq A^{\bar{s}}$ is the sequence $\bar{s} \in S^n$.

(c) The *type* of an n-ary function $f : A^{\bar{s}} \to A_t$ is the pair $\langle \bar{s}, t \rangle \in S^n \times S$ which we will write more suggestively as $\bar{s} \to t$.

(d) If $A = \dot{\bigcup}_{s \in S} A_s$ and $B = \dot{\bigcup}_{s \in S} B_s$ are sets that are partitioned into sorts, we denote by B^A the set of all functions $f : A \to B$ such that $f[A_s] \subseteq B_s$, for all $s \in S$.

(e) An *S-sorted signature* Σ is a set of relation symbols and function symbols to each of which is assigned some type.

Definition 1.4. Let Σ be an S-sorted signature. A *Σ-structure* $\mathfrak{A}$ consists of

- a family of sets A_s, for $s \in S$,
- a relation $R^{\mathfrak{A}} \subseteq A^{\bar{s}}$ for each relation symbol $R \in \Sigma$ of type $\bar{s}$, and
- a function $f^{\mathfrak{A}} : A^{\bar{s}} \to A_t$ for every function symbol $f \in \Sigma$ of type $\bar{s} \to t$.

We call A_s the *domain of sort s*. The disjoint union $A := \dot{\bigcup}_{s \in S} A_s$ of all domains is the *universe* of $\mathfrak{A}$.

Example. We can model a K-vector space V as $\{s, v\}$-sorted structure

$$\langle K, V, +, \cdot, 0^V, 0^K, 1^K \rangle$$

where

- $+ : V \times V \to V$ of type $vv \to v$ is the addition of vectors,
- $\cdot : K \times V \to V$ of type $sv \to v$ is scalar multiplication, and
- $0^V \in V$ and $0^K, 1^K \in K$ are constants of type v, s, and s, respectively.

We could also add field addition and multiplication.

Lemma 1.5. *Let Σ be a signature and $\kappa \geq \aleph_0$. Up to isomorphism there are at most $2^{\kappa \oplus |\Sigma|}$ different Σ-structures $\mathfrak{A}$ of size $|A| = \kappa$.*

Proof. For every n, there are at most $2^{\kappa^n} = 2^\kappa$ n-ary relations $R \subseteq A^n$ and at most $\kappa^{\kappa^n} = 2^\kappa$ n-ary functions $f : A^n \to A$. Hence, the number of different Σ-structures is at most $(2^\kappa)^{|\Sigma|} = 2^{\kappa \oplus |\Sigma|}$. □

Many results in algebra and logic try to shed light on the 'internal structure' of some given Σ-structure $\mathfrak{A}$. A typical result of this kind could, for instance, state that every structure in a given class is built up from smaller structures in a certain way. In the remainder of this section we look at a given structure and try to find all structures that are contained in it.

Definition 1.6. Let Σ be an S-sorted signature and $\mathfrak{A}$ and $\mathfrak{B}$ Σ-structures.
(a) We write $\mathfrak{A} \subseteq \mathfrak{B}$ if

$$\begin{aligned} & A_s \subseteq B_s\,, && \text{for each sort } s \in S\,, \\ & R^{\mathfrak{A}} = R^{\mathfrak{B}} \cap A^n, && \text{for every } n\text{-ary relation symbol } R \in \Sigma\,, \\ \text{and}\quad & f^{\mathfrak{A}} = f^{\mathfrak{B}} \cap A^{n+1}, && \text{for every } n\text{-ary function symbol } f \in \Sigma\,. \end{aligned}$$

If $\mathfrak{A} \subseteq \mathfrak{B}$ then we say that $\mathfrak{A}$ is a *substructure* of $\mathfrak{B}$ and that $\mathfrak{B}$ is an *extension* of $\mathfrak{A}$. The set of all substructures of $\mathfrak{A}$ is denoted by $\mathrm{Sub}(\mathfrak{A})$, and we set

$$\mathfrak{Sub}(\mathfrak{A}) := \langle \mathrm{Sub}(\mathfrak{A}), \subseteq \rangle\,.$$

(b) Let $X \subseteq A$. If there is a substructure $\mathfrak{B} \subseteq \mathfrak{A}$ with universe $B = X$ then we say that X *induces* the substructure $\mathfrak{B}$. We denote this substructure by $\mathfrak{A}|_X$.

Example. $\mathfrak{N} = \langle \mathbb{N}, +, 0 \rangle$ is a substructure of $\mathfrak{Z} = \langle \mathbb{Z}, +, 0 \rangle$.

Remark. (a) Note that the preceding example shows that if $\mathfrak{G} = \langle G, \cdot \rangle$ is a group and $\mathfrak{H} \subseteq \mathfrak{G}$ a substructure then $\mathfrak{H}$ is not necessarily a subgroup of $\mathfrak{G}$. If, on the other hand, we consider groups with the richer signature $\langle G, \cdot, {}^{-1}, e \rangle$ then every substructure is also a subgroup.
(b) If the signature is relational then every set induces a substructure.

(c) Since a substructure is uniquely determined by its universe we will not always distinguish between substructures and the sets inducing them.

What substructures does a given structure $\mathfrak{A}$ have?

Lemma 1.7. *Let $\mathfrak{A}$ be a Σ-structure. A set $X \subseteq A$ induces a substructure of $\mathfrak{A}$ if and only if X is closed under all functions of $\mathfrak{A}$, that is, we have*

$$f^{\mathfrak{A}}(\bar{a}) \in X\,, \quad \textit{for every } n\textit{-ary function } f \in \Sigma \textit{ and all } \bar{a} \in X^n\,.$$

Proof. Suppose that X induces the substructure $\mathfrak{A}_0 \subseteq \mathfrak{A}$. For $f \in \Sigma$ and $\bar{a} \in X^n = A_0^n$ it follows that

$$f^{\mathfrak{A}}(\bar{a}) = f^{\mathfrak{A}_0}(\bar{a}) \in A_0 = X\,.$$

Conversely, if X is closed under functions then we can define the desired substructure $\mathfrak{A}_0$ by setting

$$\begin{aligned} R^{\mathfrak{A}_0} &:= R^{\mathfrak{A}} \cap X^n, && \text{for every } n\text{-ary relation } R \in \Sigma\,, \\ f^{\mathfrak{A}_0} &:= f^{\mathfrak{A}} \cap X^{n+1}, && \text{for every } n\text{-ary function } f \in \Sigma\,. \end{aligned}$$ □

Lemma 1.8. *Let $\mathfrak{A}$ be a Σ-structure and $Z \subseteq \mathscr{P}(A)$. If every element of Z induces a substructure of $\mathfrak{A}$ then so does $\bigcap Z$.*

Proof. Let $f \in \Sigma$ be an n-ary relation symbol and $\bar{a} \in (\bigcap Z)^n$. Since every element $X \in Z$ induces a substructure of $\mathfrak{A}$ it follows that $\bar{a} \subseteq X$ implies $f^{\mathfrak{A}}(\bar{a}) \in X$. Hence, $f^{\mathfrak{A}}(\bar{a}) \in \bigcap Z$. By Lemma 1.7, it follows that $\bigcap Z$ induces a substructure. □

Since the family of substructures is closed under intersection we can use Lemma A2.4.8 to characterise $\mathrm{Sub}(\mathfrak{A})$ via a closure operator.

Definition 1.9. Let $\mathfrak{A}$ be a Σ-structure.

(a) The substructure of $\mathfrak{A}$ *generated* by a set $X \subseteq A$ is $\langle\!\langle X \rangle\!\rangle_{\mathfrak{A}} := \mathfrak{A}|_Z$ where

$$Z := \bigcap \{\, B \mid B \supseteq X \text{ induces a substructure of } \mathfrak{A} \,\}\,.$$

(b) If $\langle\!\langle X \rangle\!\rangle_{\mathfrak{A}} = \mathfrak{A}$ then we say that X *generates* $\mathfrak{A}$ and we call the elements of X *generators* of $\mathfrak{A}$. If $\mathfrak{A}$ is generated by a finite set then we call $\mathfrak{A}$ *finitely generated.*

Example. (a) The structure $\mathfrak{N} = \langle \mathbb{N}, +, 0 \rangle$ is finitely generated by $\{1\}$.

(b) Let $\mathfrak{Z} = \langle \mathbb{Z}, +, - \rangle$ be the additive group of the integers. The set $X := \{5\}$ generates the substructure

$$\mathfrak{A} := \langle\!\langle X \rangle\!\rangle_{\mathfrak{Z}} = \langle A, +, - \rangle \quad \text{with} \quad A = \{\, 5k \mid k \in \mathbb{Z} \,\}\,.$$

Note that X does not induce $\mathfrak{A}$ since $A \supset X$.

If we consider the structure $\mathfrak{Z}' = \langle \mathbb{Z}, + \rangle$ without negation then X generates the substructure

$$\mathfrak{B} := \langle\!\langle X \rangle\!\rangle_{\mathfrak{Z}'} = \langle B, + \rangle \quad \text{with} \quad B = \{\, 5k \mid k \in \mathbb{Z},\ k > 0 \,\}\,.$$

(c) Let $\mathfrak{V} = \langle V, +, (\lambda_a)_{a \in K} \rangle$ be a vector space encoded as untyped structure. If $X \subseteq V$ then $\langle\!\langle X \rangle\!\rangle_{\mathfrak{V}}$ is the subspace spanned by X. If, instead, we encode V as two-sorted structure

$$\mathfrak{V} = \langle K, V, +^V, \cdot^V, +^K, \cdot^K, 0^V, 0^K, 1^K \rangle\,,$$

where $+^V$ is vector addition, $\cdot^V$ scalar multiplication, and $+^K$ and $\cdot^K$ the field operations, then $\langle\!\langle X \rangle\!\rangle_{\mathfrak{V}}$ just consists of all linear combinations

$$\lambda_0 v_0 + \dots + \lambda_{n-1} v_{n-1}$$

where $v_0, \dots, v_{n-1} \in X$ and $\lambda_0, \dots, \lambda_{n-1} \in \mathbb{N}$.

Lemma 1.10. *Let $\mathfrak{A}$ be a Σ-structure. The function $c : X \mapsto \langle\!\langle X \rangle\!\rangle_{\mathfrak{A}}$ is a closure operator on A with finite character.*

Proof. It follows from Lemma A2.4.8 that c is a closure operator. It remains to prove that it has finite character. Let

$$Z := \bigcup \{\, \langle\!\langle X_0 \rangle\!\rangle_{\mathfrak{A}} \mid X_0 \subseteq X \text{ is finite} \,\}\,.$$

To prove that $c(X) = Z$ it is sufficient to show that Z induces a substructure of $\mathfrak{A}$. We use Lemma 1.7. Let f be an n-ary function symbol and $\bar{a} \in Z^n$. Then there exists a finite set $X_0 \subseteq X$ with $\bar{a} \subseteq \langle\!\langle X_0 \rangle\!\rangle_{\mathfrak{A}}$. Since $\langle\!\langle X_0 \rangle\!\rangle_{\mathfrak{A}}$ induces a substructure of $\mathfrak{A}$ it follows that

$$f^{\mathfrak{A}}(\bar{a}) \in \langle\!\langle X_0 \rangle\!\rangle_{\mathfrak{A}} \subseteq Z. \qquad \square$$

Corollary 1.11. *Let* $\mathfrak{A}$ *be a structure.*

(a) $\mathfrak{Sub}(\mathfrak{A})$ *forms a complete partial order.*

(b) *If* $Z \subseteq \mathrm{Sub}(\mathfrak{A})$ *then* $\bigcap Z \in \mathrm{Sub}(\mathfrak{A})$.

(c) *If* $C \subseteq \mathrm{Sub}(\mathfrak{A})$ *is a chain then* $\bigcup C \in \mathrm{Sub}(\mathfrak{A})$.

So far, we have considered structures obtained by removing elements from a given structure. Instead, we can also remove relations or functions.

Definition 1.12. (a) Let Σ and Σ^+ be signatures with $\Sigma \subseteq \Sigma^+$, and let $\mathfrak{A}$ be a Σ^+-structure. The Σ-*reduct* $\mathfrak{A}|_\Sigma$ of $\mathfrak{A}$ is the Σ-structure $\mathfrak{B}$ with the same universe as $\mathfrak{A}$ where $\xi^{\mathfrak{B}} = \xi^{\mathfrak{A}}$, for all symbols $\xi \in \Sigma$. If $\mathfrak{B} = \mathfrak{A}|_\Sigma$ we call $\mathfrak{A}$ an *expansion* of $\mathfrak{B}$.

(b) Let Σ be an S-sorted signature, $T \subseteq S$, and $\mathfrak{A}$ a Σ-structure. Let $\Gamma \subseteq \Sigma$ be the T sorted signature consisting of all elements of Σ whose type only contains sort from T. By $\mathfrak{A}|_T$ we denote the Γ-structure obtained from $\mathfrak{A}$ by removing all domains A_s with $s \in S \setminus T$ and all relations and function from $\Sigma \setminus \Gamma$.

Example. $\langle G, \cdot \rangle$ is a reduct of $\langle G, \cdot, {}^{-1}, e \rangle$. In general, a Σ-structure has $2^{|\Sigma|}$ reducts.

Remark. If $\mathfrak{A} \subseteq \mathfrak{B}$ then $\mathfrak{A}|_\Sigma \subseteq \mathfrak{B}|_\Sigma$.

Remark. Let $S \subseteq T$ be sets of sorts. Every S-sorted signature Σ is also T-sorted. Similarly, every S-sorted structure $\mathfrak{A}$ can be turned into a T-sorted structure by setting $A_t := \emptyset$, for $t \in T \setminus S$. In the following we will not distinguish between an S-sorted structure $\mathfrak{A}$ and the corresponding T-sorted one obtained in that way.

2. *Homomorphisms*

Similarly to graphs and partial orders we can compare two structures by defining a map between them. The notions of an increasing function and an isomorphism can be extended in a straightforward way to arbitrary structures. Since now we have several relations we need the symbols of the signature in order to know which relation of one structure corresponds to a given relation of the other structure.

In the following, given $\bar{a} \in A^n$ and $h : A \to B$ we will abbreviate $\langle h(a_0), \dots, h(a_{n-1}) \rangle$ by $h(\bar{a})$.

Definition 2.1. Let $\mathfrak{A}$ and $\mathfrak{B}$ be Σ-structures.

(a) A mapping $h : A \to B$ is a *homomorphism* if it satisfies the following conditions:

- $h(A_s) \subseteq B_s$, for every sort s.
- If $\bar{a} \in R^{\mathfrak{A}}$ then $h(\bar{a}) \in R^{\mathfrak{B}}$, for all $\bar{a} \subseteq A$ and every $R \in \Sigma$.
- $h(f^{\mathfrak{A}}(\bar{a})) = f^{\mathfrak{B}}(h(\bar{a}))$, for all $\bar{a} \subseteq A$ and every $f \in \Sigma$.

(b) A homomorphism $h : A \to B$ is *strict* if it further satisfies

- $\bar{a} \in R^{\mathfrak{A}}$ iff $h(\bar{a}) \in R^{\mathfrak{B}}$, for all $\bar{a} \subseteq A$ and every $R \in \Sigma$.

(c) A homomorphism $h : A \to B$ is *semi-strict* if, whenever $h(\bar{a}) \in R^{\mathfrak{B}}$ then there is some $\bar{a}' \in R^{\mathfrak{A}}$ with $h(\bar{a}') = h(\bar{a})$.

(d) An *embedding* is an injective strict homomorphism and an *isomorphism* is a bijective strict homomorphism. We write $\mathfrak{A} \cong \mathfrak{B}$ to indicate that there exists an isomorphism $\mathfrak{A} \to \mathfrak{B}$. Finally, an isomorphism $\mathfrak{A} \to \mathfrak{A}$ is called an *automorphism* of $\mathfrak{A}$.

(e) If there exists a surjective homomorphism $\mathfrak{A} \to \mathfrak{B}$, $\mathfrak{B}$ is called a *weak homomorphic image* of $\mathfrak{A}$. It is a *homomorphic image* of $\mathfrak{A}$ if the homomorphism is semi-strict.

Example. (a) Let $\mathfrak{A}$ and $\mathfrak{B}$ be partial orders. A function $f : A \to B$ is a homomorphism if and only if it is increasing, and f is a strict homomorphism if and only if it is strictly increasing.

(b) The function $\langle \omega, + \rangle \to \langle \omega, \cdot \rangle$ with $n \mapsto 2^n$ is an embedding.

(c) The function $\langle \omega, + \rangle \to \langle [5], + \rangle$ with $n \mapsto n \bmod 5$ is a strict homomorphism.

(d) If $\mathfrak{K} = \langle K, +, \cdot \rangle$ is a field and $\mathfrak{K}[x] = \langle K[x], +, \cdot \rangle$ the corresponding ring of polynomials then we have a homomorphism

$$f : K[x] \to K : p(x) \mapsto p(0)$$

mapping a polynomial to its value at $x = 0$.

Remark. A homomorphism $h : \mathfrak{A} \to \mathfrak{B}$ is strict if and only if

$$h^{-1}[R^{\mathfrak{B}}] = R^{\mathfrak{A}}, \quad \text{for every relation } R\,.$$

Similarly, h is semi-strict if and only if

$$h[R^{\mathfrak{A}}] = R^{\mathfrak{B}}, \quad \text{for every relation } R\,.$$

Exercise 2.1. Let $\mathfrak{N} = \langle \omega, \cdot \rangle$. Construct an automorphism $f : \mathfrak{N} \to \mathfrak{N}$ with $f(2) = 3$.

Lemma 2.2. *If $g : \mathfrak{A} \to \mathfrak{B}$ and $h : \mathfrak{B} \to \mathfrak{C}$ are isomorphisms then so are the functions $g^{-1} : \mathfrak{B} \to \mathfrak{A}$ and $h \circ g : \mathfrak{A} \to \mathfrak{C}$.*

Lemma 2.3. *Every injective semi-strict homomorphism $h : \mathfrak{A} \to \mathfrak{B}$ is strict.*

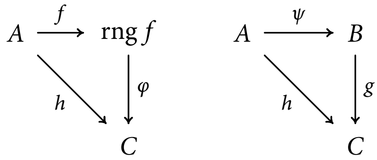
Proof. Suppose that $h(\bar{a}) \in R^{\mathfrak{B}}$. Then there is some tuple $\bar{a}' \in R^{\mathfrak{A}}$ with $h(\bar{a}') = h(\bar{a})$. Since h is injective, it follows that $\bar{a}' = \bar{a}$ and, hence, $\bar{a} \in R^{\mathfrak{A}}$. □

Definition 2.4. Let $f : \mathfrak{A} \to \mathfrak{B}$ be a function. The *kernel* of f is the relation

$$\ker f := \{ \langle a, b \rangle \in A^2 \mid f(a) = f(b) \}\,.$$

Remark. The kernel of a function is obviously an equivalence relation.

Lemma 2.5 (Factorisation Lemma). *Let $f : A \to B$, $g : B \to C$, and $h : A \to C$ be functions.*



(a) *There exists at most one function $\varphi : \operatorname{rng} f \to C$ with $h = \varphi \circ f$.*

(b) *If g is injective then there exists at most one function $\psi : A \to B$ with $h = g \circ \psi$.*

(c) *There exists a function $\varphi : \operatorname{rng} f \to C$ with $h = \varphi \circ f$ if and only if* $\ker f \subseteq \ker h$.

(d) *There exists a function $\psi : A \to B$ with $h = g \circ \psi$ if and only if* $\operatorname{rng} h \subseteq \operatorname{rng} g$.

Proof. (a) If $\varphi, \varphi' : \operatorname{rng} f \to C$ are functions such that $\varphi \circ f = g = \varphi' \circ f$ then, since $f : A \to \operatorname{rng} f$ is surjective, it follows by Lemma A2.1.10 that $\varphi = \varphi'$.

(b) If $\psi, \psi' : A \to B$ are functions such that $g \circ \psi = h = g \circ \psi'$ then, since $g : B \to C$ is injective, it follows by Lemma A2.1.10 that $\psi = \psi'$.

(c) ($\Rightarrow$) If $\langle a, a' \rangle \in \ker f$ then we have

$$h(a) = \varphi(f(a)) = \varphi(f(a')) = h(a'),$$

which implies that $\langle a, a' \rangle \in \ker h$.

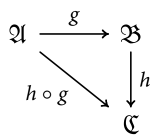
($\Leftarrow$) For $b \in \operatorname{rng} f$, select an arbitrary element $a \in f^{-1}(b)$ and set $\varphi(b) := g(a)$. We claim that $\varphi \circ f = g$. Let $a \in A$ and set $b := f(a)$. By definition of φ, we have $\varphi(b) = g(a')$, for some element $a' \in A$ with $f(a') = b$. Hence, $\langle a, a' \rangle \in \ker f \subseteq \ker g$, which implies that $g(a) = g(a')$. Consequently, we have

$$\varphi(f(a)) = \varphi(b) = g(a') = g(a).$$

(d) ($\Rightarrow$) If $c \in \operatorname{rng} h$ then there is some element $a \in A$ with $c = h(a)$ and $g(\psi(a)) = h(a) = c$ implies that $c \in \operatorname{rng} g$.

($\Leftarrow$) For $a \in A$, we have $h(a) \in \operatorname{rng} h \subseteq \operatorname{rng} g$. Hence, we can select some element $b \in g^{-1}(h(a))$ and we set $\psi(a) := b$. Then $g(\psi(a)) = g(b) = h(a)$. $\square$

Lemma 2.6. *Let $g : \mathfrak{A} \to \mathfrak{B}$ and $h : \mathfrak{B} \to \mathfrak{C}$ be functions.*



(a) *Suppose that g is a surjective semi-strict homomorphism.*

  (i) *If $h \circ g$ is a homomorphism then so is h.*

  (ii) *If $h \circ g$ is a semi-strict homomorphism then so is h.*

  (iii) *If $h \circ g$ is a strict homomorphism then so is h.*

(b) *Suppose that h is an injective semi-strict homomorphism.*

  (i) *If $h \circ g$ is a homomorphism then so is g.*

  (ii) *If $h \circ g$ is a semi-strict homomorphism then so is g.*

  (iii) *If $h \circ g$ is a strict homomorphism then so is g.*

Proof. (a) (i) Let $\bar{b} \in B^n$ and $a_i \in g^{-1}(b_i)$, for $i < n$. For an n-ary function symbol f, we have

$$f^{\mathfrak{C}}(h(\bar{b})) = f^{\mathfrak{C}}\big(h(g(\bar{a}))\big) = (h \circ g)(f^{\mathfrak{A}}(\bar{a})) = h\big(f^{\mathfrak{B}}(g(\bar{a}))\big) = h(f^{\mathfrak{B}}(\bar{b}))\,.$$

If R is an n-ary relation symbol with $\bar{b} \in R^{\mathfrak{B}}$ then, since g is semi-strict, we can find elements $a_i \in g^{-1}(b_i)$ such that $\bar{a} \in R^{\mathfrak{A}}$. This implies that $h(\bar{b}) = (h \circ g)(\bar{a}) \in R^{\mathfrak{C}}$.

(ii) For every relation R, we have $h[R^{\mathfrak{B}}] = h[g[R^{\mathfrak{A}}]] = R^{\mathfrak{C}}$.

(iii) Since g is surjective we have $g[g^{-1}[X]] = X$, for every $X \subseteq B$. It follows that

$$h^{-1}[R^{\mathfrak{C}}] = g[g^{-1}[h^{-1}[R^{\mathfrak{C}}]]] = g[R^{\mathfrak{A}}] = R^{\mathfrak{B}}.$$

(b) (i) Let $\bar{a} \in A^n$ and f an n-ary function symbol. Then we have

$$h\big(g(f^{\mathfrak{A}}(\bar{a}))\big) = f^{\mathfrak{C}}\big((h \circ g)(\bar{a})\big) = h\big(f^{\mathfrak{B}}(g(\bar{a}))\big).$$

Since h is injective it follows that $g(f^{\mathfrak{A}}(\bar{a})) = f^{\mathfrak{B}}(g(\bar{a}))$.

If R is an n-ary relation symbol with $\bar{a} \in R^{\mathfrak{A}}$ then we have $(h \circ g)(\bar{a}) \in R^{\mathfrak{C}}$ and, since h is semi-strict, there is some tuple $\bar{b} \in R^{\mathfrak{B}}$ with $h(\bar{b}) = h(g(\bar{a}))$. Since h is injective it follows that $g(\bar{a}) = \bar{b} \in R^{\mathfrak{B}}$.

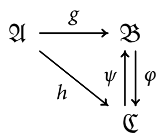
(ii) Since h is injective we have $h^{-1}[h[X]] = X$, for every $X \subseteq B$. Furthermore, injective semi-strict homomorphisms are strict. Therefore, we have

$$g[R^{\mathfrak{A}}] = h^{-1}[h[g[R^{\mathfrak{A}}]]] = h^{-1}[R^{\mathfrak{C}}] = R^{\mathfrak{B}}.$$

(iii) As in (ii) we have

$$g^{-1}[R^{\mathfrak{B}}] = g^{-1}[h^{-1}[h[R^{\mathfrak{B}}]]] = (h \circ g)^{-1}[R^{\mathfrak{C}}] = R^{\mathfrak{A}}. \qquad \square$$

Corollary 2.7. *If $g : \mathfrak{A} \to \mathfrak{B}$ and $h : \mathfrak{A} \to \mathfrak{C}$ are surjective semi-strict homomorphisms with* $\ker g = \ker h$ *then there exists a unique isomorphism $\varphi : \mathfrak{B} \to \mathfrak{C}$ with $h = \varphi \circ g$.*



Proof. By Lemmas 2.5 and 2.6 there exist unique semi-strict homomorphisms

$$\varphi : \mathfrak{B} \to \mathfrak{C} \quad \text{and} \quad \psi : \mathfrak{C} \to \mathfrak{B}$$

such that $h = \varphi \circ g$ and $g = \psi \circ h$. In the same way, $\ker g = \ker g$ implies that there exists a unique homomorphism $\eta : \mathfrak{B} \to \mathfrak{B}$ with with $g = \eta \circ g$. Since id and $\psi \circ \varphi$ both satisfy this equation it follows that $\psi \circ \varphi = \mathrm{id}$. In the same way we obtain $\varphi \circ \psi = \mathrm{id}$. Consequently, φ is an isomorphism. □

We can use a homomorphism $h : \mathfrak{A} \to \mathfrak{B}$ to compare the family of substructures of $\mathfrak{A}$ to that of $\mathfrak{B}$.

Lemma 2.8. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be Σ-structures and $h : \mathfrak{A} \to \mathfrak{B}$ a homomorphism.*

(a) *If $\mathfrak{A}_0 \subseteq \mathfrak{A}$ then $h[A_0]$ induces a substructure of $\mathfrak{B}$.*

(b) *If $\mathfrak{B}_0 \subseteq \mathfrak{B}$ then $h^{-1}[B_0]$ induces a substructure of $\mathfrak{A}$.*

(c) *If $X \subseteq A$ then $h[\langle\!\langle X \rangle\!\rangle_{\mathfrak{A}}] = \langle\!\langle h[X] \rangle\!\rangle_{\mathfrak{B}}$.*

Proof. (a) We have to show that $B_0 := h[A_0]$ is closed under all functions of $\mathfrak{B}$. Let $f \in \Sigma$ be n-ary and $b_0, \dots, b_{n-1} \in B_0$. There exist elements $a_0, \dots, a_{n-1} \in A_0$ such that $b_i = h(a_i)$, for $i < n$. Since A_0 is closed under f we have $f^{\mathfrak{A}}(\bar{a}) \in A_0$, which implies that

$$\begin{aligned} f^{\mathfrak{B}}(b_0, \dots, b_{n-1}) &= f^{\mathfrak{B}}(ha_0, \dots, ha_{n-1}) \\ &= h(f^{\mathfrak{A}}(a_0, \dots, a_{n-1})) \in B_0 \,. \end{aligned}$$

(b) Set $A_0 := h^{-1}[B_0]$. By (a) and Corollary 1.11, we know that the sets $C := \operatorname{rng} h$ and $B_1 := B_0 \cap C$ induce substructures of $\mathfrak{B}$. Note that we have $A_0 = h^{-1}[B_1]$. Let $f \in \Sigma$ be n-ary and $a_0, \dots, a_{n-1} \in A_0$. Then $h(a_i) \in B_1$ implies $f^{\mathfrak{B}}(h(a_0), \dots, h(a_{n-1})) \in B_1$. Since

$$h(f^{\mathfrak{A}}(a_0, \dots, a_{n-1})) = f^{\mathfrak{B}}(ha_0, \dots, ha_{n-1}) \in B_1$$

it follows that $f^{\mathfrak{A}}(\bar{a}) \in h^{-1}[B_1] = A_0$.

(c) By (a) we know that $h[\langle\!\langle X \rangle\!\rangle_{\mathfrak{A}}]$ induces a substructure of $\mathfrak{B}$ containing $h[X]$. Hence,

$$\langle\!\langle h[X] \rangle\!\rangle_{\mathfrak{B}} \subseteq h[\langle\!\langle X \rangle\!\rangle_{\mathfrak{A}}] \,.$$

Conversely, set $Y := \langle\!\langle h[X] \rangle\!\rangle_{\mathfrak{B}}$. By (b), $h^{-1}[Y]$ induces a substructure of $\mathfrak{A}$ with $X \subseteq h^{-1}[Y]$. Consequently, we have $\langle\!\langle X \rangle\!\rangle_{\mathfrak{A}} \subseteq h^{-1}[Y]$, which implies that

$$h[\langle\!\langle X \rangle\!\rangle_{\mathfrak{A}}] \subseteq h[h^{-1}[Y]] = Y = \langle\!\langle h[X] \rangle\!\rangle_{\mathfrak{B}}\,. \qquad \square$$

Corollary 2.9. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be Σ-structures. If $h : \mathfrak{A} \to \mathfrak{B}$ is a homomorphism then* rng h *induces a substructure of $\mathfrak{B}$.*

Definition 2.10. Let $h : \mathfrak{A} \to \mathfrak{B}$ be a homomorphism between Σ-structures $\mathfrak{A}$ and $\mathfrak{B}$. For a substructure $\mathfrak{A}_0 \subseteq \mathfrak{A}$, we denote by $h(\mathfrak{A}_0)$ the substructure of $\mathfrak{B}$ induced by $h[A_0]$.

3. *Categories*

Many algebraic properties can be expressed in terms of homomorphisms between structures. Category theory provides a general framework for doing so.

Definition 3.1. A *category* $\mathcal{C}$ consists of

- a class $\mathcal{C}^{\text{obj}}$ of *objects*,
- for each pair of objects $\mathfrak{a}, \mathfrak{b} \in \mathcal{C}^{\text{obj}}$, a set $\mathcal{C}(\mathfrak{a}, \mathfrak{b})$ of *morphisms* from $\mathfrak{a}$ to $\mathfrak{b}$, and
- for all $\mathfrak{a}, \mathfrak{b}, \mathfrak{c} \in \mathcal{C}^{\text{obj}}$, an operation

$$\circ : \mathcal{C}(\mathfrak{b}, \mathfrak{c}) \times \mathcal{C}(\mathfrak{a}, \mathfrak{b}) \to \mathcal{C}(\mathfrak{a}, \mathfrak{c})\,,$$

such that the following conditions are satisfied:

(1) If $f \in \mathcal{C}(\mathfrak{c}, \mathfrak{d})$, $g \in \mathcal{C}(\mathfrak{b}, \mathfrak{c})$, $h \in \mathcal{C}(\mathfrak{a}, \mathfrak{b})$ then

$$f \circ (g \circ h) = (f \circ g) \circ h\,.$$

(2) For every $\mathfrak{a} \in \mathcal{C}^{\mathrm{obj}}$, there is a morphism $\mathrm{id}_\mathfrak{a} \in \mathcal{C}(\mathfrak{a}, \mathfrak{a})$ such that

$$\mathrm{id}_\mathfrak{a} \circ f = f\,, \quad \text{for all } f \in \mathcal{C}(\mathfrak{b}, \mathfrak{a})\,,$$
$$f \circ \mathrm{id}_\mathfrak{a} = f\,, \quad \text{for all } f \in \mathcal{C}(\mathfrak{a}, \mathfrak{b})\,.$$

We call $\mathrm{id}_\mathfrak{a}$ the *identity morphism* of $\mathfrak{a}$.

If the category is understood we will write $f : \mathfrak{a} \to \mathfrak{b}$ to indicate that $f \in \mathcal{C}(\mathfrak{a}, \mathfrak{b})$. By $\mathcal{C}^{\mathrm{mor}}$ we denote the class of all morphisms of $\mathcal{C}$, irrespective of their end-points. Instead of $\mathfrak{a} \in \mathcal{C}^{\mathrm{obj}}$, we also simply write $\mathfrak{a} \in \mathcal{C}$.

Example. (a) The category $\mathfrak{Set}$ consists of all sets where

$$\mathfrak{Set}(A, B) := B^A$$

and $\circ$ is the usual composition of functions.

(b) $\mathfrak{Hom}(\Sigma)$ is the category of all Σ-structures where $\mathfrak{Hom}(\Sigma)(\mathfrak{A}, \mathfrak{B})$ is the set of homomorphisms $\mathfrak{A} \to \mathfrak{B}$. Similarly, we can form the category $\mathfrak{Hom}_s(\Sigma)$ of all Σ-structures where the morphisms are strict homomorphisms, and the category $\mathfrak{Emb}(\Sigma)$ of embeddings.

(c) $\mathfrak{Grp}$ is the subcategory of $\mathfrak{Hom}(\,\cdot\,, {}^{-1}, e)$ consisting of all groups.

(d) In the category $\mathfrak{Set}_*$ of *pointed sets* the objects are pairs $\langle A, a\rangle$ where A is a set and $a \in A$. A morphism $f : \langle A, a\rangle \to \langle B, b\rangle$ is a function $f : A \to B$ such that $f(a) = b$.

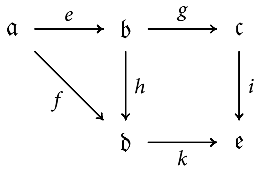
(e) Similarly, in the category $\mathfrak{Set}^2$ the objects are pairs $\langle A, A_0\rangle$ of sets with $A_0 \subseteq A$ and a morphism $f : \langle A, A_0\rangle \to \langle B, B_0\rangle$ is a function $f : A \to B$ such that $f[A_0] \subseteq B_0$.

(f) We have categories $\mathfrak{Top}$ and $\mathfrak{Top}^2$ of topological spaces and pairs of such spaces where the morphisms are continuous functions.

(g) We can consider every partial order $\mathfrak{A} = \langle A, \leq\rangle$ as a category where the objects are the elements of $\mathfrak{A}$ and the morphisms are

$$\mathfrak{A}(a, b) := \begin{cases} \{\langle a, b\rangle\} & \text{if } a \leq b\,, \\ \varnothing & \text{otherwise}\,. \end{cases}$$

Almost all statements in category theory involve equations of the form $f \circ g = h \circ k$. When there are many of them a graphical presentation comes handy. Usually, we will use diagrams of the form



We say that such a diagram *commutes* if, for every pair of paths starting at the same object and ending at the same one, the equation

$$f_m \circ \cdots \circ f_0 = g_n \circ \cdots \circ g_0$$

holds, where $f_0, \ldots, f_m$ and $g_0, \ldots, g_n$ are the respective labels along the two paths. For example, the above diagram commutes if the following equations hold:

$$h \circ e = f\,, \quad i \circ g = k \circ h\,, \quad i \circ g \circ e = k \circ f\,.$$

(The last one is actually redundant.)

Lemma 3.2. *Let $\mathcal{C}$ be a category. For each object $\mathfrak{a} \in \mathcal{C}^{\mathrm{obj}}$, there is a unique identity morphism* $\mathrm{id}_\mathfrak{a} \in \mathcal{C}(\mathfrak{a}, \mathfrak{a})$.

Proof. If $\mathrm{id}_\mathfrak{a}$ and $\mathrm{id}'_\mathfrak{a}$ are identity morphisms of $\mathfrak{a}$ then

$$\mathrm{id}_\mathfrak{a} = \mathrm{id}_\mathfrak{a} \circ \mathrm{id}'_\mathfrak{a} = \mathrm{id}'_\mathfrak{a}\,. \qquad \square$$

Although the morphisms of a category need not to be functions we can generalise many concepts from functions to arbitrary categories. For instance, we can use the characterisation of Lemma A2.1.10 to generalise the notion of injectivity and surjectivity.

Definition 3.3. (a) A morphism $f : \mathfrak{a} \to \mathfrak{b}$ is a *monomorphism* if, for all morphisms g and h,

$$f \circ g = f \circ h \quad \text{implies} \quad g = h\,.$$

And f is an *epimorphism* if

$$g \circ f = h \circ f \quad \text{implies} \quad g = h\,.$$

(b) If $f : \mathfrak{a} \to \mathfrak{b}$ and $g : \mathfrak{b} \to \mathfrak{a}$ are morphisms with $g \circ f = \mathrm{id}_{\mathfrak{a}}$, we call g a *left inverse* of f and f a *right inverse* of g. In this situation we also say that f is a *section* and g is a *retraction*. An *inverse* of f is a morphism g that is both a left and a right inverse of f. If $f : \mathfrak{a} \to \mathfrak{b}$ has an inverse, we denote it by $f^{-1} : \mathfrak{b} \to \mathfrak{a}$ and we call f an *isomorphism* between $\mathfrak{a}$ and $\mathfrak{b}$.

Example. In many categories where the morphisms are actual functions, monomorphisms correspond to injective functions and epimorphisms correspond to surjective functions. For instance, in $\mathfrak{Set}$ and in $\mathfrak{Hom}(\Sigma)$ this is the case. But there are also examples where monomorphisms are not injective or epimorphisms are not surjective. For instance, in the category of all rings the inclusion homomorphism $h : \mathbb{Z} \to \mathbb{Q}$ is an epimorphism since a homomorphism $f : \mathbb{Q} \to \mathfrak{R}$ is uniquely determined by its restriction $f \restriction \mathbb{Z}$. Similarly, in the category of all Hausdorff spaces with continuous maps as morphisms a morphism $f : \mathfrak{X} \to \mathfrak{Y}$ is an epimorphism if, and only if, its image $\mathrm{rng}\, f$ is dense in Y.

Lemma 3.4. (a) *Every section is a monomorphism.*

(b) *Every retraction an epimorphism.*

(c) *Every epimorphism with a left inverse is an isomorphism.*

(d) *Every monomorphism with a right inverse is an isomorphism.*

(e) *If a morphism f has a left inverse g and a right inverse h then f is an isomorphism and $g = h$.*

Proof. (a) and (b) are left as an exercise.

(c) Let $f : \mathfrak{a} \to \mathfrak{b}$ be an epimorphism with left inverse $g : \mathfrak{b} \to \mathfrak{a}$. Then $g \circ f = \mathrm{id}_{\mathfrak{a}}$ implies that $f \circ g \circ f = f = \mathrm{id}_{\mathfrak{b}} \circ f$. As f is an epimorphism, this implies that $f \circ g = \mathrm{id}_{\mathfrak{b}}$. Hence, g is an inverse of f.

(d) follows in the same way as (c).

(e) We have $g = g \circ \mathrm{id}_{\mathfrak{b}} = g \circ (f \circ h) = (g \circ f) \circ h = \mathrm{id}_{\mathfrak{a}} \circ h = h$. □

Exercise 3.1. Let $f : \mathfrak{a} \to \mathfrak{b}$ and $g : \mathfrak{b} \to \mathfrak{c}$ be morphisms. Show that

(a) if f and g are monomorphisms then so is $g \circ f$;

(b) if f and g are epimorphisms then so is $g \circ f$.

Most statements of category theory also hold if every morphism is reversed. To avoid duplicating proofs we introduce the notion of the opposite of a category.

Definition 3.5. Let $\mathcal{C}$ be a category. The *opposite* of $\mathcal{C}$ is the category $\mathcal{C}^{\mathrm{op}}$ with the same objects as $\mathcal{C}$. For each morphism $f : \mathfrak{a} \to \mathfrak{b}$ of $\mathcal{C}$ there exists the morphism $f^{\mathrm{op}} : \mathfrak{b} \to \mathfrak{a}$ in $\mathcal{C}^{\mathrm{op}}$. The composition of such morphisms is defined by

$$g^{\mathrm{op}} \circ f^{\mathrm{op}} := (f \circ g)^{\mathrm{op}}.$$

Definition 3.6. An object $\mathfrak{a} \in \mathcal{C}$ is *initial* if, for every $\mathfrak{b} \in \mathcal{C}$, there exists a unique morphism $\mathfrak{a} \to \mathfrak{b}$. Similarly, we call $\mathfrak{a}$ *terminal* if there exist unique morphisms $\mathfrak{b} \to \mathfrak{a}$, for all $\mathfrak{b} \in \mathcal{C}$.

Example. (a) $\mathfrak{Set}$ contains one initial object $\varnothing$, while every singleton $\{x\}$ is terminal.

(b) The trivial group $\{e\}$ is both initial and terminal in $\mathfrak{Grp}$.

The importance of initial and terminal objects stems from the fact that, up to isomorphism, they are unique.

Lemma 3.7. *Let $\mathcal{C}$ be a category. All initial objects of $\mathcal{C}$ are isomorphic and all terminal objects are isomorphic.*

Proof. Note that a terminal object in $\mathcal{C}$ is an initial object in $\mathcal{C}^{\mathrm{op}}$. Therefore, it is sufficient to prove the claim for initial objects. Suppose that $\mathfrak{a}$ and $\mathfrak{b}$ are initial objects in $\mathcal{C}$. Then there exist unique morphisms $f : \mathfrak{a} \to \mathfrak{b}$ and $g : \mathfrak{b} \to \mathfrak{a}$. Let $h := g \circ f$. Then $h : \mathfrak{a} \to \mathfrak{a}$ and h is the only morphism $\mathfrak{a} \to \mathfrak{a}$ since $\mathfrak{a}$ is initial. It follows that $h = \mathrm{id}_{\mathfrak{a}}$. By a symmetric argument, it follows that $f \circ g = \mathrm{id}_{\mathfrak{b}}$. Consequently, g is an inverse of f and f is an isomorphism. □

To compare two categories we need the notion of a 'homomorphism' between categories.

Definition 3.8. (a) A (*covariant*) *functor* F from a category $\mathcal{C}$ to a category $\mathcal{D}$ consists of two functions

$$F^{\mathrm{obj}} : \mathcal{C}^{\mathrm{obj}} \to \mathcal{D}^{\mathrm{obj}} \quad \text{and} \quad F^{\mathrm{mor}} : \mathcal{C}^{\mathrm{mor}} \to \mathcal{D}^{\mathrm{mor}}$$

such that the following conditions are satisfied:

- F^{mor} maps each morphism $f : \mathfrak{a} \to \mathfrak{b}$ in $\mathcal{C}$ to a morphism

  $$F^{\mathrm{mor}}(f) : F^{\mathrm{obj}}(\mathfrak{a}) \to F^{\mathrm{obj}}(\mathfrak{b}) \quad \text{in } \mathcal{D}.$$

- $F^{\mathrm{mor}}(\mathrm{id}_{\mathfrak{a}}) = \mathrm{id}_{F^{\mathrm{obj}}(\mathfrak{a})}$, for all $\mathfrak{a} \in \mathcal{C}^{\mathrm{obj}}$.
- $F^{\mathrm{mor}}(g \circ f) = F^{\mathrm{mor}}(g) \circ F^{\mathrm{mor}}(f)$, for all $f : \mathfrak{a} \to \mathfrak{b}$ and $g : \mathfrak{b} \to \mathfrak{c}$ in $\mathcal{C}^{\mathrm{mor}}$.

Usually we will omit the superscripts and just write F instead of F^{obj} and F^{mor}.

(b) A functor $F : \mathcal{C} \to \mathcal{D}$ is called *faithful* if, for every pair $\mathfrak{a}, \mathfrak{b} \in \mathcal{C}$, the induced map

$$F : \mathcal{C}(\mathfrak{a}, \mathfrak{b}) \to \mathcal{D}(F(\mathfrak{a}), F(\mathfrak{b}))$$

is injective. Similarly, $F : \mathcal{C} \to \mathcal{D}$ is called *full* if, for every pair $\mathfrak{a}, \mathfrak{b} \in \mathcal{C}$, the induced map

$$F : \mathcal{C}(\mathfrak{a}, \mathfrak{b}) \to \mathcal{D}(F(\mathfrak{a}), F(\mathfrak{b}))$$

is surjective.

(c) A *contravariant functor* form $\mathcal{C}$ to $\mathcal{D}$ is a covariant functor from $\mathcal{C}^{\mathrm{op}}$ to $\mathcal{D}$.

(d) The *opposite* of a functor $F : \mathcal{C} \to \mathcal{D}$ is the functor $F^{\mathrm{op}} : \mathcal{C}^{\mathrm{op}} \to \mathcal{D}^{\mathrm{op}}$ with

$$\begin{aligned} F^{\mathrm{op}}(\mathfrak{a}) &:= F(\mathfrak{a}), && \text{for } \mathfrak{a} \in \mathcal{C}^{\mathrm{obj}}, \\ F^{\mathrm{op}}(f^{\mathrm{op}}) &:= F(f)^{\mathrm{op}}, && \text{for } f \in \mathcal{C}^{\mathrm{mor}}. \end{aligned}$$

Example. (a) For a signature Σ, the *forgetful functor* $F : \mathfrak{Hom}(\Sigma) \to \mathfrak{Set}$ maps every structure $\mathfrak{A}$ to its universe A and every homomorphism $h : \mathfrak{A} \to \mathfrak{B}$ to the corresponding function $h : A \to B$ between the universes. This functor is faithful, but in general not full.

(b) Let $G : \mathfrak{Set} \to \mathfrak{Hom}(\varnothing)$ be the functor mapping a set X to the structure $\langle X \rangle$ over the empty signature. This functor is full and faithful. The forgetful functor $F : \mathfrak{Hom}(\varnothing) \to \mathfrak{Set}$ is an inverse of G. It follows that the categories $\mathfrak{Set}$ and $\mathfrak{Hom}(\varnothing)$ are isomorphic.

Definition 3.9. Let $F : \mathcal{C} \to \mathcal{D}$ be a functor and let P be a property of objects or morphisms.

(a) We say that F *preserves* P if, whenever x is an object or morphism with property P, then $F(x)$ also has this property.

(b) We say that F *reflects* P if, whenever x is an object or morphism such that $F(x)$ has property P, x also has this property.

Lemma 3.10. (a) *Every functor preserves sections, retractions, and isomorphisms.*

(b) *Faithful functors reflect monomorphisms and epimorphisms.*

(c) *Full and faithful functors reflect sections, retractions, and isomorphisms.*

Proof. Let $F : \mathcal{C} \to \mathcal{D}$ be a functor.

(a) Let $f : \mathfrak{a} \to \mathfrak{b}$ and $g : \mathfrak{b} \to \mathfrak{a}$ be morphisms of $\mathcal{C}$ such that $g \circ f = \mathrm{id}_{\mathfrak{a}}$. Then

$$F(g) \circ F(f) = F(g \circ f) = F(\mathrm{id}_{\mathfrak{a}}) = \mathrm{id}_{F(\mathfrak{a})}.$$

Hence, $F(g)$ is a left inverse of $F(f)$ and $F(f)$ is a right inverse of $F(g)$.

(b) Suppose that F is faithful and let $f : \mathfrak{a} \to \mathfrak{b}$ be a morphism such that $F(f)$ is a monomorphism. To show that f is also a monomorphism, consider morphisms $g, h : \mathfrak{c} \to \mathfrak{a}$ with $f \circ g = f \circ h$. Then

$$F(f) \circ F(g) = F(f \circ g) = F(f \circ h) = F(f) \circ F(h) .$$

Since $F(f)$ is a monomorphism, it follows that $F(g) = F(h)$. Because F is faithful, this implies that $g = h$.

In the same way it follows that F reflects epimorphisms.

(c) Suppose that F is faithful and full and let $F(f) : F(\mathfrak{a}) \to F(\mathfrak{b})$ be a section with left inverse $g : F(\mathfrak{b}) \to F(\mathfrak{a})$. As F is full, there exists a morphism $g_o : \mathfrak{b} \to \mathfrak{a}$ with $F(g_o) = g$. Hence,

$$F(\mathrm{id}_{\mathfrak{a}}) = \mathrm{id}_{F(\mathfrak{a})} = F(g_o) \circ F(f) = F(g_o \circ f).$$

Since F is faithful, this implies that $g_o \circ f = \mathrm{id}_{\mathfrak{a}}$. Consequently, f is a section. The cases where f is a retraction or an isomorphism follow in the same way. □

Let us briefly present some operations on categories.

Definition 3.11. Let $\mathcal{C}$ and $\mathcal{D}$ be categories.

(a) $\mathcal{C}$ is a *subcategory* of $\mathcal{D}$ if

- $\mathcal{C}^{\mathrm{obj}} \subseteq \mathcal{D}^{\mathrm{obj}}$ and $\mathcal{C}^{\mathrm{mor}} \subseteq \mathcal{D}^{\mathrm{mor}}$,
- the identity morphisms of $\mathcal{C}$ are the identity morphisms of $\mathcal{D}$,
- the composition $g \circ h$ of two morphisms of $\mathcal{C}$ gives the same result in both categories.

A subcategory $\mathcal{C} \subseteq \mathcal{D}$ is *full* if

$$\mathcal{C}(\mathfrak{a}, \mathfrak{b}) = \mathcal{D}(\mathfrak{a}, \mathfrak{b}) , \quad \text{for all } \mathfrak{a}, \mathfrak{b} \in \mathcal{C}^{\mathrm{obj}}.$$

The *inclusion functor* $I : \mathcal{C} \to \mathcal{D}$ from a subcategory $\mathcal{C}$ to $\mathcal{D}$ maps each object and morphism of $\mathcal{C}$ to itself.

(b) The *product* of $\mathcal{C}$ and $\mathcal{D}$ is the category $\mathcal{C} \times \mathcal{D}$ where

$$(\mathcal{C} \times \mathcal{D})^{\text{obj}} := \mathcal{C}^{\text{obj}} \times \mathcal{D}^{\text{obj}},$$

and $\quad (\mathcal{C} \times \mathcal{D})(\langle \mathfrak{a}_0, \mathfrak{a}_1 \rangle, \langle \mathfrak{b}_0, \mathfrak{b}_1 \rangle) := \mathcal{C}(\mathfrak{a}_0, \mathfrak{b}_0) \times \mathcal{D}(\mathfrak{a}_1, \mathfrak{b}_1)$,

for objects $\langle \mathfrak{a}_0, \mathfrak{a}_1 \rangle, \langle \mathfrak{b}_0, \mathfrak{b}_1 \rangle \in \mathcal{C} \times \mathcal{D}$. The composition of morphisms is defined componentwise:

$$\langle f_0, f_1 \rangle \circ \langle g_0, g_1 \rangle := \langle f_0 \circ g_0,\ f_1 \circ g_1 \rangle .$$

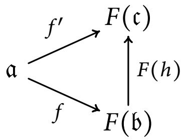
With each product $\mathcal{C} \times \mathcal{D}$ are associated two *projection functors*

$$P_0 : \mathcal{C} \times \mathcal{D} \to \mathcal{C} \quad \text{and} \quad P_1 : \mathcal{C} \times \mathcal{D} \to \mathcal{D},$$

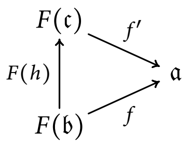
where P_i maps an object $\langle \mathfrak{a}_0, \mathfrak{a}_1 \rangle$ to $\mathfrak{a}_i$ and a morphism $\langle f_0, f_1 \rangle$ to f_i.

(c) Given an object $\mathfrak{a} \in \mathcal{D}$ and a functor $F : \mathcal{C} \to \mathcal{D}$, we define the *comma category* $(\mathfrak{a} \downarrow F)$ whose objects are all pairs $\langle f, \mathfrak{b} \rangle$ consisting of an object $\mathfrak{b} \in \mathcal{C}$ and a morphism $f : \mathfrak{a} \to F(\mathfrak{b})$ of $\mathcal{D}$. A morphism $h : \langle f, \mathfrak{b} \rangle \to \langle f', \mathfrak{c} \rangle$ from $f : \mathfrak{a} \to F(\mathfrak{b})$ to $f' : \mathfrak{a} \to F(\mathfrak{c})$ is a morphism $h : \mathfrak{b} \to \mathfrak{c}$ of $\mathcal{C}$ such that

$$f' = F(h) \circ f .$$



Similarly, we can define the *comma category* $(F \downarrow \mathfrak{a})$ consisting of all pairs $\langle \mathfrak{b}, f \rangle$ consisting of an object $\mathfrak{b} \in \mathcal{C}$ and a morphism $f : F(\mathfrak{b}) \to \mathfrak{a}$ of $\mathcal{D}$, where a morphism $h : \langle \mathfrak{b}, f \rangle \to \langle \mathfrak{c}, f' \rangle$ consists of a morphism $h \in \mathcal{C}^{\text{mor}}$ such that $f = f' \circ F(h)$.



More generally, given two functors $F : \mathcal{I} \to \mathcal{D}$ and $G : \mathcal{J} \to \mathcal{D}$, we define the *comma category* $(F \downarrow G)$ of all triples $\langle \mathfrak{a}, f, \mathfrak{b} \rangle$ where $\mathfrak{a} \in \mathcal{I}$, $\mathfrak{b} \in \mathcal{J}$, and $f : F(\mathfrak{a}) \to G(\mathfrak{b})$. A morphism $\varphi : \langle \mathfrak{a}, f, \mathfrak{b} \rangle \to \langle \mathfrak{a}', f', \mathfrak{b}' \rangle$ from $f : F(\mathfrak{a}) \to G(\mathfrak{b})$ to $f' : F(\mathfrak{a}') \to G(\mathfrak{b}')$ consists of a pair $\varphi = \langle g, h \rangle$ of morphisms $g : \mathfrak{a} \to \mathfrak{a}'$ and $h : \mathfrak{b} \to \mathfrak{b}'$ such that

$$F(h) \circ f = f' \circ F(g) .$$

$$\begin{array}{ccc} F(\mathfrak{a}') & \xrightarrow{\;f'\;} & F(\mathfrak{b}') \\ {\scriptstyle F(g)} \uparrow & & \uparrow {\scriptstyle F(h)} \\ F(\mathfrak{a}) & \xrightarrow[\;f\;]{} & F(\mathfrak{b}) \end{array}$$

To simplify notation, we will usually just write $f : F(\mathfrak{a}) \to G(\mathfrak{b})$ for an object $\langle \mathfrak{a}, f, \mathfrak{b} \rangle$.

Example. Consider the identity functor $I : \mathfrak{Emb}(\Sigma) \to \mathfrak{Emb}(\Sigma)$. For $\mathfrak{A} \in \mathfrak{Emb}(\Sigma)$, the comma category $(I \downarrow \mathfrak{A})$ consists of all embeddings $\mathfrak{C} \to \mathfrak{A}$ of a substructure into $\mathfrak{A}$.

Remark. The general definition of a comma category $(F \downarrow G)$ covers the special cases $(\mathfrak{a} \downarrow F)$ and $(F \downarrow \mathfrak{a})$ by using the functor $G : [1] \to \mathcal{D}$ from the single object category $[1]$ to $\mathcal{D}$ which maps the unique object of $[1]$ to $\mathfrak{a}$.

Exercise 3.2. Prove that the product $\mathcal{C} \times \mathcal{D}$ of two categories is universal in the sense that, given any category $\mathcal{E}$ and two functors $F : \mathcal{E} \to \mathcal{C}$ and $G : \mathcal{E} \to \mathcal{D}$, there exists a functor $H : \mathcal{E} \to \mathcal{C} \times \mathcal{D}$ such that $F = P_0 \circ H$ and $G = P_1 \circ H$. (For sets we have proved a corresponding statement in Lemma A2.2.2).

To compare two functors we define the notion of a 'homomorphism between functors'. In particular, we want to define when two functors are 'basically the same'.

Definition 3.12. (a) Let F and G be two functors from $\mathcal{C}$ to $\mathcal{D}$. A *natural transformation* from F to G is a family $\eta = (\eta_{\mathfrak{a}})_{\mathfrak{a} \in \mathcal{C}^{\text{obj}}}$ of morphisms

$$\eta_{\mathfrak{a}} \in \mathcal{D}(F(\mathfrak{a}), G(\mathfrak{a}))\,, \qquad \text{for } \mathfrak{a} \in \mathcal{C}^{\text{obj}}\,,$$

such that, for every morphism $f : \mathfrak{a} \to \mathfrak{b}$ of $\mathcal{C}$, the diagram

$$\begin{array}{ccc} F(\mathfrak{a}) & \xrightarrow{F(f)} & F(\mathfrak{b}) \\ {\scriptstyle \eta_{\mathfrak{a}}}\downarrow & & \downarrow{\scriptstyle \eta_{\mathfrak{b}}} \\ G(\mathfrak{a}) & \xrightarrow[G(f)]{} & G(\mathfrak{b}) \end{array}$$

commutes. If each $\eta_{\mathfrak{a}}$ is an isomorphism we call the transformation a *natural isomorphism*. In this case we write $\eta : F \cong G$.

(b) A functor $F : \mathcal{C} \to \mathcal{D}$ is an *equivalence* between the categories $\mathcal{C}$ and $\mathcal{D}$ if there exist a functor $G : \mathcal{D} \to \mathcal{C}$ and natural isomorphisms $\eta : \mathrm{id}_{\mathcal{D}} \cong F \circ G$ and $\rho : G \circ F \cong \mathrm{id}_{\mathcal{C}}$, where id denotes the identity functor. In this case we call $\mathcal{C}$ and $\mathcal{D}$ *equivalent*. If $\mathcal{C}$ is equivalent to $\mathcal{D}^{\text{op}}$, we say that the categories $\mathcal{C}$ and $\mathcal{D}$ are *dual*.

Example. Let V be a finite dimensional K-vector space. The *dual* $V^\vee$ of V consists of all linear maps $V \to K$. $V^\vee$ is again a K-vector space and we have $(V^\vee)^\vee \cong V$. For every linear map $h : V \to W$, we obtain a linear map $h^\vee : W^\vee \to V^\vee$ by setting $h^\vee(\lambda) := \lambda \circ h$. Consequently, the mapping $F : V \mapsto V^\vee$ forms a contravariant functor from the category of all finite dimensional K-vector spaces into itself. Furthermore, the family of isomorphisms $\pi_V : (V^\vee)^\vee \to V$ forms a natural isomorphism between $F \circ F$ and the identity functor. Hence, we can say that 'up to isomorphism' $F = F^{-1}$.

Lemma 3.13. *An equivalence $F : \mathcal{C} \to \mathcal{D}$ preserves and reflects monomorphisms, epimorphisms, initial objects, and terminal objects.*

Exercise 3.3. Prove the preceding lemma.

The next theorem provides an alternative characterisation of equivalences between categories. It also contains an important relationship between the two natural isomorphisms η and ρ associated with an equivalence.

Theorem 3.14. *Let $F : \mathcal{C} \to \mathcal{D}$ be a functor. The following statements are equivalent:*

(1) *F is an equivalence.*

(2) *F is full and faithful, and every object of $\mathcal{D}$ is isomorphic to one in* $\operatorname{rng} F^{\mathrm{obj}}$.

(3) *There exist a functor $G : \mathcal{D} \to \mathcal{C}$ and two natural isomorphisms $\eta : \mathrm{id}_{\mathcal{D}} \cong F \circ G$ and $\rho : G \circ F \cong \mathrm{id}_{\mathcal{C}}$ satisfying*

$$F(\rho_{\mathfrak{a}}) = \eta^{-1}_{F(\mathfrak{a})} \quad \textit{and} \quad G(\eta_{\mathfrak{b}}) = \rho^{-1}_{G(\mathfrak{b})} \,.$$

Proof. (3) ⇒ (1) is trivial.

(1) ⇒ (2) Suppose that there exist a functor $G : \mathcal{D} \to \mathcal{C}$ and two natural isomorphisms $\eta : \mathrm{id}_{\mathcal{D}} \cong F \circ G$ and $\rho : G \circ F \cong \mathrm{id}_{\mathcal{C}}$ with the above properties. For every object $\mathfrak{b} \in \mathcal{D}$, we have the isomorphism

$$\eta_{\mathfrak{b}} : \mathfrak{b} \cong F(G(\mathfrak{b})) \in \operatorname{rng} F^{\mathrm{obj}}.$$

To show that F is faithful, let $f, f' : \mathfrak{a} \to \mathfrak{b}$ be morphisms with $F(f) = F(f')$. Then

$$\begin{aligned} f = f \circ \rho_{\mathfrak{a}} \circ \rho_{\mathfrak{a}}^{-1} &= \rho_{\mathfrak{b}} \circ G(F(f)) \circ \rho_{\mathfrak{a}}^{-1} \\ &= \rho_{\mathfrak{b}} \circ G(F(f')) \circ \rho_{\mathfrak{a}}^{-1} = f' \circ \rho_{\mathfrak{a}} \circ \rho_{\mathfrak{a}}^{-1} = f' \,. \end{aligned}$$

In the same way, it follows that G is faithful.

It remains to show that F is full. Let $f : F(\mathfrak{a}) \to F(\mathfrak{b})$ be a morphism of $\mathcal{D}$. Setting

$$g := \rho_{\mathfrak{b}} \circ G(f) \circ \rho_{\mathfrak{a}}^{-1} \,,$$

we have

$$\rho_{\mathfrak{b}} \circ G(f) \circ \rho_{\mathfrak{a}}^{-1} = g = g \circ \rho_{\mathfrak{a}} \circ \rho_{\mathfrak{a}}^{-1} = \rho_{\mathfrak{b}} \circ G(F(g)) \circ \rho_{\mathfrak{a}}^{-1} .$$

As $\rho_{\mathfrak{b}}$ and $\rho_{\mathfrak{a}}$ are isomorphisms, this implies that $G(f) = G(F(g))$. We have shown above that G is faithful. Consequently, it follows that $f = F(g) \in \operatorname{rng} F^{\mathrm{mor}}$.

(2) ⇒ (3) By (2), we can choose, for every $\mathfrak{b} \in \mathcal{D}^{\mathrm{obj}}$, some object $G(\mathfrak{b}) \in \mathcal{C}$ and an isomorphism $\eta_{\mathfrak{b}} : \mathfrak{b} \cong F(G(\mathfrak{b}))$. This defines the object part of the functor G.

It remains to define the morphism part G^{mor}. Since F is full and faithful, it induces bijections

$$\psi_{\mathfrak{a},\mathfrak{b}} := F \upharpoonright \mathcal{C}(\mathfrak{a},\mathfrak{b}) : \mathcal{C}(\mathfrak{a},\mathfrak{b}) \to \mathcal{D}(F(\mathfrak{a}), F(\mathfrak{b})) , \quad \text{for } \mathfrak{a},\mathfrak{b} \in \mathcal{C} .$$

For a morphism $f : \mathfrak{a} \to \mathfrak{b}$ of $\mathcal{D}$, we set

$$G(f) := \psi^{-1}_{G(\mathfrak{a}),G(\mathfrak{b})}(\eta_{\mathfrak{b}} \circ f \circ \eta_{\mathfrak{a}}^{-1}) .$$

Since $F(g \circ f) = F(g) \circ F(f)$, we have

$$\psi^{-1}_{\mathfrak{a},\mathfrak{c}}(g \circ f) = \psi^{-1}_{\mathfrak{b},\mathfrak{c}}(g) \circ \psi^{-1}_{\mathfrak{a},\mathfrak{b}}(f) ,$$

for $f : F(\mathfrak{a}) \to F(\mathfrak{b})$ and $g : F(\mathfrak{b}) \to F(\mathfrak{c})$. Consequently,

$$\begin{aligned}
G(g \circ f) &= \psi^{-1}_{G(\mathfrak{a}),G(\mathfrak{c})}(\eta_{\mathfrak{c}} \circ g \circ f \circ \eta_{\mathfrak{a}}^{-1}) \\
&= \psi^{-1}_{G(\mathfrak{a}),G(\mathfrak{c})}(\eta_{\mathfrak{c}} \circ g \circ \eta_{\mathfrak{b}}^{-1} \circ \eta_{\mathfrak{b}} \circ f \circ \eta_{\mathfrak{a}}^{-1}) \\
&= \psi^{-1}_{G(\mathfrak{b}),G(\mathfrak{c})}(\eta_{\mathfrak{c}} \circ g \circ \eta_{\mathfrak{b}}^{-1}) \circ \psi^{-1}_{G(\mathfrak{a}),G(\mathfrak{b})}(\eta_{\mathfrak{b}} \circ f \circ \eta_{\mathfrak{a}}^{-1}) \\
&= G(g) \circ G(f) ,
\end{aligned}$$

and G is a functor.

We have choosen each morphism $\eta_{\mathfrak{a}}$ to be an isomorphism. Hence, to show that η is a natural isomorphism, it is sufficient to prove that

$$F(G(f)) \circ \eta_{\mathfrak{a}} = \eta_{\mathfrak{b}} \circ f , \quad \text{for all } f : \mathfrak{a} \to \mathfrak{b} \text{ in } \mathcal{D}^{\mathrm{mor}}.$$

For a morphism $f : \mathfrak{a} \to \mathfrak{b}$, we have

$$\begin{aligned} F(G(f)) \circ \eta_\mathfrak{a} &= F\big(\psi^{-1}_{G(\mathfrak{a}),G(\mathfrak{b})}(\eta_\mathfrak{b} \circ f \circ \eta_\mathfrak{a}^{-1})\big) \circ \eta_\mathfrak{a} \\ &= \eta_\mathfrak{b} \circ f \circ \eta_\mathfrak{a}^{-1} \circ \eta_\mathfrak{a} \\ &= \eta_\mathfrak{b} \circ f\,, \end{aligned}$$

as desired.

To conclude the proof, we define

$$\rho_\mathfrak{a} := \psi^{-1}_{G(F(\mathfrak{a})),\mathfrak{a}}(\eta^{-1}_{F(\mathfrak{a})})\,, \quad \text{for } \mathfrak{a} \in \mathcal{C}\,.$$

Then $\rho := (\rho_\mathfrak{a})_{\mathfrak{a}\in\mathcal{C}}$ is a natural transformation since, for $f : \mathfrak{a} \to \mathfrak{b}$ in $\mathcal{C}$,

$$\begin{aligned} &\rho_\mathfrak{b} \circ G(F(f)) \\ &= \psi^{-1}_{G(F(\mathfrak{b})),\mathfrak{b}}(\eta^{-1}_{F(\mathfrak{b})}) \circ \psi^{-1}_{G(F(\mathfrak{a})),G(F(\mathfrak{b}))}(\eta_{F(\mathfrak{b})} \circ F(f) \circ \eta^{-1}_{F(\mathfrak{a})}) \\ &= \psi^{-1}_{G(F(\mathfrak{a})),\mathfrak{b}}(\eta^{-1}_{F(\mathfrak{b})} \circ \eta_{F(\mathfrak{b})} \circ F(f) \circ \eta^{-1}_{F(\mathfrak{a})}) \\ &= \psi^{-1}_{\mathfrak{a},\mathfrak{b}}(F(f)) \circ \psi^{-1}_{G(F(\mathfrak{a})),\mathfrak{a}}(\eta^{-1}_{F(\mathfrak{a})}) \\ &= f \circ \rho_\mathfrak{a}^{-1}\,. \end{aligned}$$

Furthermore, each component $\rho_\mathfrak{a}$ is an ismorphism since $F(\rho_\mathfrak{a}) = \eta^{-1}_{F(\mathfrak{a})}$ is an isomorphism and the functor F is full and faithful. Finally, note that

$$\begin{aligned} G(\eta_\mathfrak{b}) &= \psi^{-1}_{G(\mathfrak{b}),G(F(G(\mathfrak{b})))}(\eta_{F(G(\mathfrak{b}))} \circ \eta_\mathfrak{b} \circ \eta_\mathfrak{b}^{-1}) \\ &= \psi^{-1}_{G(\mathfrak{b}),G(F(G(\mathfrak{b})))}(\eta_{F(G(\mathfrak{b}))}) \\ &= \big(\psi^{-1}_{G(F(G(\mathfrak{b}))),G(\mathfrak{b})}(\eta^{-1}_{F(G(\mathfrak{b}))})\big)^{-1} = \rho^{-1}_{G(\mathfrak{b})}\,. \end{aligned}$$ □

4. Congruences and quotients

Sometimes we do not want to distinguish between certain elements of a structure. In these situations we can use congruences to obtain a more abstract view of the given structure.

Definition 4.1. Let $\mathfrak{A}$ be a Σ-structure.

(a) An equivalence relation $\sim$ on the universe A is a *weak congruence relation* if it satisfies the following properties:

- If $a \sim b$ then there is some sort s such that $a, b \in A_s$.
- If $f \in \Sigma$ is an n-ary function and $a_0 \sim b_0, \dots, a_{n-1} \sim b_{n-1}$ then

$$f^{\mathfrak{A}}(a_0, \dots, a_{n-1}) \sim f^{\mathfrak{A}}(b_0, \dots, b_{n-1}) .$$

(b) A (*strong*) *congruence relation* is a weak congruence relation $\sim$ with the additional property that

- if $R \in \Sigma$ is an n-ary relation and $a_0 \sim b_0, \dots, a_{n-1} \sim b_{n-1}$ then

$$\langle a_0, \dots, a_{n-1} \rangle \in R^{\mathfrak{A}} \quad \text{iff} \quad \langle b_0, \dots, b_{n-1} \rangle \in R^{\mathfrak{A}}.$$

(c) We denote the set of all congruence relations of $\mathfrak{A}$ by $\mathrm{Cong}(\mathfrak{A})$, and we set

$$\mathfrak{Cong}(\mathfrak{A}) := \langle \mathrm{Cong}(\mathfrak{A}), \subseteq \rangle .$$

Similarly, $\mathrm{Cong}_{\mathrm{w}}(\mathfrak{A})$ is the set of all weak congruences and

$$\mathfrak{Cong}_{\mathrm{w}}(\mathfrak{A}) := \langle \mathrm{Cong}_{\mathrm{w}}(\mathfrak{A}), \subseteq \rangle$$

is the corresponding partial order.

Example. (a) If $\mathfrak{A} = \langle A, \leq \rangle$ is a linear order then $\mathrm{Cong}(\mathfrak{A}) = \{\mathrm{id}\}$ while $\mathrm{Cong}_{\mathrm{w}}(\mathfrak{A})$ contains all equivalence relations over A.

(b) Let $\mathfrak{V} = \langle V, +, (\lambda_a)_a \rangle$ be a vector space. If $\sim$ is a congruence of $\mathfrak{V}$ then $[0]_\sim$ forms a linear subspace of $\mathfrak{V}$. Conversely, if $\mathfrak{U} \subseteq \mathfrak{V}$ is a linear subspace then the relation

$$a \sim b \quad :\text{iff} \quad a - b \in U$$

is a congruence of $\mathfrak{V}$ with $[0]_\sim = U$. It follows that the map $\sim \mapsto [0]_\sim$ is an isomorphism between $\mathfrak{Cong}(\mathfrak{V})$ and the class of all linear subspaces of $\mathfrak{V}$ ordered by inclusion.

(c) Let $\mathfrak{Z} = \langle \mathbb{Z}, + \rangle$ and $\mathfrak{D} = \langle \mathbb{N}, \sqsubseteq \rangle$ where

$$x \sqsubseteq y \quad \text{:iff} \quad y \mid x$$

is the reverse divisibility order. We claim that $\mathfrak{Cong}(\mathfrak{Z}) \cong \mathfrak{D}$. For $k \in \mathbb{N}$, set

$$x \sim_k y \quad \text{:iff} \quad x - y = kz \text{ for some } z \in \mathbb{Z}\,.$$

We show that $\mathrm{Cong}(\mathfrak{Z}) = \{ \sim_k \mid k \in \mathbb{N} \}$. Since

$$\sim_k \subseteq \sim_m \quad \text{iff} \quad m \mid k$$

it then follows that the function $\sim_k \mapsto k$ is the desired isomorphism.

Clearly, every relation $\sim_k$ is a congruence of $\mathfrak{Z}$. Conversely, let $\approx$ be a congruence of $\mathfrak{Z}$. If $\approx \neq \sim_0$ then there are numbers $x < y$ with $x \approx y$. Since $-x \approx -x$ it follows that

$$0 = x + -x \approx y + -x > 0\,.$$

Let k be the minimal number such that $k > 0$ and $0 \approx k$. We claim that $\approx = \sim_k$. Since $0 \approx k$ we have $0 \approx kz$, for all $z \in \mathbb{Z}$. Hence, $\sim_k \subseteq \approx$. Conversely, if $x \approx y$ then we have seen that $|y - x| \approx 0$. Suppose that

$$|y - x| \equiv m \pmod{k}\,, \quad \text{for } 0 \leq m < k\,.$$

Since $0 \approx k$ it follows that $m \approx 0$. By choice of k, we have $m = 0$. Hence, $x \sim_k y$.

Before turning to quotients let us take a closer look at the structure of $\mathfrak{Cong}(\mathfrak{A})$.

Lemma 4.2. $\mathrm{Cong}(\mathfrak{A})$ *is an initial segment of* $\mathfrak{Cong}_{\mathrm{w}}(\mathfrak{A})$.

Proof. Let $\approx \in \mathrm{Cong}(\mathfrak{A})$ and $\sim \in \mathrm{Cong}_{\mathrm{w}}(\mathfrak{A})$ with $\sim \subseteq \approx$. Let R be an n-ary relation symbol of $\mathfrak{A}$. If $a_0 \sim b_0, \dots, a_{n-1} \sim b_{n-1}$ then $\sim \subseteq \approx$ implies $a_i \approx b_i$, for all i. Hence, we have

$$\bar{a} \in R^{\mathfrak{A}} \quad \text{iff} \quad \bar{b} \in R^{\mathfrak{A}}.$$

Consequently, $\sim \in \mathrm{Cong}(\mathfrak{A})$. ☐

Lemma 4.3. *Let $\mathfrak{A}$ be a Σ-structure and $X \subseteq \mathrm{Cong_w}(\mathfrak{A})$ nonempty. Set*

$$E_- := \bigcap X \quad \textit{and} \quad E_+ := \mathrm{TC}(\bigcup X)\,.$$

(a) *E_- and E_+ are weak congruence relations on A.*

(b) *If $X \subseteq \mathrm{Cong}(\mathfrak{A})$ then we have $E_-, E_+ \in \mathrm{Cong}(\mathfrak{A})$.*

Proof. We have already seen in Corollary A2.4.17 that E_- and E_+ are equivalence relations. It remains to prove that they are (weak) congruences.

Suppose that $\langle a_i, b_i \rangle \in E_-$, for $i < n$, and fix some $F \in X$. Let f be an n-ary function. Since $\langle a_i, b_i \rangle \in F$ it follows that

$$\langle f(\bar{a}), f(\bar{b}) \rangle \in F\,.$$

Hence, $\langle f(\bar{a}), f(\bar{b}) \rangle \in \bigcap X$.

For (b), we also have to consider n-ary relations R. Fix a congruence $F \in X \subseteq \mathrm{Cong}(\mathfrak{A})$. Then $\langle a_i, b_i \rangle \in F$ implies

$$\langle a_0, \dots, a_{n-1} \rangle \in R \quad \text{iff} \quad \langle b_0, \dots, b_{n-1} \rangle \in R\,.$$

The proof for E_+ is slightly more involved. Suppose that $\langle a_i, b_i \rangle \in E_+$, for $i < n$. For every $i < n$, there is a sequence $c^i_0, \dots, c^i_{l_i}$, with $l_i < \omega$, such that

$$c^i_0 = a_i\,, \quad c^i_{l_i} = b_i\,, \quad \text{and} \quad \langle c^i_j, c^i_{j+1} \rangle \in \bigcup X\,, \quad \text{for all } j < l_i\,.$$

Let f be an n-ary function. For every $i < n$ and all $j < l_i$, we have

$$\big\langle f(b_0, \dots, b_{i-1}, c^i_j, a_{i+1}, \dots, a_{n-1}), \\ f(b_0, \dots, b_{i-1}, c^i_{j+1}, a_{i+1}, \dots, a_{n-1}) \big\rangle \in \bigcup X\,.$$

This implies that

$$\big\langle f(b_0, \dots, b_{i-1}, a_i, a_{i+1}, \dots, a_{n-1}), \\ f(b_0, \dots, b_{i-1}, b_i, a_{i+1}, \dots, a_{n-1}) \big\rangle \in \mathrm{TC}(\bigcup X)\,,$$

and, by induction, it follows that

$$\begin{aligned}
&\big\langle f(\bar a), f(b_0, a_1, a_2, \dots, a_{n-1})\big\rangle \in E_+ ,\\
&\big\langle f(\bar a), f(b_0, b_1, a_2, \dots, a_{n-1})\big\rangle \in E_+ ,\\
&\qquad\qquad \dots\\
&\big\langle f(\bar a), f(b_0, \dots, b_{n-2}, a_{n-1})\big\rangle \in E_+ ,\\
&\big\langle f(\bar a), f(b_0, \dots, b_{n-2}, b_{n-1})\big\rangle \in E_+ .
\end{aligned}$$

Similarly, if R is an n-ary relation then we have, for all $i < n$ and $j < l_i$,

$$\begin{aligned}
&\langle b_0, \dots, b_{i-1}, c^i_j, a_{i+1}, \dots, a_{n-1}\rangle \in R\\
\text{iff}\quad &\langle b_0, \dots, b_{i-1}, c^i_{j+1}, a_{i+1}, \dots, a_{n-1}\rangle \in R ,
\end{aligned}$$

and it follows that

$$\begin{aligned}
&\langle b_0, \dots, b_{i-1}, a_i, a_{i+1}, \dots, a_{n-1}\rangle \in R\\
\text{iff}\quad &\langle b_0, \dots, b_{i-1}, b_i, a_{i+1}, \dots, a_{n-1}\rangle \in R .
\end{aligned}$$

As above we can conclude that $\bar a \in R$ iff $\bar b \in R$. ☐

Theorem 4.4. *Let $\mathfrak{A}$ be a structure. $\mathfrak{Cong}_w(\mathfrak{A})$ and $\mathfrak{Cong}(\mathfrak{A})$ form complete partial orders where, for every nonempty set X, we have*

$$\inf X = \bigcap X \quad \textit{and} \quad \sup X = \mathrm{TC}\big(\bigcup X\big) .$$

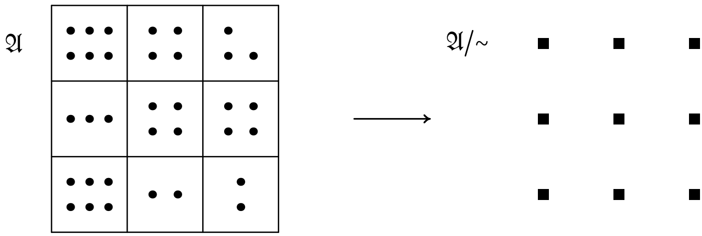
Proof. We have seen in Corollary A2.4.17 that the partial order of equivalence relations on A is complete. Consequently, the claim follows from Lemma 4.3 and Corollary A2.3.11. ☐

Every weak congruence defines an abstraction operation on structures.

Definition 4.5. Let $\mathfrak{A}$ be a Σ-structure and $\sim$ a weak congruence of $\mathfrak{A}$.

(a) The *quotient* $\mathfrak{A}/{\sim}$ of $\mathfrak{A}$ is the Σ-structure where the domain of sort s is $A_s/{\sim}$, for each n-ary relation symbol $R \in \Sigma$, we have the relation

$$R^{\mathfrak{A}/\sim} := \big\{ \langle [a_0]_\sim, \dots, [a_{n-1}]_\sim\rangle \bigm| \langle a_0, \dots, a_{n-1}\rangle \in R^{\mathfrak{A}} \big\} ,$$



and, for every n-ary function symbol $f \in \Sigma$, the function

$$f^{\mathfrak{A}/\sim}([a_0]_\sim, \dots, [a_{n-1}]_\sim) := [f^{\mathfrak{A}}(a_0, \dots, a_{n-1})]_\sim .$$

We also say that we obtain $\mathfrak{A}/\sim$ from $\mathfrak{A}$ by *factorisation by* $\sim$.

(b) The function $\pi : \mathfrak{A} \to \mathfrak{A}/\sim$ with $\pi(a) := [a]_\sim$ is called the *canonical projection.*

Remark. The structure $\mathfrak{A}/\sim$ is well-defined since, by definition, if we have $a_0 \sim b_0, \dots, a_{n-1} \sim b_{n-1}$ then

$$f^{\mathfrak{A}}(a_0, \dots, a_{n-1}) \sim f^{\mathfrak{A}}(b_0, \dots, b_{n-1}) ,$$

which implies that

$$[f^{\mathfrak{A}}(a_0, \dots, a_{n-1})]_\sim = [f^{\mathfrak{A}}(b_0, \dots, b_{n-1})]_\sim .$$

Example. $\mathfrak{On} = \langle \mathrm{Wo}, \leq \rangle / \cong$ and $\mathrm{ord} : \langle \mathrm{Wo}, \leq \rangle \to \mathfrak{On}$ is a homomorphism.

There is a strong connection between congruence relations and homomorphisms.

Lemma 4.6. *Let $\mathfrak{A}$ be a Σ-structure, $\sim$ a weak congruence on $\mathfrak{A}$, and $\pi : \mathfrak{A} \to \mathfrak{A}/\sim$ the canonical projection.*

(a) *π is a surjective semi-strict homomorphism with* $\ker \pi = \sim$.

(b) *If $\sim$ is a congruence then π is a surjective strict homomorphism.*

Proof. (a) π is surjective since

$$A/{\sim} = \{ [a]_{\sim} \mid a \in A \} = \{ \pi(a) \mid a \in A \} = \operatorname{rng} \pi .$$

It is a homomorphism since, for all n-ary functions symbols $f \in \Sigma$, we have

$$\begin{aligned} \pi f^{\mathfrak{A}}(a_0, \dots, a_{n-1}) &= [f^{\mathfrak{A}}(a_0, \dots, a_{n-1})]_{\sim} \\ &= f^{\mathfrak{A}/\sim}([a_0]_{\sim}, \dots, [a_{n-1}]_{\sim}) \\ &= f^{\mathfrak{A}/\sim}(\pi a_0, \dots, \pi a_{n-1}) , \end{aligned}$$

and, for each n-ary relation symbols $R \in \Sigma$,

$$\begin{aligned} \langle a_0, \dots, a_{n-1} \rangle \in R^{\mathfrak{A}} \quad &\Rightarrow \quad \langle [a_0]_{\sim}, \dots, [a_{n-1}]_{\sim} \rangle \in R^{\mathfrak{A}/\sim} \\ &\Rightarrow \quad \langle \pi a_0, \dots, \pi a_{n-1} \rangle \in R^{\mathfrak{A}/\sim} . \end{aligned}$$

To show that π is semi-strict let $\langle [a_0], \dots, [a_{n-1}] \rangle \in R^{\mathfrak{A}/\sim}$. By definition of $\mathfrak{A}/\sim$ there are elements $b_i \sim a_i$, $i < n$, with $\bar{b} \in R^{\mathfrak{A}}$. This implies that $\pi(\bar{b}) = \pi(\bar{a})$.

(b) We have already seen in (a) that π is a surjective homomorphism. It is strict since, for each n-ary relation symbols $R \in \Sigma$, we have

$$\begin{aligned} \langle a_0, \dots, a_{n-1} \rangle \in R^{\mathfrak{A}} \quad &\text{iff} \quad \langle [a_0]_{\sim}, \dots, [a_{n-1}]_{\sim} \rangle \in R^{\mathfrak{A}/\sim} \\ &\text{iff} \quad \langle \pi a_0, \dots, \pi a_{n-1} \rangle \in R^{\mathfrak{A}/\sim} . \end{aligned}$$

□

Lemma 4.7. *Let $h : \mathfrak{A} \to \mathfrak{B}$ be a function.*

(a) *If h is a homomorphism then* $\ker h$ *is a weak congruence of* $\mathfrak{A}$.

(b) *If h is a strict homomorphism then* $\ker h$ *is a congruence.*

Proof. (a) $\ker h$ is an equivalence relation since $=$ is reflexive, symmetric, and transitive. Furthermore, $h(a) = h(b)$ implies that a and b are of the same sort. Suppose that $\langle a_0, b_0 \rangle, \dots, \langle a_{n-1}, b_{n-1} \rangle \in \ker h$. If $f \in \Sigma$ is an n-ary function symbol then

$$h(f^{\mathfrak{A}}(\bar{a})) = f^{\mathfrak{B}}(h(\bar{a})) = f^{\mathfrak{B}}(h(\bar{b})) = h(f^{\mathfrak{A}}(\bar{b}))$$

implies that $\langle f^{\mathfrak{A}}(\bar{a}), f^{\mathfrak{A}}(\bar{b})\rangle \in \ker h$.

(b) If $R \in \Sigma$ is an n-ary relation symbol then we have

$$\bar{a} \in R^{\mathfrak{A}} \quad \text{iff} \quad h(\bar{a}) \in R^{\mathfrak{B}} \quad \text{iff} \quad h(\bar{b}) \in R^{\mathfrak{B}} \quad \text{iff} \quad \bar{b} \in R^{\mathfrak{A}}. \quad \square$$

Corollary 4.8. *Let $\mathfrak{A}$ be a Σ-structure and $\sim \subseteq A \times A$ a binary relation.*

(a) *$\sim$ is a weak congruence relation if and only if there exists a homomorphism $h : \mathfrak{A} \to \mathfrak{B}$ such that $\sim = \ker h$.*

(b) *$\sim$ is a congruence relation if and only if there exists a strict homomorphism $h : \mathfrak{A} \to \mathfrak{B}$ such that $\sim = \ker h$.*

(c) *Let $\mathfrak{B}$ be a Σ-structure. There exists a weak congruence $\sim$ such that $\mathfrak{B} \cong \mathfrak{A}/{\sim}$ if and only if $\mathfrak{B}$ is a homomorphic image of $\mathfrak{A}$.*

Proof. We prove all three claims simultaneously. The direction $(\Leftarrow)$ follows immediately from Lemma 4.7. For $(\Rightarrow)$ we can take $\mathfrak{B} := \mathfrak{A}/{\sim}$ and $h : a \mapsto [a]_\sim$, by Lemma 4.6. $\square$

Definition 4.9. Let $h : \mathfrak{A} \to \mathfrak{B}$ be a homomorphism and $\sim$ a weak congruence on $\mathfrak{B}$. We set

$$h^{-1}(\sim) := \{\, \langle a, b\rangle \in A \times A \mid h(a) \sim h(b) \,\}.$$

Lemma 4.10. *Let $h : \mathfrak{A} \to \mathfrak{B}$ be a homomorphism and $\sim$ a weak congruence on $\mathfrak{B}$.*

(a) *$h^{-1}(\sim)$ is a weak congruence on $\mathfrak{A}$.*

(b) *If h is strict and $\sim \in \mathrm{Cong}(\mathfrak{B})$ then $h^{-1}(\sim) \in \mathrm{Cong}(\mathfrak{A})$.*

Proof. If $\pi : \mathfrak{B} \to \mathfrak{B}/{\sim}$ is the canonical projection then we have

$$h^{-1}(\sim) = \ker(\pi \circ h).$$

Hence, the claims follow from Lemma 4.7. $\square$

Theorem 4.11. (a) *There exists a contravariant functor*

$$\mathcal{F} : \mathfrak{Hom}(\Sigma) \to \mathfrak{Hom}(\subseteq) : \mathfrak{A} \mapsto \mathfrak{Cong}_{\mathrm{w}}(\mathfrak{A})$$

with $\mathcal{F}(f) : \sim \mapsto f^{-1}(\sim)$, *for homomorphisms* $f : \mathfrak{A} \to \mathfrak{B}$.

(b) *There exists a contravariant functor*

$$\mathcal{G} : \mathfrak{Hom}_s(\Sigma) \to \mathfrak{Hom}(\subseteq) : \mathfrak{A} \mapsto \mathfrak{Cong}(\mathfrak{A})$$

with $\mathcal{G}(f) : \sim \mapsto f^{-1}(\sim)$, *for strict homomorphisms* $f : \mathfrak{A} \to \mathfrak{B}$.

Proof. (a) If $f : \mathfrak{A} \to \mathfrak{B}$ is a homomorphism and $\sim \subseteq \approx$ are weak congruences of $\mathfrak{B}$ then we have

$$\mathcal{F}(f)(\sim) = f^{-1}(\sim) \subseteq f^{-1}(\approx) = \mathcal{F}(f)(\approx)\,.$$

Hence, $\mathcal{F}(f)$ is a homomorphism. Furthermore, we have

$$\mathcal{F}(\mathrm{id}_{\mathfrak{A}})(\sim) = \sim\,, \qquad \text{for all } \sim \in \mathrm{Cong}_w(\mathfrak{A})\,,$$

which implies that $\mathcal{F}(\mathrm{id}_{\mathfrak{A}}) = \mathrm{id}_{\mathfrak{Cong}_w(\mathfrak{A})}$. Finally, if $f : \mathfrak{A} \to \mathfrak{B}$ and $g : \mathfrak{B} \to \mathfrak{C}$ are homomorphisms then we have

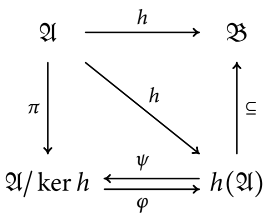
$$\begin{aligned}\mathcal{F}(g \circ f)(\sim) &= (g \circ f)^{-1}(\sim) \\ &= f^{-1}(g^{-1}(\sim)) = (\mathcal{F}(f) \circ \mathcal{F}(g))(\sim)\,.\end{aligned}$$

(b) is shown in exactly the same way replacing 'homomorphism' by 'strict homomorphism' and 'weak congruence' by 'congruence'. □

Theorem 4.12 (Homomorphism Theorem)**.** *For every semi-strict homomorphism* $h : \mathfrak{A} \to \mathfrak{B}$, *there exists a unique isomorphism*

$$\varphi : \mathfrak{A}/\ker h \to h(\mathfrak{A})$$

such that the following diagram commutes.



Proof. Let $\pi : \mathfrak{A} \to \mathfrak{A}/\ker h$ be the canonical projection. The existence of $\varphi : \mathfrak{A}/\ker h \to h(\mathfrak{A})$ follows immediately from Corollary 2.7 since both homomorphisms π and $h : \mathfrak{A} \to h(\mathfrak{A})$ are semi-strict and surjective and we have $\ker \pi = \ker h$. □

Corollary 4.13. *Every strict homomorphism $h : \mathfrak{A} \to \mathfrak{B}$ can be factorised as $h = \varphi \circ \pi$ where π is a surjective strict homomorphism and φ is an injective strict homomorphism.*

Example. Let $h : \mathfrak{G} \to \mathfrak{H}$ be a homomorphism between groups. Let $N := \ker h$ be the (normal subgroup corresponding to the) kernel of h. Then there exists a homomorphism $\varphi : \mathfrak{G}/N \to \mathfrak{H}$ such that $h = \varphi \circ \pi$ where $\pi : \mathfrak{G} \to \mathfrak{G}/N$ is the canonical projection.

Corollary 4.14. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be structures.*

(a) *There exists a surjective strict homomorphism $\mathfrak{A} \to \mathfrak{B}$ if and only if $\mathfrak{B} \cong \mathfrak{A}/\sim$, for some congruence relation $\sim$.*

(b) *There exists a strict homomorphism $\mathfrak{A} \to \mathfrak{B}$ if and only if there is a substructure $\mathfrak{B}_0 \subseteq \mathfrak{B}$ and a congruence relation $\sim$ on $\mathfrak{A}$ such that $\mathfrak{B}_0 \cong \mathfrak{A}/\sim$.*

We conclude this section with an investigation of the relationship between quotients $\mathfrak{A}/\sim$ and $\mathfrak{A}/\approx$ of the same structures.

Remark. For weak congruences $\sim \subseteq \approx$, we have $[a]_\sim \subseteq [a]_\approx$. Hence, every $\approx$-class is partitioned by $\sim$ into one or several $\sim$-classes.

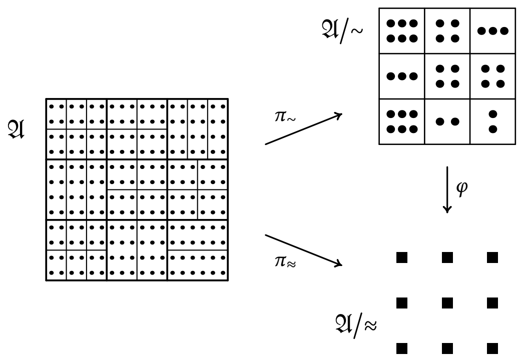
Definition 4.15. For weak congruences $\sim \subseteq \approx$ on $\mathfrak{A}$ we define

$$\approx/\sim := \{ \langle [a]_\sim, [b]_\sim \rangle \in A/\sim \times A/\sim \mid a \approx b \} .$$

Remark. If $\sim \subseteq \approx$ are weak congruences on $\mathfrak{A}$ then $\sim$ is also a weak congruence of $\langle \mathfrak{A}, \approx \rangle$ and we have

$$\langle \mathfrak{A}, \approx \rangle/\sim = \langle \mathfrak{A}/\sim, \approx/\sim \rangle .$$

Furthermore, if $\sim$ is a congruence on $\mathfrak{A}$ then $\sim$ is also a congruence of $\langle \mathfrak{A}, \approx \rangle$.



Lemma 4.16. *Let* $\sim\ \subseteq\ \approx$ *be weak congruences on* $\mathfrak{A}$ *and let* $\pi_\sim : \mathfrak{A} \to \mathfrak{A}/{\sim}$ *and* $\pi_\approx : \mathfrak{A} \to \mathfrak{A}/{\approx}$ *be the corresponding canonical projections.*

We have ${\approx}/{\sim} = \ker \varphi$ *where* $\varphi : \mathfrak{A}/{\sim} \to \mathfrak{A}/{\approx}$ *is the unique semi-strict homomorphism with* $\pi_\approx = \varphi \circ \pi_\sim$.

Proof. Since $\ker \pi_\sim = {\sim} \subseteq {\approx} = \ker \pi_\approx$ it follows by Lemmas 2.5 and 2.6 that there exists a unique semi-strict homomorphism $\varphi : \mathfrak{A}/{\sim} \to \mathfrak{A}/{\approx}$ with $\pi_\approx = \varphi \circ \pi_\sim$. For $[a]_\sim, [b]_\sim \in A/{\sim}$, we have

$$\begin{aligned} \varphi[a]_\sim = \varphi[b]_\sim \quad &\text{iff} \quad (\varphi \circ \pi_\sim)(a) = (\varphi \circ \pi_\sim)(b) \\ &\text{iff} \quad \pi_\approx(a) = \pi_\approx(b) \\ &\text{iff} \quad a \approx b \\ &\text{iff} \quad [a]_\sim \mathrel{{\approx}/{\sim}} [b]_\sim \,. \end{aligned}$$

□

Corollary 4.17. *Let* $\sim\ \subseteq\ \approx$ *be weak congruences on* $\mathfrak{A}$*.*

(a) ${\approx}/{\sim}$ *is a weak congruence on* $\mathfrak{A}/{\sim}$*.*

(b) *If* $\approx$ *is a congruence then so is* ${\approx}/{\sim}$*.*

Proof. (a) follows immediately from Lemma 4.16. For (b) note that, if $\approx$ is a congruence then $\pi_\approx$ is strict and it follows by Lemma 2.6 that φ is a strict homomorphism. □

Theorem 4.18. *Let $\sim\ \subseteq\ \approx$ be weak congruences on $\mathfrak{A}$. There exists an isomorphism*

$$(\mathfrak{A}/{\sim})\big/({\approx}/{\sim}) \cong \mathfrak{A}/{\approx}\,.$$

Proof. According to Lemma 4.16 there exists a semi-strict homomorphism $\varphi : \mathfrak{A}/{\sim} \to \mathfrak{A}/{\approx}$ with $\ker \varphi = {\approx}/{\sim}$. By the Homomorphism Theorem, it follows that there exists an isomorphism

$$\psi : (\mathfrak{A}/{\sim})\big/({\approx}/{\sim}) \to \mathfrak{A}/{\approx}\,. \qquad \square$$

Example. Let $\mathfrak{N} \subseteq \mathfrak{U} \subseteq \mathfrak{G}$ be normal subgroups of $\mathfrak{G}$. Then $\mathfrak{N}$ is also a normal subgroup of $\mathfrak{U}$ and we have

$$\mathfrak{G}/U \cong (\mathfrak{G}/N)\big/(U/N)\,.$$

Theorem 4.19. *Let $\mathfrak{A}$ be a structure and $\sim\ \in$ Cong$(\mathfrak{A})$. The function*

$$h : {\Uparrow}{\sim} \to \mathrm{Cong}(\mathfrak{A}/{\sim}) \quad \textit{with} \quad h(\approx) := {\approx}/{\sim}$$

defines an isomorphism between Cong$(\mathfrak{A}/{\sim})$ *and the final segment* ${\Uparrow}{\sim}$ *of* Cong$(\mathfrak{A})$.

Proof. Let $\rho, \sigma \in {\Uparrow}{\sim}$. It follows immediately from the definition that we have

$$\rho/{\sim} \subseteq \sigma/{\sim} \quad \text{iff} \quad \rho \subseteq \sigma\,.$$

Therefore, h is a strict homomorphism.

It remains to show that it is bijective. Suppose that $\rho \neq \sigma$. By symmetry, we may assume that there is some pair $\langle a, b\rangle \in \rho \setminus \sigma$. It follows that

$$\langle [a]_\sim, [b]_\sim\rangle \in \rho/{\sim} = h(\rho) \quad \text{and} \quad \langle [a]_\sim, [b]_\sim\rangle \notin \sigma/{\sim} = h(\sigma)\,.$$

Hence, we have $h(\rho) \neq h(\sigma)$ and h is injective. For surjectivity, let $\rho \in \mathrm{Cong}(\mathfrak{A}/{\sim})$ and define

$$\sigma := \big\{ \langle a, b\rangle \in A \times A \;\big|\; \langle [a]_\sim, [b]_\sim\rangle \in \rho \big\}\,.$$

Then we have $h(\sigma) = \rho$. $\square$

B2. *Trees and lattices*

1. *Trees*

Recall that, for an ordinal α, we denote by $A^{<\alpha}$ the set of all sequences $f : \beta \to A$ with $\beta < \alpha$. To simplify notation we will write finite sequences $\bar{a} = \langle a_0, \dots, a_n \rangle$ without braces and commas as $\bar{a} = a_0 \dots a_n$. We can equip $A^{<\alpha}$ with the following operations.

Definition 1.1. Let $x, y \in A^{<\alpha}$.

(a) The *length* of x is the ordinal $|x| := \operatorname{dom} x$.

(b) The *concatenation* $x \cdot y$ of x and y is the sequence $z : |x| + |y| \to A$ with

$$z_\beta := \begin{cases} x_\beta & \text{if } \beta < |x|\,, \\ y_\gamma & \text{if } \beta = |x| + \gamma\,. \end{cases}$$

Usually, we omit the dot and simply write xy instead of $x \cdot y$. For sets $X, Y \subseteq A^{<\alpha}$, we introduce the usual abbreviations

$$XY := \{ xy \mid x \in X, y \in Y \} \qquad \text{and} \qquad xY := \{ xy \mid y \in Y \}\,.$$

(c) The *prefix order* $\preceq$ on $A^{<\alpha}$ is defined by

$$x \preceq y \quad :\text{iff} \quad |x| \leq |y| \text{ and } y \restriction |x| = x\,.$$

If $x \preceq y$ then x is called a *prefix* of y.

(d) If we are given a linear order $\sqsubseteq$ on A then we can define the *lexicographic order* $\leq_{\text{lex}}$ on $A^{<\alpha}$ by

$$x \leq_{\text{lex}} y \quad :\text{iff} \quad x \preceq y \text{ or there are } z \in A^{<\alpha} \text{ and } a \sqsubset b \in A \text{ such that } za \preceq x \text{ and } zb \preceq y\,.$$

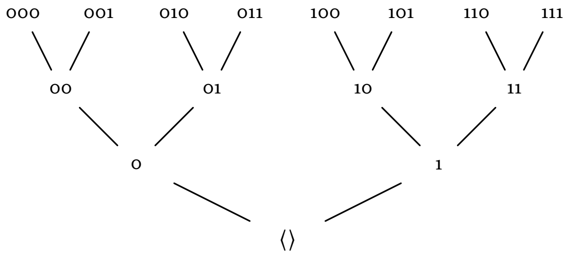


Figure 1.. $\langle 2^{<4}, \leq \rangle$

Example. (a) If $x = a_0 \dots a_{m-1}$ and $y = b_0 \dots b_{n-1}$ then

$$xy = a_0 \dots a_{m-1} b_0 \dots b_{n-1} .$$

In particular, $x \leq xy$.

(b) We have $x \cdot \langle\rangle = x = \langle\rangle \cdot x$, for all $x \in A^{<\alpha}$.

(c) The prefix order $\langle 2^{<4}, \leq \rangle$ is depicted in Figure 1, while the lexicographic ordering $\langle 2^{<4}, \leq_{\mathrm{lex}} \rangle$ is

$$\begin{aligned} &\langle\rangle < 0 < 00 < 000 < 001 < 01 < 010 < 011 \\ &\quad < 1 < 10 < 100 < 101 < 11 < 110 < 111 . \end{aligned}$$

This order corresponds to a so-called 'pre-order' or 'depth-first' traversal of the tree $\langle 2^{<4}, \leq \rangle$.

Exercise 1.1. Prove that $x \leq y$ iff there exists some z such that $y = xz$.

Note that, if $x, y \in A^{<\alpha}$ then $xy \in A^{<\alpha 2}$, but it might be the case that $xy \notin A^{<\alpha}$. Since $\operatorname{dom} xy = \operatorname{dom} x + \operatorname{dom} y$ we can use Lemma A3.4.25 to obtain a characterisation of all ordinals α such that $A^{<\alpha}$ is closed under concatenation.

Lemma 1.2. *Let $\alpha \in \mathrm{On}$. The set $A^{<\alpha}$ is closed under concatenation if and only if $\alpha = 0$ or $\alpha = \omega^{(\eta)}$, for some η.*

Remark. It follows that, for every α, the structure $\langle A^{<\omega^{(\alpha)}}, \cdot, \langle\rangle \rangle$ forms a monoid.

Trees play a prominent role in mathematics and computer science. Firstly, they have many pleasant algebraic and algorithmic properties, and secondly, many processes and structures can be modelled as a tree. For instance, consider an inductive fixed-point iteration that, starting with some basic elements, combines them in every step to form new elements. Every element is built up from one or several other elements that, in turn, consist of even more primitive elements, and so on until a basic element is reached. To model such hierarchical dependencies we will frequently use families $(a_v)_{v \in T}$ indexed by a tree T.

Definition 1.3. (a) A *tree* is a partial order $\mathfrak{T} = \langle T, \leq \rangle$ such that

- the set $\downarrow v$ is well-ordered, for every $v \in T$, and
- each pair $u, v \in T$ has a greatest lower bound $u \sqcap v := \inf \{u, v\}$.

(b) The elements of a tree are usually called *nodes* or *vertices.* A maximal element of a tree is called a *leaf,* all other elements of T are *inner vertices,* and the least element is the *root.*

(c) A vertex v is a *successor* of the vertex u if $u < v$ and there is no vertex w with $u < w < v$.

(d) A chain $C \subseteq T$ is a *path* if $u, v \in C$ implies that $w \in C$, for all $u \leq w \leq v$. A maximal path is called a *branch.*

Remark. (a) Note that every tree is a well-founded partial order.

(b) By convention, we will usually depict trees upside down with the root at the top.

The partial order $\langle 2^{<4}, \preceq \rangle$ in Figure 1 is a tree. In fact, the prefix order $\preceq$ always forms a tree and we will see below that every tree can be obtained in this way.

Lemma 1.4. *$\langle A^{<\alpha}, \preceq \rangle$ is a tree, for all A and α.*

The only thing preventing a tree from being a complete partial order is the lack of a greatest element.

Lemma 1.5. *Let $\mathfrak{T} = \langle T, \leq \rangle$ be a tree. If $X \subseteq T$ is nonempty then there are elements $a, b \in X$ with* $\inf X = a \sqcap b$. *In particular, X has an infimum.*

Proof. Fix some element $a \in X$. The set

$$Y := \{ a \sqcap x \mid x \in X \}$$

is a nonempty subset of $\Downarrow a$. Hence, it has a least element $c \in Y$. This element is a lower bound of X since we have

$$c \leq a \sqcap x \leq x , \qquad \text{for every } x \in X .$$

Fix some element $b \in X$ with $c = a \sqcap b$. If d is another lower bound of X then $d \leq a$ and $d \leq b$ implies $d \leq a \sqcap b = c$. Consequently, we have $c = a \sqcap b = \inf X$. □

Definition 1.6. Let $\mathfrak{T} = \langle T, \leq \rangle$ be a tree and $v \in T$ a vertex.

(a) The *subtree* of $\mathfrak{T}$ *rooted* at v is the substructure $\mathfrak{T}_v := \mathfrak{T}|_{\Uparrow v}$ induced by $\Uparrow v$.

(b) The *level* of a vertex v is the ordinal

$$|v| := \operatorname{ord} \langle \downarrow v, \leq \rangle .$$

The *height* of $\mathfrak{T}$ is the least ordinal greater than all levels

$$\sup \{ |v| + 1 \mid v \in T \} .$$

Example. Let $\mathfrak{T} = \langle A^{<\alpha}, \leq \rangle$. The level of $v \in A^{<\alpha}$ is the length of v. (That is the reason why we denote both by $|v|$.) It follows that the height of $\mathfrak{T}$ is α.

Lemma 1.7. *For every tree $\mathfrak{T} = \langle T, \leq \rangle$ of height α, there exists an initial segment $X \subseteq |T|^{<\alpha}$ such that $\mathfrak{T} \cong \langle X, \preceq \rangle$.*

Proof. For $\beta \in \mathrm{On}$, define $T_\beta := \{ v \in T \mid |v| < \beta \}$. Let α be the minimal ordinal such that $T_\alpha = T$ and set $\kappa := |T|$. To prove the claim it is sufficient to define an embedding $h : T \to \kappa^{<\alpha}$ such that $X := \operatorname{rng} h$

forms an initial segment. By induction on β, we construct an increasing sequence $h_1 \subseteq h_2 \subseteq \dots$ of embeddings $h_\beta : T_\beta \to \kappa^{<\beta}$. The desired function $h : T \to \kappa^{<\alpha}$ will be obtained as the limit $h := \bigcup_{\beta<\alpha} h_\beta$.

Let v be the root of T. Since v is the only vertex of length 0 we can set

$$h_1 : \{v\} \to \{\langle\rangle\} : v \mapsto \langle\rangle\,.$$

For the inductive step, suppose that h_γ is already defined for all $\gamma < \beta$. If β is a limit ordinal then we can set $h_\beta := \bigcup_{\gamma<\beta} h_\gamma$. Therefore, suppose that $\beta = \gamma + 1$ is a successor. For every vertex $v \in T$ of length $|v| < \gamma$, we set $h_\beta(v) := h_\gamma(v)$. It remains to consider the case that $|v| = \gamma$. First, suppose that $\gamma = \eta + 1$ is a successor. For each vertex $u \in T$ of length $|u| = \eta$, we fix an injective function $g_u : S_u \to \kappa$ from the set S_u of successors of u into κ. If $|v| = \gamma$ then $v \in S_u$, for some u, and we can set

$$h_\beta(v) := h_\gamma(u) \cdot \langle g_u(v)\rangle\,.$$

Finally, suppose that γ is a limit ordinal. We set $h_\beta(v) := x$ where $x : \gamma \to \kappa^{<\gamma+1}$ is the sequence with

$$x_\eta := h_\gamma(u)\,, \quad \text{for the vertex } u \leq v \text{ with } |u| = \eta\,. \qquad \square$$

We conclude this section with an investigation of the connection between trees and fixed-point inductions. First, we characterise those trees that contain an infinite path. Then we show that those without can be generated bottom-up in a recursive way.

Definition 1.8. The *branching degree* of a tree $\mathfrak{T}$ is the minimal cardinal κ such that there exists an embedding of $\mathfrak{T}$ into $\kappa^{<\alpha}$, for some ordinal α. We say that $\mathfrak{T}$ is *finitely branching* if every vertex $v \in T$ has only finitely many successors.

Example. The branching degree of $\langle A^{<\alpha}, \leq\rangle$ is $|A|$.

Remark. (a) Note that there are finitely branching trees of branching degree $\aleph_0$. For instance, the tree $\langle T, \leq\rangle$ with

$$T := \{\, \bar{a} \in \aleph_0^{<\omega} \mid a_n \leq n \text{ for } n < \omega \,\}\,,$$

is finitely branching. Every vertex $\bar{a}$ of length $|\bar{a}| = n$ has $n + 1$ successors.

(b) The branching degree of a tree $\mathfrak{T}$ is at most $|T|$, by the above lemma.

Lemma 1.9 (Kőnig). *Every infinite tree that is finitely branching contains an infinite branch.*

Proof. By induction, we construct an infinite branch $v_0 < v_1 < \dots$ such that $\Uparrow v_n$ is infinite, for all n. Let v_0 be the root of $\mathfrak{T}$. By assumption, $\Uparrow v_0 = T$ is infinite. For the inductive step, suppose that we have already defined the path $v_0 < \dots < v_n$ such that $\Uparrow v_n$ is infinite. Since v_n has only finitely many successors $u_0, \dots, u_k$ and

$$\Uparrow v_n = \{v_n\} \cup \Uparrow u_0 \cup \dots \cup \Uparrow u_k\,,$$

there must be at least one successor u_i such that $\Uparrow u_i$ is infinite. We set $v_{n+1} := u_i$. □

If we compute a set X as the inductive fixed point of some operation then we can associate with the elements of X a rank that measures at which stage of the induction the element entered the fixed point.

Definition 1.10. Let $f : \wp(A) \to \wp(A)$ be a function that is inductive over $\varnothing$ and let $F : \mathrm{On} \to \wp(A)$ be the corresponding fixed-point induction. We associate with every element $a \in A$ a *rank* as follows. For elements $a \in F(\infty)$, we define the *rank* of a as the ordinal α such that $a \in F(\alpha + 1) \setminus F(\alpha)$. For $a \notin F(\infty)$, we set the rank of a to ∞.

Example. The power-set operation $\wp : \mathbb{S} \to \mathbb{S}$ is inductive over $\varnothing$. The corresponding notion of rank coincides with the rank $\rho(a)$ introduced in Definition A3.2.24.

Let us define a rank for trees.

Definition 1.11. Let $\mathfrak{T} = \langle T, \leq \rangle$ be a tree. The *foundation rank* $\mathrm{frk}(v)$ of a vertex $v \in T$ is the rank corresponding to the fixed-point operator $f : \wp(T) \to \wp(T)$ with

$$f(X) := \{\, v \in T \mid \uparrow v \subseteq X \,\}\,.$$

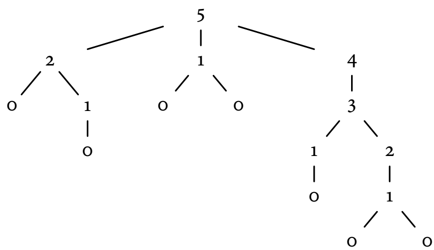
The *rank* $\mathrm{frk}(\mathfrak{T})$ of $\mathfrak{T}$ is the rank of its root.

Remark. We have $\operatorname{frk}(v) = 0$ if and only if v is a leaf of T.

In the course of this book we will introduce several ranks. Since it is cumbersome to define them in terms of fixed-point operations we will usually give more informal definitions. For a given ordinal α, we will just specify all elements a such that $a \notin F(\alpha)$. For instance, for the foundation rank the definition would have the following format:

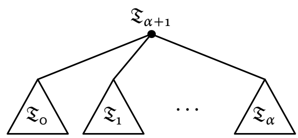
- $\operatorname{frk}(v) \geq 0$, for all $v \in T$.
- For successor ordinals, we have $\operatorname{frk}(v) \geq \alpha + 1$ if and only if there is some $u > v$ with $\operatorname{frk}(u) \geq \alpha$.
- If δ is a limit ordinal then $\operatorname{frk}(v) \geq \delta$ iff $\operatorname{frk}(v) \geq \alpha$, for all $\alpha < \delta$.

Example. (a) The tree



has foundation rank 5.

(b) For every ordinal α, we can construct a tree $\mathfrak{T}_\alpha$ of foundation rank α. $\mathfrak{T}_0$ consists just of a single vertex. If $\alpha > 0$ then we can construct $\mathfrak{T}_\alpha$ by taking the disjoint union of all $\mathfrak{T}_\beta$, $\beta < \alpha$, and adding a new vertex as the root:



Lemma 1.12. *Let $\mathfrak{T}$ be a tree and $u, v \in T$. If $u < v$ then we have*

$$\operatorname{frk}(u) > \operatorname{frk}(v) \quad \text{or} \quad \operatorname{frk}(u) = \operatorname{frk}(v) = \infty\,.$$

Lemma 1.13. *Let $\mathfrak{T}$ be a tree and $v \in T$.*

(a) $\mathrm{frk}(v) = \sup \{ \mathrm{frk}(u) + 1 \mid u$ *is a successor of* $v \}$.

(b) *We have* $\mathrm{frk}(v) = \infty$ *if and only if* $\Uparrow v$ *contains an infinite path.*

Proof. (a) Let F be the fixed point induction used to define $\mathrm{frk}(v)$. If u is a successor of v then $u \in F(\mathrm{frk}(u) + 1) \setminus F(\mathrm{frk}(u))$ and $u \in \uparrow v$ implies that $v \notin F(\mathrm{frk}(u) + 1)$. Hence, $\mathrm{frk}(v) \geq \mathrm{frk}(u) + 1$. For the converse, suppose that $\mathrm{frk}(v) > \alpha$, i.e., $v \notin F(\alpha + 1)$. There exists some vertex $w > v$ with $w \notin F(\alpha)$. Let u be the successor of v such that $v < u \leq w$. If $u < w$ then, by definition of $F(\alpha+1)$, it follows that $u \notin F(\alpha+1) \supseteq F(\alpha)$. Otherwise, we have $u = w \notin F(\alpha)$. Consequently, for every $\alpha < \mathrm{frk}(v)$, there exists some successor u with $\mathrm{frk}(u) \geq \alpha$.

(b) If $\mathrm{frk}(v) = \infty$ then (a) implies that there is some successor u of v with $\mathrm{frk}(u) = \infty$. Hence, we can inductively construct an infinite path $v = v_0 < v_1 < \dots$ such that $\mathrm{frk}(v_n) = \infty$, for all n.

Conversely, if $v_0 < v_1 < \dots$ is an infinite path then it follows by induction on α that $v_n \notin F(\alpha)$, for all n. Therefore, we have $\mathrm{frk}(v_n) = \infty$. □

Corollary 1.14. *Let $\mathfrak{T} = \langle T, \leq \rangle$. We have* $\mathrm{frk}(\mathfrak{T}) < \infty$ *if and only if the partial order $\mathfrak{T}^{\mathrm{op}} := \langle T, \geq \rangle$ is well-founded.*

Lemma 1.15. *Let $T \subseteq \kappa^{<\alpha}$. If* $\mathrm{frk}(T) < \infty$ *then* $\mathrm{frk}(T) < \kappa^+$.

Proof. Suppose, for a contradiction that $\kappa^+ \leq \mathrm{frk}(T) < \infty$. By the preceding corollary, we know that the inverse ordering $\geq$ is well-founded. Hence, there exists a maximal vertex $v \in T$ such that $\mathrm{frk}(v) \geq \kappa^+$. Let S be the set of successors of v. By maximality and Lemma 1.13, it follows that

$$\kappa^+ = \mathrm{frk}(v) = \sup \{ \mathrm{frk}(u) + 1 \mid u \in S \},$$

where $\mathrm{frk}(u) < \kappa^+$. Hence, κ^+ is the supremum of a set of $|S| < \kappa^+$ ordinals each of which is less then κ^+. This contradicts the fact that every successor cardinal is regular. □

2. Lattices

Lattices are partial orders that, although not necessarily complete, enjoy a certain weak completeness property. Instead of requiring that every subset has a supremum and an infimum we only do so for all finite sets.

Definition 2.1. (a) A partial order $\mathfrak{L} = \langle L, \sqsubseteq \rangle$ is a *lower semilattice* if every pair $a, b \in L$ has a greatest lower bound $\inf\{a, b\}$. Analogously we call $\mathfrak{L}$ an *upper semilattice* if every pair $a, b \in L$ has a least upper bound $\sup\{a, b\}$.

(b) A *lattice* is a structure $\mathfrak{L} = \langle L, \sqcup, \sqcap, \sqsubseteq \rangle$ where $\sqsubseteq$ is a partial order and

$$a \sqcap b = \inf\{a, b\} \quad \text{and} \quad a \sqcup b = \sup\{a, b\}, \quad \text{for } a, b \in L.$$

A lattice $\mathfrak{L}$ is *bounded* if it has a least element $\perp$ and a greatest element $\top$.

Remark. (a) If $\langle L, \sqsubseteq \rangle$ is both an upper and a lower semilattice then there exists a unique expansion $\langle L, \sqcap, \sqcup, \sqsubseteq \rangle$ to a lattice. Informally we will therefore also call the order $\langle L, \sqsubseteq \rangle$ a lattice. But note that by a homomorphism between lattices we always mean a homomorphism with respect to the full signature.

Similarly, we will also call structures of the form $\langle L, \sqcap, \sqsubseteq \rangle$ with

$$a \sqcap b = \inf\{a, b\}$$

a lower semilattice, and structures $\langle L, \sqcup, \sqsubseteq \rangle$ with

$$a \sqcup b = \sup\{a, b\}$$

an upper semilattice.

(b) All complete partial orders and all linear orders are lattices.

Example. (a) The divisibility order $\langle \mathbb{N}, \mid \rangle$ is a lattice where $m \sqcap n$ is the greatest common divisor of m and n and $m \sqcup n$ is their least common multiple.

(b) $\mathfrak{Cong}(\mathfrak{A})$ and $\mathfrak{Sub}(\mathfrak{A})$ are lattices.

(c) Let $\mathfrak{A}$ be a structure and S the family of all finitely generated substructures of $\mathfrak{A}$. Then $\langle S, \subseteq \rangle$ is a lattice.

Exercise 2.1. (a) Let $\mathfrak{L}$ be a lattice and $a, b \in L$. Prove that the interval $[a, b]$ induces a sublattice.

(b) Prove that every substructure of a lattice is a lattice.

The ordering $\sqsubseteq$ is actually redundant since it can be defined with the help of $\sqcap$ or $\sqcup$.

Lemma 2.2. *Let* $\mathfrak{L} = \langle L, \sqcap, \sqcup, \sqsubseteq \rangle$ *be a lattice.*

(a) *For* $a, b \in L$, *we have*

$$a \sqsubseteq b \quad \text{iff} \quad a \sqcap b = a \quad \text{iff} \quad a \sqcup b = b\,.$$

(b) *If* $b \sqsubseteq c$ *then*

$$a \sqcap b \sqsubseteq a \sqcap c \quad \text{and} \quad a \sqcup b \sqsubseteq a \sqcup c\,.$$

Proof. (a) is trivial. For (b), we have

$$a \sqcap b = a \sqcap (b \sqcap c) = (a \sqcap a) \sqcap (b \sqcap c) = (a \sqcap b) \sqcap (a \sqcap c),$$

by (a). Again by (a), it follows that $a \sqcap b \sqsubseteq a \sqcap c$. The other inequality is proved in the same way. □

Lemma 2.3. *A structure* $\mathfrak{L} = \langle L, \sqcap, \sqsubseteq \rangle$ *is a lower semilattice if and only if, for all* $a, b, c \in L$, *we have*

$$a \sqsubseteq b \quad \text{iff} \quad a \sqcap b = a\,,$$

$$a \sqcap a = a\,, \qquad \textit{(idempotence)}$$

$$a \sqcap b = b \sqcap a\,, \qquad \textit{(commutativity)}$$

$$a \sqcap (b \sqcap c) = (a \sqcap b) \sqcap c\,. \qquad \textit{(associativity)}$$

Proof. ($\Rightarrow$) If $\mathfrak{L}$ is a lower semilattice then the above conditions follow immediately from the definition of the infimum.

($\Leftarrow$) Suppose that $\mathfrak{L}$ satisfies the above conditions. First we show that $\sqsubseteq$ is a partial order. It is reflexive since $a \sqcap a = a$ implies that $a \sqsubseteq a$. For antisymmetry, note that $a \sqsubseteq b$ and $b \sqsubseteq a$ implies that

$$a = a \sqcap b = b \sqcap a = b\,.$$

Finally, for transitivity suppose that $a \sqsubseteq b$ and $b \sqsubseteq c$. Then we have $a \sqcap b = a$ and $b \sqcap c = b$. It follows that

$$a \sqcap c = (a \sqcap b) \sqcap c = a \sqcap (b \sqcap c) = a \sqcap b = a\,.$$

Hence, we have $a \sqsubseteq c$.

It remains to prove that $a \sqcap b = \inf\{a, b\}$. We have

$$(a \sqcap b) \sqcap b = a \sqcap (b \sqcap b) = a \sqcap b\,,$$

which implies that $a \sqcap b \sqsubseteq b$. Similarly, we obtain $a \sqcap b \sqsubseteq a$. Consequently, $a \sqcap b$ is a lower bound of $\{a, b\}$. Furthermore, if c is some element with $c \sqsubseteq a$ and $c \sqsubseteq b$ then we have $c \sqcap a = c$ and $c \sqcap b = c$ and it follows that

$$c \sqcap (a \sqcap b) = (c \sqcap a) \sqcap b = c \sqcap b = c\,.$$

Hence, $c \sqsubseteq a \sqcap b$ and $a \sqcap b$ is the greatest lower bound of $\{a, b\}$. □

As an immediate consequence we obtain the following characterisation of lattices.

Lemma 2.4. *A structure $\mathfrak{L} = \langle L, \sqcap, \sqcup, \sqsubseteq \rangle$ is a lattice if and only if, for all $a, b, c \in L$, we have*

$$a \sqsubseteq b \quad \text{iff} \quad a \sqcap b = a$$

$$\text{and} \qquad a \sqcap a = a \qquad a \sqcup a = a \qquad \textit{(idempotence)}$$

$$a \sqcap b = b \sqcap a \qquad a \sqcup b = b \sqcup a \qquad \textit{(commutativity)}$$

$$a \sqcap (a \sqcup b) = a \qquad a \sqcup (a \sqcap b) = a \qquad \textit{(absorption)}$$

$$a \sqcap (b \sqcap c) = (a \sqcap b) \sqcap c \qquad \textit{(associativity)}$$

$$a \sqcup (b \sqcup c) = (a \sqcup b) \sqcup c$$

We conclude this section with a look at three important subclasses of lattices.

Definition 2.5. Let $\mathfrak{L} = \langle L, \sqcap, \sqcup, \sqsubseteq \rangle$ be a lattice.

(a) $\mathfrak{L}$ is *modular* if, for all $a, b, c \in L$, we have that

$$a \sqsubseteq b \quad \text{implies} \quad a \sqcup (b \sqcap c) = b \sqcap (a \sqcup c)\,.$$

(b) $\mathfrak{L}$ is *distributive* if, for all $a, b, c \in L$, we have

$$a \sqcap (b \sqcup c) = (a \sqcap b) \sqcup (a \sqcap c)$$

$$\text{and} \quad a \sqcup (b \sqcap c) = (a \sqcup b) \sqcap (a \sqcup c)\,.$$

(c) $\mathfrak{L}$ is *boolean* if it is distributive, bounded, and, for every $a \in L$ there is some element $a^* \in L$ such that

$$a \sqcap a^* = \bot \quad \text{and} \quad a \sqcup a^* = \top\,.$$

The element a^* is called the *complement* of a. If $\mathfrak{L}$ is a boolean lattice then we call the structure $\langle L, \sqcap, \sqcup, {}^* \rangle$ a *boolean algebra.*

Example. For every set A, $\langle \mathscr{P}(A), \cap, \cup, {}^* \rangle$ forms a boolean algebra with $X^* := A \setminus X$.

Remark. Note that every sublattice of a power-set lattice $\langle \mathscr{P}(A), \subseteq \rangle$ is distributive.

Exercise 2.2. Prove that every sublattice of a distributive lattice is distributive and that every sublattice of a modular lattice is modular.

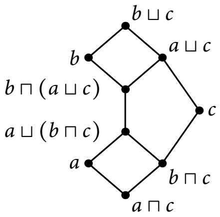


Figure 2.. The general situation

To better understand the modularity condition we have shown in Figure 2 the corresponding situation in an arbitrary lattice. (Some of the depicted elements might coincide.)

Lemma 2.6. *If $a \sqsubseteq b$ then we have*

$$a \sqsubseteq a \sqcup (b \sqcap c) \sqsubseteq b \sqcap (a \sqcup c) \sqsubseteq b\,.$$

Proof. The first and the last inequality follow immediately from the definition of $\sqcup$ and $\sqcap$. For the remaining inequality, note that

$$\begin{aligned} & a \sqsubseteq b \quad \text{and } b \sqcap c \sqsubseteq b && \text{implies} \quad a \sqcup (b \sqcap c) \sqsubseteq b\,, \\ \text{and} \quad & a \sqsubseteq a \sqcup c \text{ and } b \sqcap c \sqsubseteq c \sqsubseteq a \sqcup c && \text{implies} \quad a \sqcup (b \sqcap c) \sqsubseteq a \sqcup c\,. \end{aligned}$$ □

In general the distributive laws also hold only in one direction.

Lemma 2.7. *In every lattice $\mathfrak{L}$, we have*

$$\begin{aligned} & a \sqcap (b \sqcup c) \sqsupseteq (a \sqcap b) \sqcup (a \sqcap c) \\ \textit{and} \quad & a \sqcup (b \sqcap c) \sqsubseteq (a \sqcup b) \sqcap (a \sqcup c)\,, \end{aligned}$$

for all $a, b, c \in L$.

Lemma 2.8. *Every distributive lattice is modular.*

Proof. $a \sqsubseteq b$ implies $a \sqcup b = b$. Consequently, we have

$$a \sqcup (b \sqcap c) = (a \sqcup b) \sqcap (a \sqcup c) = b \sqcap (a \sqcup c)\,.$$ □

Lemma 2.9. *A lattice $\mathfrak{L}$ is modular if, and only if,*

$$a \sqsubseteq b \text{ and } a \sqcup c = b \sqcup c \quad \text{implies} \quad a \sqcup (b \sqcap c) = b\,.$$

Proof. ($\Rightarrow$) If $a \sqsubseteq b$ and $a \sqcup c = b \sqcup c$, modularity implies that

$$b = b \sqcap (b \sqcup c) = b \sqcap (a \sqcup c) = a \sqcup (b \sqcap c)\,.$$

($\Leftarrow$) Suppose that $a \sqsubseteq b$. To show that

$$a \sqcup (b \sqcap c) = b \sqcap (a \sqcup c)$$

we consider the element $x := b \sqcap (a \sqcup c)$. Note that $a \sqsubseteq x \sqsubseteq a \sqcup c$ implies $a \sqcup c = x \sqcup c$. By assumption, it therefore follows that

$$a \sqcup (x \sqcap c) = x\,.$$

Furthermore, by Lemma 2.6 we have

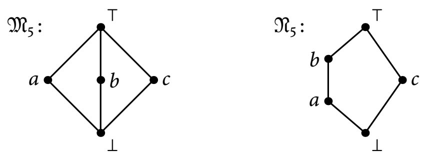
$$b \sqcap c \sqsubseteq a \sqcup (b \sqcap c) \sqsubseteq x \sqsubseteq b\,,$$

which implies that $x \sqcap c = b \sqcap c$. Hence,

$$a \sqcup (b \sqcap c) = a \sqcup (x \sqcap c) = x\,.$$ □

Distributive and modular lattices can be characterised in terms of forbidden configurations.

Definition 2.10. Let $\mathfrak{M}_5$ and $\mathfrak{N}_5$ be the following lattices:



Theorem 2.11. *Let $\mathfrak{L}$ be a lattice.*

(a) *$\mathfrak{L}$ is modular iff there exists no embedding $\mathfrak{N}_5 \to \mathfrak{L}$.*

(b) *$\mathfrak{L}$ is distributive iff there exists neither an embedding $\mathfrak{M}_5 \to \mathfrak{L}$ nor an embedding $\mathfrak{N}_5 \to \mathfrak{L}$.*

Proof. (a) ($\Rightarrow$) Suppose that $h : \mathfrak{N}_5 \to \mathfrak{L}$ is an embedding. Then $h(a) \sqsubseteq h(b)$ but

$$\begin{aligned} h(a) \sqcup (h(b) \sqcap h(c)) &= h(a) \sqcup h(\bot) = h(a) \\ &\neq h(b) = h(b) \sqcap h(\top) \\ &= h(b) \sqcap (h(a) \sqcup h(c)) . \end{aligned}$$

Hence, $\mathfrak{L}$ is not modular.

($\Leftarrow$) Suppose that $\mathfrak{L}$ is not modular. Then there exist elements $x, y, z \in L$, such that $x \sqsubseteq y$ but $x \sqcup (y \sqcap z) \neq y \sqcap (x \sqcup z)$. Set

$$\begin{aligned} a &:= x \sqcup (y \sqcap z) , \qquad & d &:= b \sqcup z , \\ b &:= y \sqcap (x \sqcup z) , \qquad & e &:= a \sqcap z . \end{aligned}$$

We claim that the inclusion map $\{a, b, z, d, e\} \to L$ is the desired embedding.

Note that $x \sqsubseteq y$ and $x \sqsubseteq x \sqcup z$ implies

$$a = x \sqcup (y \sqcap z) \sqsubseteq x \sqcup (y \sqcap (x \sqcup z)) = y \sqcap (x \sqcup z) = b .$$

Hence, we have $e \sqsubseteq a \sqsubset b \sqsubseteq d$ and $e \sqsubseteq z \sqsubseteq d$. It remains to prove that $a \not\sqsubseteq z \not\sqsubseteq b$. If $a \sqsubseteq z$ then we have

$$z = a \sqcup z = (x \sqcup (y \sqcap z)) \sqcup z = x \sqcup ((y \sqcap z) \sqcup z) = x \sqcup z$$

which implies that

$$a \sqsubset b = y \sqcap (x \sqcup z) = y \sqcap z \sqsubseteq x \sqcup (y \sqcap z) = a .$$

A contradiction. The assumption that $z \sqsubseteq b$ leads to a similar contradiction.

(b) By (a) it is sufficient to prove that a modular lattice $\mathfrak{L}$ is distributive if and only if there is no embedding $\mathfrak{M}_5 \to \mathfrak{L}$.

($\Rightarrow$) Suppose that $h : \mathfrak{M}_5 \to \mathfrak{L}$ is an embedding. Then we have

$$\begin{aligned} h(a) \sqcup (h(b) \sqcap h(c)) &= h(a) \sqcup h(\bot) = h(a) \\ &\neq h(\top) = h(\top) \sqcap h(\top) \\ &= (h(a) \sqcup h(b)) \sqcap (h(a) \sqcup h(c)) \,. \end{aligned}$$

Hence, $\mathfrak{L}$ is not distributive.

($\Leftarrow$) Suppose that $\mathfrak{L}$ is not distributive. Then we can find elements $x, y, z \in L$ such that

$$x \sqcup (y \sqcap z) \sqsubset (x \sqcup y) \sqcap (x \sqcup z) \,.$$

Set

$$\begin{aligned} d &:= (x \sqcap y) \sqcup (x \sqcap z) \sqcup (y \sqcap z) \,, & a &:= (x \sqcap e) \sqcup d \,, \\ e &:= (x \sqcup y) \sqcap (x \sqcup z) \sqcap (y \sqcup z) \,, & b &:= (y \sqcap e) \sqcup d \,, \\ & & c &:= (z \sqcap e) \sqcup d \,. \end{aligned}$$

By definition we have $d \sqsubseteq a, b, c \sqsubseteq e$. We claim that $\{a, b, c, d, e\}$ induce a copy of $\mathfrak{M}_5$. By absorption, we have

$$x \sqcup d = x \sqcup x \sqcup (y \sqcap z) = x \sqcup (y \sqcap z) \,.$$

On the other hand, since $\mathfrak{L}$ is modular and $x \sqsubseteq (x \sqcup y) \sqcap (x \sqcup z)$ we have

$$\begin{aligned} x \sqcup e &= x \sqcup [(x \sqcup y) \sqcap (x \sqcup z) \sqcap (y \sqcup z)] \\ &= [(x \sqcup y) \sqcap (x \sqcup z)] \sqcap [x \sqcup (y \sqcup z)] \\ &= (x \sqcup y) \sqcap (x \sqcup z) \,. \end{aligned}$$

Hence, $x \sqcup d \sqsubset x \sqcup e$ which implies that $d \sqsubset e$. It remains to prove that

$$a \sqcap b = a \sqcap c = b \sqcap c = d \,,$$

and $$a \sqcup b = a \sqcup c = b \sqcup c = e \,.$$

By symmetry and duality, we only need to show that $a \sqcap b = d$. Applying the absorption law twice we have

$$\begin{aligned}(a \sqcap b) \sqcap d &= ((x \sqcap e) \sqcup d) \sqcap ((y \sqcap e) \sqcup d) \sqcap d \\ &= ((x \sqcap e) \sqcup d) \sqcap d = d\,.\end{aligned}$$

Finally, note that the elements a, b, c are distinct since $a = b$ would imply that $d = a \sqcap b = a = a \sqcup b = e$. □

3. *Ideals and filters*

The notions of a normal subgroup or an ideal of a ring can be generalised to lattices.

Definition 3.1. Let $\mathfrak{L} = \langle L, \sqcap, \sqcup, \sqsubseteq \rangle$ be a lattice.

(a) A nonempty initial segment $\mathfrak{a} \subseteq L$ is an *ideal* if $a, b \in \mathfrak{a}$ implies $a \sqcup b \in \mathfrak{a}$. Similarly, we call a nonempty final segment $\mathfrak{u} \subseteq L$ a *filter* if $a, b \in \mathfrak{u}$ implies $a \sqcap b \in \mathfrak{u}$.

(b) An ideal or filter is *proper* if it is a proper subset of L. A proper ideal or filter $\mathfrak{a}$ is *maximal* if there exists no proper ideal or filter $\mathfrak{b}$ such that $\mathfrak{a} \subset \mathfrak{b} \subset L$. Ideals of the form $\Downarrow a$, for some $a \in L$, and filters of the form $\Uparrow a$ are called *principal*.

Example. (a) In every bounded lattice we have the *trivial ideal* $\{\bot\}$ and the *trivial filter* $\{\top\}$.

(b) Consider $\langle \mathscr{P}(A), \subseteq \rangle$. We can define an ideal $\mathfrak{a}$ and a filter $\mathfrak{u}$ by

$$\begin{aligned}\mathfrak{a} &:= \{\, X \subseteq A \mid X \text{ is finite} \,\}\,, \\ \mathfrak{u} &:= \{\, X \subseteq A \mid A \setminus X \text{ is finite} \,\}\,.\end{aligned}$$

They are proper if and only if A is infinite.

(c) Let K be a field and consider the lattice of all polynomials over K with leading coefficient 1 ordered by the inverse divisibility relation

$$p \sqsubseteq q \quad :\text{iff} \quad q \mid p\,.$$

We have $\bot = 0$ and $\top = 1$. $p \sqcap q$ is the least common multiple of p and q and $p \sqcup q$ is their greatest common divisor. For every subset $A \subseteq K$, we obtain the ideal

$$I(A) := \{ p \in K[x] \mid p(a) = 0 \text{ for all } a \in A \} .$$

Remark. To every lattice $\mathfrak{L} = \langle L, \sqcap, \sqcup, \sqsubseteq \rangle$ we can associate the *opposite lattice* $\mathfrak{L}^{\mathrm{op}} = \langle L, \sqcup, \sqcap, \sqsupseteq \rangle$ where the order is reversed. Obviously, this functions maps filters of $\mathfrak{L}$ to ideals of $\mathfrak{L}^{\mathrm{op}}$ and ideals of $\mathfrak{L}$ to filters. Therefore, we will state and prove many results only in one version, either for filters or for ideals. The other half can be obtained by duality.

Ideal and filters can be characterised in terms of a suitable closure operator.

Definition 3.2. Let $\mathfrak{L}$ be a lattice and $X \subseteq L$. We define

$$\mathrm{cl}_\downarrow(X) := \{ b \in L \mid b \sqsubseteq a_0 \sqcup \cdots \sqcup a_n \text{ for some } a_0, \ldots, a_n \in X, n < \omega \},$$
$$\mathrm{cl}_\uparrow(X) := \{ b \in L \mid b \sqsupseteq a_0 \sqcap \cdots \sqcap a_n \text{ for some } a_0, \ldots, a_n \in X, n < \omega \}.$$

Lemma 3.3. *Let $\mathfrak{L}$ be a lattice.*

(a) *If $\mathfrak{L}$ is bounded then* $\mathrm{cl}_\downarrow$ *and* $\mathrm{cl}_\uparrow$ *are closure operators on L with finite character.*

(b) *A nonempty set $X \subseteq L$ is an ideal if and only if it is* $\mathrm{cl}_\downarrow$*-closed.*

(c) *A nonempty set $X \subseteq L$ is a filter if and only if it is* $\mathrm{cl}_\uparrow$*-closed.*

Corollary 3.4. *The set of all ideals of a bounded lattice $\mathfrak{L}$ forms a complete partial order. It is closed under arbitrary intersections and under unions of chains.*

Corollary 3.5. *Let $\mathfrak{L}$ be a lattice. If $\mathfrak{a}$ is a proper ideal and $\mathfrak{u}$ a proper filter with $\mathfrak{a} \cap \mathfrak{u} = \varnothing$ then the set*

$$\mathcal{I} := \{ \mathfrak{b} \mid \mathfrak{b} \text{ a proper ideal with } \mathfrak{a} \subseteq \mathfrak{b} \text{ and } \mathfrak{b} \cap \mathfrak{u} = \varnothing \}$$

contains a maximal element.

Proof. We show that $\mathcal{I}$ is inductively ordered. Then it contains a maximal element by Zorn's Lemma. Let $C \subseteq \mathcal{I}$ be a chain. Then $\mathfrak{c} := \bigcup C$ is an ideal. Since $\mathfrak{a} \cap \mathfrak{u} = \varnothing$, for all $\mathfrak{a} \in \mathcal{I}$, we have $\mathfrak{c} \cap \mathfrak{u} = \varnothing$. In particular, $\mathfrak{c}$ is proper. Consequently, $\mathfrak{c} \in \mathcal{I}$. □

Lemma 3.6. *Let $\mathfrak{L}$ be a lattice. The following statements are equivalent:*

(1) *Every ideal of $\mathfrak{L}$ is principal.*

(2) *Every strictly increasing sequence $\mathfrak{a}_0 \subset \mathfrak{a}_1 \subset \dots$ of ideals of $\mathfrak{L}$ is finite.*

(3) *The inverse subset relation is a well-order on the set of all ideals of $\mathfrak{L}$.*

Proof. Clearly, (2) is equivalent to (3). Let us prove that (2) implies (1). Suppose that there exists an ideal $\mathfrak{a}$ that is not principal. We select a sequence $(a_n)_{n<\omega}$ of elements of $\mathfrak{a}$ as follows. Let $a_0 \in \mathfrak{a}$ be arbitrary. If $a_0, \dots, a_n \in \mathfrak{a}$ have already been chosen then, since $\mathfrak{a}$ is not principal, we can find an element $a_{n+1} \in \mathfrak{a} \setminus \Downarrow(a_0 \sqcup \dots \sqcup a_n) \neq \varnothing$. This way we obtain an infinite strictly increasing sequence of ideals

$$\Downarrow a_0 \subset \Downarrow(a_0 \sqcup a_1) \subset \dots \subset \Downarrow(a_0 \sqcup \dots \sqcup a_n) \subset \dots,$$

as desired.

It remains to prove the converse. Suppose that $\mathfrak{a}_0 \subset \mathfrak{a}_1 \subset \dots$ is an infinite strictly increasing sequence of ideals. Their union $\mathfrak{b} := \bigcup_n \mathfrak{a}_n$ is again an ideal. We claim that $\mathfrak{b}$ is not principal. Suppose otherwise. Then $\mathfrak{b} = \Downarrow b$, for some $b \in \mathfrak{b}$. Since $\mathfrak{b} = \bigcup_n \mathfrak{a}_n$ there is some index n such that $b \in \mathfrak{a}_n$. It follows that $\mathfrak{b} = \Downarrow b \subseteq \mathfrak{a}_n \subset \mathfrak{a}_{n+1} \subseteq \mathfrak{b}$. Contradiction. □

Ideals and filters in lattices play the same role with regard to homomorphisms and congruences as normal subgroups in group theory or ideals in ring theory. The main difference is that, since the lattice operations are not invertible, there might be several congruences inducing the same ideal.

Lemma 3.7. *Let $h : \mathfrak{L} \to \mathfrak{K}$ be a homomorphism between lattices and let $\mathfrak{a} \subseteq K$ be an ideal of $\mathfrak{K}$. If $h^{-1}[\mathfrak{a}]$ is nonempty then it is an ideal of $\mathfrak{L}$.*

Proof. Suppose that $a \in h^{-1}[\mathfrak{a}]$ and $b \sqsubseteq a$. Since h is a homomorphism it follows that $h(b) \sqsubseteq h(a) \in \mathfrak{a}$. Consequently, we have $h(b) \in \mathfrak{a}$ and $b \in h^{-1}[\mathfrak{a}]$.

Similarly, if $a, b \in h^{-1}[\mathfrak{a}]$ then $h(a), h(b) \in \mathfrak{a}$ implies that $h(a \sqcup b) = h(a) \sqcup h(b) \in \mathfrak{a}$. Hence, we have $a \sqcup b \in h^{-1}[\mathfrak{a}]$. □

Corollary 3.8. *Let $h : \mathfrak{L} \to \mathfrak{K}$ be a surjective homomorphism between lattices where $\mathfrak{K}$ is bounded.*

(a) $h^{-1}(\bot)$ *is an ideal.*

(b) $h^{-1}(\top)$ *is a filter.*

Corollary 3.9. *Let $\mathfrak{L}$ be a bounded lattice. If $\sim$ is a congruence of $\mathfrak{L}$ then $[\bot]_\sim$ is an ideal and $[\top]_\sim$ is a filter.*

There are important cases where we would like to apply lattice theory but which do not fall under the above definition of a lattice because the underlying 'order' $\sqsubseteq$ fails to be a partial order. A prominent example is given by rings like $\langle \mathbb{Z}, \mid \rangle$ and $\langle \mathbb{R}[x], \mid \rangle$ where the divisibility relation $\mid$ is not antisymmetric. In the ring of integers, for instance, we have

$$1 \mid -1 \quad \text{and} \quad -1 \mid 1\,.$$

Definition 3.10. A graph $\langle V, E \rangle$ is a *preorder* if E is reflexive and transitive.

Example. If R is a ring then the divisibility relation

$$x \mid y \quad :\text{iff} \quad y = axb\,, \text{ for some } a, b \in R$$

forms a preorder on R.

Every preorder has a quotient that is a partial order.

Lemma 3.11. *Let $\mathfrak{P} = \langle P, \leq \rangle$ be a preorder and define*

$$x \sim y \quad :\text{iff} \quad x \leq y \text{ and } y \leq x\,.$$

$\sim$ is a congruence on $\mathfrak{P}$ and the quotient $\langle P, \leq \rangle / {\sim}$ is a partial order.

Proof. By definition, $\sim$ is symmetric. And since $\leq$ is a preorder it follows that $\sim$ is reflexive and transitive. Therefore, $\sim$ is an equivalence relation. Suppose that $x \sim x'$ and $y \sim y'$. If $x \leq y$ then $x' \leq x \leq y \leq y'$ implies $x' \leq y'$. Hence, $\sim$ is a congruence.

It is easy to see that $\mathfrak{P}/{\sim}$ is a preorder. It remains to show that it is also antisymmetric. Let $[x]_\sim, [y]_\sim \in P/{\sim}$ with $[x]_\sim \leq [y]_\sim$ and $[y]_\sim \leq [x]_\sim$. Then $x \leq y$ and $y \leq x$ implies $x \sim y$. Hence, $[x]_\sim = [y]_\sim$. □

We can generalise many concepts of lattice theory to preorders.

Definition 3.12. (a) A *prelattice* is a preorder $\langle L, \leq \rangle$ such that the corresponding partial order $\langle L, \leq \rangle/{\sim}$ is a lattice.

(b) Let $\mathfrak{L}$ be a prelattice and $\pi : \mathfrak{L} \to \mathfrak{L}/{\sim}$ the canonical projection to the corresponding lattice. An *ideal* of $\mathfrak{L}$ is a set of the form $\pi^{-1}[\mathfrak{a}]$ where $\mathfrak{a}$ is an ideal of $\mathfrak{L}/{\sim}$. Similarly, if $\mathfrak{u}$ is a filter of $\mathfrak{L}/{\sim}$ then we call the set $\pi^{-1}[\mathfrak{u}]$ a *filter* of $\mathfrak{L}$. In the same way we can generalise other notions to prelattices, like proper and principal ideals.

Example. Let $\langle R, +, -, \cdot, 0, 1 \rangle$ be a commutative factorial ring. The divisibility order $\langle R, \mid \rangle$ is a prelattice and a subset $I \subseteq R$ is a ring-theoretic ideal if, and only if, it is a filter of $\langle R, \mid \rangle$.

4. *Prime ideals and ultrafilters*

Definition 4.1. A proper ideal $\mathfrak{a}$ is a *prime ideal* if

$$x \sqcap y \in \mathfrak{a} \quad \text{implies} \quad x \in \mathfrak{a} \text{ or } y \in \mathfrak{a}\,.$$

Similarly, we call a proper filter $\mathfrak{u}$ an *ultrafilter* if

$$x \sqcup y \in \mathfrak{u} \quad \text{implies} \quad x \in \mathfrak{u} \text{ or } y \in \mathfrak{u}\,.$$

In the special case that the lattice in question is the power-set algebra $\langle \mathcal{P}(X), \cup, \cap, \subseteq \rangle$ we call $\mathfrak{u}$ an *ultrafilter on X*.

Example. (a) Let $\mathfrak{N} := \langle \mathbb{N}, \mid \rangle$. A filter $\mathfrak{u} \subseteq \mathbb{N}$ is an ultrafilter if, and only if, either $\mathfrak{u} = \{0\}$ or there exists a prime number p such that

$$\mathfrak{u} = \{\, kp \mid k \in \mathbb{N} \,\} .$$

(b) Let $\mathfrak{F} = \langle F, \subseteq \rangle$ where

$$F := \{\, X \subseteq \omega \mid X \text{ or } \omega \setminus X \text{ is finite} \,\} .$$

Then $\mathfrak{F}$ is a lattice and we have the following ultrafilters:

$$\begin{aligned} \mathfrak{u}_n &:= \Uparrow\{n\} , \qquad \text{for } n < \omega , \\ \mathfrak{u}_\infty &:= \{\, X \subseteq \omega \mid \omega \setminus X \text{ is finite} \,\} . \end{aligned}$$

Lemma 4.2. *A set $X \subseteq L$ is a prime ideal if, and only if, its complement $L \setminus X$ is an ultrafilter.*

Proof. By duality it is sufficient to prove one direction. Let $\mathfrak{a} \subseteq L$ be a prime ideal. We claim that $\mathfrak{u} := L \setminus \mathfrak{a}$ is an ultrafilter. Since $\mathfrak{a}$ is proper and nonempty so is $\mathfrak{u}$. If $a \sqsubseteq b$ then $b \in \mathfrak{a}$ implies $a \in \mathfrak{a}$. Consequently, $a \in \mathfrak{u}$ implies $b \in \mathfrak{u}$ and $\mathfrak{u}$ is a final segment. If $a \sqcap b \in \mathfrak{a}$ then we have $a \in \mathfrak{a}$ or $b \in \mathfrak{a}$ since $\mathfrak{a}$ is prime. Thus, $a, b \in \mathfrak{u}$ implies $a \sqcap b \in \mathfrak{u}$ and $\mathfrak{u}$ is a filter. Finally, $a, b \in \mathfrak{a}$ implies $a \sqcup b \in \mathfrak{a}$. Hence, if $a \sqcup b \in \mathfrak{u}$ then we have $a \in \mathfrak{u}$ or $b \in \mathfrak{u}$. □

Prime ideals can be characterised in terms of homomorphisms.

Definition 4.3. Let $\mathfrak{B}_2$ denote the lattice with universe $[2]$ and ordering $0 \leq 1$. And $\mathfrak{B}_{2\times 2}$ is the lattice with universe $[2] \times [2]$ and ordering

$$\langle i, k \rangle \leq \langle j, l \rangle \quad :\text{iff} \quad i \leq j \text{ and } k \leq l .$$

Remark. $\mathfrak{B}_2$ and $\mathfrak{B}_{2\times 2}$ are boolean lattices.

Lemma 4.4. *Let $h : \mathfrak{L} \to \mathfrak{B}_2$ be a surjective lattice homomorphism.*

(a) $h^{-1}(0)$ *is a prime ideal.*

(b) $h^{-1}(1)$ *is an ultrafilter.*

Proof. Let $\mathfrak{a} := h^{-1}(0)$. We have already seen in Lemma 3.8 that $\mathfrak{a}$ is an ideal. To show that it is prime suppose that $a \sqcap b \in \mathfrak{a}$. Then $h(a) \sqcap h(b) = h(a \sqcap b) = 0$ implies that $h(a) = 0$ or $h(b) = 0$. Hence, $a \in \mathfrak{a}$ or $b \in \mathfrak{a}$. □

Lemma 4.5. *Let* $\mathfrak{L}$ *be a lattice,* $\mathfrak{a}$ *a prime ideal, and* $\mathfrak{u}$ *an ultrafilter with* $\mathfrak{a} \cap \mathfrak{u} = \varnothing$.

(a) *There exists a homomorphism* $h : \mathfrak{L} \to \mathfrak{B}_2$ *with* $h^{-1}(0) = \mathfrak{a}$.

(b) *There exists a homomorphism* $h : \mathfrak{L} \to \mathfrak{B}_2$ *with* $h^{-1}(1) = \mathfrak{u}$.

(c) *There exists a homomorphism* $h : \mathfrak{L} \to \mathfrak{B}_{2\times 2}$ *with* $h^{-1}(\langle 0, 0\rangle) = \mathfrak{a}$ *and* $h^{-1}(\langle 1, 1\rangle) = \mathfrak{u}$.

Proof. (a) We claim that the function

$$h(x) := \begin{cases} 0 & \text{if } x \in \mathfrak{a}\,, \\ 1 & \text{if } x \notin \mathfrak{a}\,. \end{cases}$$

is the desired homomorphism. By definition we have $\mathfrak{a} = h^{-1}(0)$. Therefore, we only need to check that h is indeed a homomorphism.

If $x, y \notin \mathfrak{a}$ then we have $x \sqcap y \notin \mathfrak{a}$ since $\mathfrak{a}$ is prime. It follows that

$$h(x \sqcap y) = 1 = 1 \sqcap 1 = h(x) \sqcap h(y)\,.$$

Otherwise, we may assume, by symmetry, that $x \in \mathfrak{a}$. Since $x \sqcap y \sqsubseteq x$ we have $x \sqcap y \in \mathfrak{a}$ and

$$h(x \sqcap y) = 0 = 0 \sqcap h(y) = h(x) \sqcap h(x)\,.$$

The claim that $h(x \sqcup y) = h(x) \sqcup h(y)$ is shown analogously. If $x, y \in \mathfrak{a}$ then $x \sqcup y \in \mathfrak{a}$ and we have $h(x \sqcup y) = 0 = h(x) \sqcup h(y)$. Otherwise, by symmetry, we may assume that $x \notin \mathfrak{a}$. Hence, $x \sqcup y \notin \mathfrak{a}$ which implies that $h(x \sqcup y) = 1 = h(x) \sqcup h(y)$.

(b) follows from (a) by duality.

(c) Let $h_0, h_1 : \mathfrak{L} \to \mathfrak{B}_2$ be the homomorphisms from (a) and (b) with $h_0^{-1}(0) = \mathfrak{a}$ and $h_1^{-1}(1) = \mathfrak{u}$. We define

$$h(x) := \langle h_0(x), h_1(x) \rangle .$$

Since $\mathfrak{a} \cap \mathfrak{u} = \varnothing$ it follows that $h^{-1}(\langle 0, 0 \rangle) = \mathfrak{a}$ and $h^{-1}(\langle 1, 1 \rangle) = \mathfrak{u}$. Furthermore, h is a homomorphism since

$$\begin{aligned} h(x) \sqcup h(y) &= \big\langle h_0(x), h_1(x) \big\rangle \sqcup \big\langle h_0(y), h_1(y) \big\rangle \\ &= \big\langle h_0(x) \sqcup h_0(y),\ h_1(x) \sqcup h_1(y) \big\rangle = h(x \sqcup y) , \end{aligned}$$

and similarly for $\sqcap$. □

Corollary 4.6. *Let $\mathfrak{L}$ be a lattice. A subset $X \subseteq L$ is a prime ideal if and only if $X = h^{-1}(0)$ for some surjective homomorphism $h : \mathfrak{L} \to \mathfrak{B}_2$.*

The prime ideals in distributive and boolean lattices are especially well-behaved. We will show that for these lattices every maximal ideal is prime and that, for boolean lattices, the converse also holds. Note that in general there may be non-prime maximal ideals. For instance, the lattice $\mathfrak{M}_5$ has three maximal ideals none of which is prime.

Theorem 4.7. *Let $\mathfrak{L}$ be a distributive lattice, $\mathfrak{a}$ an ideal, and $\mathfrak{u}$ a filter with $\mathfrak{a} \cap \mathfrak{u} = \varnothing$. There exists a maximal ideal $\mathfrak{b} \supseteq \mathfrak{a}$ with $\mathfrak{b} \cap \mathfrak{u} = \varnothing$ and this ideal is prime.*

Proof. The existence of $\mathfrak{b}$ was already proved in Corollary 3.5. It remains to show that it is prime. Suppose otherwise. Then there are elements $x, y \in L \setminus \mathfrak{b}$ with $x \sqcap y \in \mathfrak{b}$. By maximality of $\mathfrak{b}$, it follows that

$$\mathrm{cl}_{\downarrow}(\mathfrak{b} \cup \{x\}) \cap \mathfrak{u} \neq \varnothing \quad \text{and} \quad \mathrm{cl}_{\downarrow}(\mathfrak{b} \cup \{y\}) \cap \mathfrak{u} \neq \varnothing .$$

Therefore, there are elements $a, b \in \mathfrak{b}$ with $a \sqcup x \in \mathfrak{u}$ and $b \sqcup y \in \mathfrak{u}$. Consequently,

$$z := (a \sqcup x) \sqcap (b \sqcup y) \in \mathfrak{u} .$$

On the other hand, by distributivity we have

$$z = \underbrace{(a \sqcap b)}_{\in \mathfrak{b}} \sqcup \underbrace{(a \sqcap y)}_{\in \mathfrak{b}} \sqcup \underbrace{(x \sqcap b)}_{\in \mathfrak{b}} \sqcup \underbrace{(x \sqcap y)}_{\in \mathfrak{b}} .$$

Thus, $z \in \mathfrak{b} \cap \mathfrak{u} \neq \varnothing$. Contradiction. □

Corollary 4.8. *Every maximal ideal in a distributive lattice is prime.*

As a consequence of Theorem 4.7 we obtain a simple condition for the existence of ultrafilters containing given elements.

Definition 4.9. A set $X \subseteq L$ has the *finite intersection property* if

$$\bigsqcap X_0 \neq \bot , \quad \text{for all finite } X_0 \subseteq X .$$

If L has no least element then every subset has the finite intersection property.

Corollary 4.10. *Let $\mathfrak{L}$ be a bounded distributive lattice and $X \subseteq L$. There exists an ultrafilter $\mathfrak{u} \supseteq X$ if, and only if, X has the finite intersection property.*

Proof. X has the finite intersection property if and only if $\bot \notin \mathrm{cl}_\uparrow(X)$. By (the dual of) Theorem 4.7, $\bot \notin \mathrm{cl}_\uparrow(X)$ implies that there exists an ultrafilter $\mathfrak{u} \supseteq \mathrm{cl}_\uparrow(X)$. □

In boolean lattices the structure of the prime ideals is especially simple.

Theorem 4.11. *Let $\mathfrak{B}$ be a boolean lattice and $\mathfrak{a} \subseteq B$ an ideal. The following statements are equivalent:*

(1) $\mathfrak{a}$ *is maximal.*

(2) $\mathfrak{a}$ *is prime.*

(3) *For every $x \in B$, we have either $x \in \mathfrak{a}$ or $x^* \in \mathfrak{a}$.*

Proof. (1) ⇒ (2) was shown in Corollary 4.8.

(2) ⇒ (3) We have $x \sqcap x^* = \bot \in \mathfrak{a}$. Since $\mathfrak{a}$ is a prime ideal it follows that $x \in \mathfrak{a}$ or $x^* \in \mathfrak{a}$. Clearly, we cannot have both since, otherwise, $\top = x \sqcup x^* \in \mathfrak{a}$ and $\mathfrak{a}$ would not be proper.

(3) ⇒ (1) Let $\mathfrak{b} \supset \mathfrak{a}$ be an ideal. We have to show that $\mathfrak{b}$ is nonproper. Fix some $x \in \mathfrak{b} \setminus \mathfrak{a}$. By assumption, $x^* \in \mathfrak{a} \subseteq \mathfrak{b}$. Hence, $\top = x \sqcup x^* \in \mathfrak{b}$ and $\mathfrak{b} = B$ is nonproper. □

Corollary 4.12. *A bounded distributive lattice* $\mathfrak{L}$ *is boolean if, and only if, there are no prime ideals* $\mathfrak{a}, \mathfrak{b}$ *with* $\mathfrak{a} \subset \mathfrak{b}$.

Proof. (⇒) By Theorem 4.11, every prime ideal is maximal.

(⇐) We have to show that every element $a \in L$ has a complement a^*. Suppose that some element a has none. The sets

$$\mathfrak{u} := \{\, b \in L \mid a \sqcup b = \top \,\},$$
$$\mathfrak{v} := \{\, b \in L \mid b \sqsupseteq a \sqcap d \text{ for some } d \in \mathfrak{u} \,\}$$

are filters. If $\bot \in \mathfrak{v}$ then $\bot = a \sqcap d$ for some d with $a \sqcup d = \top$, and d would be a complement of a. Consequently, $\mathfrak{v}$ is proper. By Theorem 4.7 it follows that there exists a prime ideal $\mathfrak{a}$ with $\mathfrak{a} \cap \mathfrak{v} = \varnothing$. The ideal

$$\mathfrak{b} := \{\, b \in L \mid b \sqsubseteq a \sqcup c \text{ for some } c \in \mathfrak{a} \,\}$$

is proper since $\top = a \sqcup c$, for some $c \in \mathfrak{a}$ would imply that $c \in \mathfrak{a} \cap \mathfrak{u} \neq \varnothing$. Choose some prime ideal $\mathfrak{c} \supseteq \mathfrak{b}$. Since $\mathfrak{b} \supset \mathfrak{a}$ we have found two comparable prime ideals $\mathfrak{a} \subset \mathfrak{c}$. Contradiction. □

Let us compute the number of ultrafilters in a boolean lattice of the form $\langle \mathscr{P}(A), \subseteq \rangle$.

Theorem 4.13. *For every infinite set A there are $2^{2^{|A|}}$ ultrafilters on A.*

Proof. Set $\kappa := |A|$. As every ultrafilter is a subset of $\mathscr{P}(A)$, there are at most $|\mathscr{P}(\mathscr{P}(A))| = 2^{2^\kappa}$ ultrafilters on A. Thus, we only need to prove a lower bound.

We call a family $F \subseteq \wp(A)$ *independent* if every non-trivial finite boolean combination of sets in F has cardinality $|A|$, that is, for all pairwise distinct sets $X_0, \dots, X_{m-1}, Y_0, \dots, Y_{n-1} \in F$, $m, n < \omega$, we have

$$\big|X_0 \cap \dots \cap X_{m-1} \cap (A \setminus Y_0) \cap \dots \cap (A \setminus Y_{n-1})\big| = |A| \,.$$

We will prove below that there exists an independent family $F \subseteq \wp(A)$ of size $|F| = 2^\kappa$. Using such a family F we can construct 2^{2^κ} ultrafilters as follows. For each subset $K \subseteq F$, set

$$S_K := K \cup \{\, A \setminus X \mid X \in F \setminus K \,\} \,.$$

Note that $S_K \subseteq F$ has the finite intersection property since F is independent. Therefore, we can use Corollary 4.10 to extend S_K to an ultrafilter $\mathfrak{u}_K \supseteq S_K$.

Since $|\wp(F)| = 2^{|F|} = 2^{2^\kappa}$, it remains to prove that $\mathfrak{u}_K \neq \mathfrak{u}_L$ for $K \neq L$. Thus, let $K \neq L$. By symmetry, we may assume that there is some set $X \in K \setminus L$. Then $X \in S_K \subseteq \mathfrak{u}_K$ and $A \setminus X \in S_L \subseteq \mathfrak{u}_L$. Consequently, $\mathfrak{u}_K \neq \mathfrak{u}_L$.

It remains to construct the desired family $F \subseteq \wp(A)$. Let W be the set of all pairs $\langle B, H\rangle$ where $B \subseteq A$ is finite and H is a finite set of finite subsets of A. Then $|W| = |A|^{<\aleph_0} \otimes (|A|^{<\aleph_0})^{<\aleph_0} = |A|$ and there exists a bijection $\varphi : W \to A$. It is sufficient to find an independent family $F \subseteq \wp(W)$ of size 2^κ since we can apply φ to F to obtain the desired subsets of $\wp(A)$. For $s \subseteq A$, let

$$P_s := \{\, \langle B, H\rangle \in W \mid B \cap s \in H \,\} \,.$$

We claim that

$$F := \{\, P_s \mid s \subseteq A \,\}$$

is the desired independent family.

To show that it has size 2^κ, consider distinct subsets $s, t \subseteq A$. By symmetry we may assume that $s \nsubseteq t$. Fixing some element $a \in s \setminus t$, it follows that

$$\big\langle \{a\}, \{\{a\}\} \big\rangle \in P_s \setminus P_t \,, \quad \text{which implies that} \quad P_s \neq P_t \,.$$

To show that F is independent, let $s_0, \dots, s_{m-1}, t_0, \dots, t_{n-1} \subseteq A$ be pairwise distinct. For every pair $\langle i, k \rangle \in [m] \times [n]$, we fix some element

$$a_{ik} \in (s_i \setminus t_k) \cup (t_k \setminus s_i) .$$

Let Q be the set of all finite subsets of A that contain all chosen elements a_{ik}, for $i < m$, $k < n$. By choice of a_{ik} we have

$$B \cap s_i \neq B \cap t_k , \quad \text{for all } B \in Q .$$

Setting $H_B := \{ B \cap s_i \mid i < m \}$ this implies that

$$\langle B, H_B \rangle \in P_{s_i} \quad \text{and} \quad \langle B, H_B \rangle \notin P_{t_k} , \quad \text{for all } i < m \text{ and } k < n .$$

Consequently,

$$\langle B, H_B \rangle \in P_{s_0} \cap \dots \cap P_{s_{m-1}} \cap (W \setminus P_{t_0}) \cap \dots \cap (W \setminus P_{t_{n-1}}) ,$$

for all $B \in Q$. This implies that

$$\begin{aligned} &\left| P_{s_0} \cap \dots \cap P_{s_{m-1}} \cap (W \setminus P_{t_0}) \cap \dots \cap (W \setminus P_{t_{n-1}}) \right| \\ &\quad \geq |Q| = \kappa = |W| . \end{aligned}$$ □

Exercise 4.1. How many ultrafilters are there on a finite set A?

We conclude this section with a result stating that ultrafilters of a subalgebra have several extensions to ultrafilters of the whole algebra.

Proposition 4.14. *Let $\mathfrak{A} \subseteq \mathfrak{B}$ be boolean algebras. If, for every ultrafilter $\mathfrak{u}$ of $\mathfrak{A}$, there exists a unique ultrafilter $\mathfrak{v}$ of $\mathfrak{B}$ with $\mathfrak{u} \subseteq \mathfrak{v}$, then $\mathfrak{A} = \mathfrak{B}$.*

Proof. Let $\mathfrak{A} \subseteq \mathfrak{B}$ be boolean algebras such that every ultrafilter of $\mathfrak{A}$ can be extended to a unique ultrafilter of $\mathfrak{B}$. Consider some element $b \in B$. In order to show that $b \in A$, we prove the following statements.

(1) For every ultrafilter $\mathfrak{v}$ of $\mathfrak{B}$ with $A \cap \Uparrow b \subseteq \mathfrak{v}$, the set $(\mathfrak{v} \cap A) \cup \{b\}$ has the finite intersection property.

(2) There is no ultrafilter $\mathfrak{v}$ of $\mathfrak{B}$ containing $A \cap \Uparrow b$ and b^*.

(3) There is some element $a \in A \cap \Uparrow b$ with $a \sqsubseteq b$.

Note that the proposition follows from (3) since $a \in \Uparrow b$ implies $b \sqsubseteq a$. Hence, $b = a \in A$. It remains to prove the claims.

(1) For a contradiction, suppose that there is some ultrafilter $\mathfrak{v}$ such that $A \cap \Uparrow b \subseteq \mathfrak{v}$, but $(\mathfrak{v} \cap A) \cup \{b\}$ does not have the finite intersection property. Since $\mathfrak{v} \cap A$ is closed under the infimum operation $\sqcap$, it follows that there is some element $a \in \mathfrak{v} \cap A$ such that $a \sqcap b = \bot$. Hence, $b \sqsubseteq a^*$, which implies that $a^* \in A \cap \Uparrow b \subseteq \mathfrak{v}$ and $\bot = a \sqcap a^* \in \mathfrak{v}$. A contradiction.

(2) For a contradiction, suppose that there is some ultrafilter $\mathfrak{v}$ of $\mathfrak{B}$ with $(A \cap \Uparrow b) \cup \{b^*\} \subseteq \mathfrak{v}$. By (1) and Corollary 4.10, there is some ultrafilter $\mathfrak{v}'$ containing $(\mathfrak{v} \cap A) \cup \{b\}$. By assumption, $\mathfrak{v}' \cap A = \mathfrak{v} \cap A$ implies $\mathfrak{v}' = \mathfrak{v}$. But $b \in \mathfrak{v}'$ while $b^* \in \mathfrak{v}$. A contradiction.

(3) According to (2) there is no ultrafilter containing $(A \cap \Uparrow b) \cup \{b^*\}$. By Corollary 4.10, it follows that this set does not have the finite intersection property. Since $A \cap \Uparrow b$ is closed under the infimum operation $\sqcap$, we can therefore find an element $a \in A \cap \Uparrow b$ such that $a \sqcap b^* = \bot$. Consequently, $a \sqsubseteq b$. □

5. *Atomic lattices and partition rank*

In this section we take a closer look at those elements of a lattice that are near to the bottom. The distance of an element from $\bot$ can be measured in different ways. A simple but coarse measure is the *height* of an element.

Definition 5.1. Let $\mathfrak{L}$ be a lattice.

(a) The *height* of an element $a \in L$ is

$$\mathrm{ht}(a) := \sup \{ |C| \mid C \subseteq \downarrow a \text{ is a chain} \} .$$

Elements of height 1 are called *atoms*.

(b) $\mathfrak{L}$ is *atomless* if it has no atoms. It is *atomic* if $\Downarrow a$ contains an atom, for every element $a \neq \bot$.

Example. Let $\mathfrak{V}$ be a vector space and let $\mathfrak{L}$ be the set of all linear subspaces of $\mathfrak{V}$. Note that $\mathfrak{L}$ consists of all fixed points of the closure operator mapping a set $X \subseteq V$ to the subspace spanned by X. Hence, $\mathfrak{L}$ forms a complete lattice where $U \sqcap W = U \cap W$ and $U \sqcup W = U \oplus W$ is the subspace spanned by $U \cup W$. This lattice is atomic. The height of an element $U \in L$ coincides with its dimension.

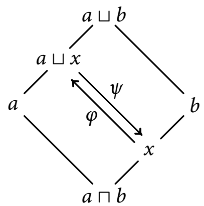
The notion of height is mainly meaningful for modular lattices where it is well-behaved, at least for elements of finite height.

Lemma 5.2. *Let $\mathfrak{L}$ be a modular lattice and $a, b \in L$. The function*

$$\varphi : [a \sqcap b,\ b] \to [a,\ a \sqcup b] : x \mapsto a \sqcup x$$

is strictly increasing and surjective. Its inverse is given by the function

$$\varphi : [a,\ a \sqcup b] \to [a \sqcap b,\ b] : x \mapsto b \sqcap x\,.$$



Proof. Clearly, φ and ψ are increasing and we have $\operatorname{rng}\varphi \subseteq \Uparrow a$ and $\operatorname{rng}\psi \subseteq \Downarrow b$. Furthermore, $x \sqsubseteq b \sqsubseteq a \sqcup b$ implies that $\varphi(x) = a \sqcup x \sqsubseteq a \sqcup b$. Hence, $\operatorname{rng}\varphi \subseteq \Downarrow(a \sqcup b)$. Similarly, it follows that $\operatorname{rng}\psi \subseteq \Uparrow(a \sqcap b)$.

It remains to show that ψ is the inverse of φ. Note that if $\mathfrak{L}$ is modular then so is $\mathfrak{L}^{\mathrm{op}}$. It is therefore sufficient to prove that $\varphi \circ \psi = \mathrm{id}$, the equation $\psi \circ \varphi = \mathrm{id}$ then follows by duality. For $a \sqsubseteq x \sqsubseteq a \sqcup b$, modularity implies that

$$\varphi(\psi(x)) = a \sqcup (b \sqcap x) = x \sqcap (a \sqcup b) = x\,,$$

as desired. □

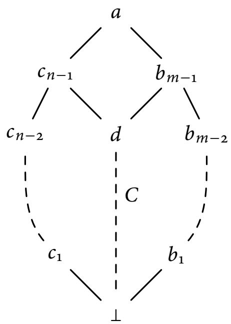


Figure 3.. Proof of Lemma 5.3

Lemma 5.3. *Let $\mathfrak{L}$ be a modular lattice and $a \in L$ an element of height $n < \aleph_0$. Every maximal chain in $\Downarrow a$ has size $n + 1$.*

Proof. We prove by induction on n that, if $b_0 \sqsubset \cdots \sqsubset b_m$ is a maximal chain with $b_m = a$, then $m = n$. Since a has height n, there exists a chain $c_0 \sqsubset \cdots \sqsubset c_n$ of size $n + 1$ with $c_0 = \bot$ and $c_n = a$. If $b_{m-1} = c_{n-1}$ then the claim follows by inductive hypothesis. Suppose that $b_{m-1} \neq c_{n-1}$. Set $d := b_{m-1} \sqcap c_{n-1}$ and let $C \subseteq \Downarrow d$ be a maximal chain. Then $|C| = \text{ht}(d) + 1 < \text{ht}(c_{n-1}) + 1 = n$.

By Lemma 5.2, there is no element x with $d \sqsubset x \sqsubset c_{n-1}$ because, otherwise, we would have $c_{n-1} \sqsubset c_{n-1} \sqcup x \sqsubset c_n$ in contradiction to the minimality of n. Consequently, $C \cup \{c_{n-1}\}$ is a maximal chain in $\Downarrow c_{n-1}$ and, by inductive hypothesis, it follows that $|C| + 1 = n$.

Similarly, there is no element x with $d \sqsubset x \sqsubset b_{m-1}$. Hence, $C \cup \{b_{m-1}\}$ is a maximal chain in $\Downarrow b_{m-1}$ and we have $|C| + 1 = m$. It follows that $m = |C| + 1 = n$, as desired. □

Example. For infinite heights the lemma fails. Consider the real interval $I := [0, 1]$ and its subset $K := I \cap \mathbb{Q}$. We order the product $L := I \times K$ by $(a, b) \leq (c, d)$ iff $a \leq b$ and $c \leq d$. Then L is a modular lattice with maximal chains

$$C := (\{0\} \times K) \cup (I \times \{1\}) \quad \text{and} \quad C' := \{ (x, x) \mid x \in K \} .$$

But $|C| = 2^{\aleph_0}$ while $|C'| = \aleph_0$.

Lemma 5.4. *Let $\mathfrak{L}$ be a modular lattice and $a \sqsubseteq b$ elements of finite height. The size of a maximal chain $C \subseteq [a, b]$ is* $\mathrm{ht}(b) - \mathrm{ht}(a) + 1$.

Proof. Every chain in $C \subseteq [a, b]$ can be extended to a chain in $\Downarrow b$ of size $|C| + \mathrm{ht}(a)$. Therefore, the size of such chains is bounded by $\mathrm{ht}(b) - \mathrm{ht}(a) + 1$. Conversely, fix maximal chains $C' \subseteq [a, b]$ and $C'' \subseteq [\bot, a]$. Then $C' \cup C''$ is also maximal. By Lemma 5.3, it follows that $|C' \cup C''| = \mathrm{ht}(b) + 1$. Since $|C''| = \mathrm{ht}(a) + 1$ and $C' \cap C'' = \{a\}$ it follows that $|C'| = \mathrm{ht}(b) - \mathrm{ht}(a) + 1$. □

Theorem 5.5. *Let $\mathfrak{L}$ be a modular lattice. If $a, b \in L$ are elements with* $\mathrm{ht}(a \sqcup b) < \aleph_0$ *then*

$$\mathrm{ht}(a) + \mathrm{ht}(b) = \mathrm{ht}(a \sqcup b) + \mathrm{ht}(a \sqcap b) .$$

Proof. Set $I_0 := [a \sqcap b, a]$ and $I_1 := [b, a \sqcup b]$. The partial orders $\mathfrak{J}_0 := \langle I_0, \sqsubseteq \rangle$ and $\mathfrak{J}_1 := \langle I_1, \sqsubseteq \rangle$ are modular lattices and, by Lemma 5.2, there exists an isomorphism $\varphi : \mathfrak{J}_0 \to \mathfrak{J}_1$. By Lemma 5.4, the height of the top element of $\mathfrak{J}_0$ is $\mathrm{ht}(a) - \mathrm{ht}(a \sqcap b) + 1$ and the height of the top element of $\mathfrak{J}_1$ is $\mathrm{ht}(a \sqcup b) - \mathrm{ht}(b) + 1$. Since $\mathfrak{J}_0 \cong \mathfrak{J}_1$ it follows that

$$\mathrm{ht}(a) - \mathrm{ht}(a \sqcap b) + 1 = \mathrm{ht}(a \sqcup b) - \mathrm{ht}(b) + 1 .$$ □

Remark. The above equation is called the *modular law.* It can be used to characterise modular lattices. If $\mathfrak{L}$ is a lattice where every element has finite height then $\mathfrak{L}$ is modular if and only if every pair $a, b \in L$ of elements satisfies the modular law.

Example. For the subspace lattice of a vector space, we obtain the well-known dimension formula:

$$\dim U + \dim W = \dim(U \cap W) + \dim(U \oplus W) .$$

For boolean algebras the structure of the elements of finite height is especially simple.

Lemma 5.6. *Let $\mathfrak{B}$ be a boolean algebra. If $b \sqsubset c$ are elements of finite height then there exists an atom $a \in \Downarrow c \setminus \Downarrow b$.*

Proof. Let $b' := c \sqcap b^*$. Since c has finite height there exists a finite chain $C \subseteq \Downarrow b'$ of maximal size. This chain contains an atom a. Note that $a \sqsubseteq b$ would imply $a \sqsubseteq b \sqcap b' = \bot$ which is impossible since a is an atom. Hence, $a \in \Downarrow c \setminus \Downarrow b$. □

Lemma 5.7. *Let $\mathfrak{B}$ be a boolean algebra and $a \in B$ an element of height $n < \aleph_0$. Then there are exactly n atoms in $\Downarrow a$.*

Proof. By Lemma 5.6, if $c_0 \sqsubset \dots \sqsubset c_n$ is a chain of length $n+1$ with $c_n = a$ then there are at least n atoms below c_n. Conversely, suppose that $b_0, \dots, b_{n-1} \in \Downarrow a$ are atoms. Set $c_0 := \bot$ and $c_{i+1} := c_i \sqcup b_i$. Then $c_0 \sqsubset \dots \sqsubset c_n$ forms a chain of length $n+1$ in $\Downarrow a$. Consequently, the height of a is at least n. □

Corollary 5.8. *Let $\mathfrak{B}$ be a boolean algebra. Every element $a \in B$ with finite height is the supremum of finitely many atoms.*

Proof. Let P be the set of all atoms in $\Downarrow a$. It is sufficient to show that $a = \sup P$. Suppose otherwise. Then $c := \sup P \sqsubset a$. By Lemma 5.6, there exists an atom $b \in \Downarrow a \setminus \Downarrow c$. By definition of P, it follows that $b \in P$. But $b \not\sqsubseteq c = \sup P$. Contradiction. □

Example. The previous lemma cannot be generalised to infinite heights. Let A be an uncountable set and define

$$F := \{ X \subseteq A \mid X \text{ or } A \setminus X \text{ is finite} \}.$$

Then $\langle F, \subseteq \rangle$ is a boolean algebra and we have

$$\operatorname{ht}(X) = \begin{cases} |X| & \text{if } X \text{ is finite}, \\ \aleph_0 & \text{otherwise}. \end{cases}$$

But every infinite set $X \in F$ is uncountable. Hence, there are uncountably many atoms below X.

Let us introduce a second measure of the distance between an element and $\perp$ that allows a finer classification of elements of infinite height. Basically, instead of considering all chains in $\Downarrow a$ we only look at strictly decreasing sequences.

Definition 5.9. Let $\mathfrak{L}$ be a lattice with least element $\perp$.

(a) A *partition* of an element $a \in L$ is a set $P \subseteq \Downarrow a$ with $\perp \notin P$ such that $p \sqcap q = \perp$, for all $p, q \in P$ with $p \neq q$.

(b) The *partition rank* of an element $a \in L$ is defined as follows:

- $\mathrm{rk}_{\mathrm{P}}(a) = -1$ iff $a = \perp$.
- $\mathrm{rk}_{\mathrm{P}}(a) \geq 0$ iff $a \neq \perp$.
- $\mathrm{rk}_{\mathrm{P}}(a) \geq \alpha + 1$ iff there exists an infinite partition P of a such that $\mathrm{rk}_{\mathrm{P}}(p) \geq \alpha$, for all $p \in P$.
- For limit ordinals δ, we set $\mathrm{rk}_{\mathrm{P}}(a) \geq \delta$ iff $\mathrm{rk}_{\mathrm{P}}(a) \geq \alpha$, for all $\alpha < \delta$.

Exercise 5.1. Let $\mathfrak{B}$ be a boolean algebra and $a \in B$ an element of height $0 < \mathrm{ht}(a) < \aleph_0$. Show that $\mathrm{rk}_{\mathrm{P}}(a) = 1$.

Lemma 5.10. $a \sqsubseteq b$ *implies* $\mathrm{rk}_{\mathrm{P}}(a) \leq \mathrm{rk}_{\mathrm{P}}(b)$.

Lemma 5.11. *If $\mathfrak{L}$ is a distributive lattice then*

$$\mathrm{rk}_{\mathrm{P}}(a \sqcup b) = \max\{\mathrm{rk}_{\mathrm{P}}(a), \mathrm{rk}_{\mathrm{P}}(b)\}.$$

Proof. By the preceding lemma, we have $\mathrm{rk}_{\mathrm{P}}(a \sqcup b) \geq \mathrm{rk}_{\mathrm{P}}(a), \mathrm{rk}_{\mathrm{P}}(b)$. It remains to show that $\mathrm{rk}_{\mathrm{P}}(a \sqcup b) \geq \alpha$ implies $\mathrm{rk}_{\mathrm{P}}(a) \geq \alpha$ or $\mathrm{rk}_{\mathrm{P}}(b) \geq \alpha$. We proceed by induction on α.

If $\alpha = -1$ then $a \sqcup b = \perp$ implies $a = \perp$ and $b = \perp$. For limit ordinals α, there is nothing to do. Suppose that $\mathrm{rk}_{\mathrm{P}}(a \sqcup b) \geq \alpha + 1$. Then there exists an infinite partition P of $a \sqcup b$ such that $\mathrm{rk}_{\mathrm{P}}(p) \geq \alpha$, for all $p \in P$. For $p \in P$, set $a_p := a \sqcap p$ and $b_p := b \sqcap p$. Then

$$a_p \sqcup b_p = (a \sqcap p) \sqcup (b \sqcap p) = (a \sqcup b) \sqcap p = p.$$

By inductive hypothesis, we know that

$$\mathrm{rk_P}(a_p \sqcup b_p) = \mathrm{rk_P}(p) \geq \alpha$$

implies that $\mathrm{rk_P}(a_p) \geq \alpha$ or $\mathrm{rk_P}(b_p) \geq \alpha$. Set

$$P_a := \{ p \in P \mid \mathrm{rk_P}(a_p) \geq \alpha \}$$
$$\text{and} \quad P_b := \{ p \in P \mid \mathrm{rk_P}(b_p) \geq \alpha \} .$$

Then $P_a \cup P_b = P$ and at least one of the sets is infinite. By symmetry, let us assume that P_a is infinite. Then P_a is an infinite partition of a with $\mathrm{rk_P}(q) \geq \alpha$, for all $q \in P_a$. Consequently, $\mathrm{rk_P}(a) \geq \alpha + 1$. □

Lemma 5.12. *Let $h : \mathfrak{A} \to \mathfrak{B}$ be an injective homomorphism between boolean algebras. Then*

$$\mathrm{rk_P}(a) \leq \mathrm{rk_P}(h(a)) , \quad \textit{for all } a \in A .$$

Proof. If $\mathfrak{A} \subseteq \mathfrak{B}$ then it follows immediately from the definition that the rank of an element $a \in A$ in $\mathfrak{A}$ is less than or equal to its rank in $\mathfrak{B}$. Therefore, it is sufficient to prove that every injective homomorphism between boolean algebras is an embedding.

Suppose that $h(a) \leq h(b)$. Then $\bot = h(a) \sqcap h(b)^* = h(a \sqcap b^*)$. Since h is injective it follows that $a \sqcap b^* = \bot$. Hence, $a \leq b$. □

As usual for ranks defined by inductive fixed points the maximal non-infinite rank is bounded by the cardinality of the underlying set.

Lemma 5.13. *Let $\mathfrak{L}$ be a lattice.* $\mathrm{rk_P}(a) \geq |L|^+$ *implies that* $\mathrm{rk_P}(a) = \infty$.

Proof. Let $\kappa := |L|$ and set $X_\alpha := \{ a \in L \mid \mathrm{rk_P}(a) \geq \alpha \}$. Then $X_\alpha \supseteq X_\beta$, for $\alpha \leq \beta$. Consequently, there is some $\alpha < \kappa^+$ such that $X_\alpha = X_{\alpha+1}$. This implies that $X_\alpha = X_{\kappa^+} = X_\infty$. □

The next lemma shows that it is possible to split elements of infinite rank into an arbitrary number of elements whose rank is again infinite. This will be useful to prove the existence of many different ultrafilters in Corollary B5.7.4 below.

Definition 5.14. Let $\mathfrak{L}$ be a lattice with least element $\perp$, and let κ be a cardinal and α an ordinal. An *embedding* of the tree $\kappa^{<\alpha}$ is a family $(a_w)_{w \in \kappa^{<\alpha}}$ of elements $a_w \in L$ such that

$$\begin{aligned} \perp \sqsubset a_w \sqsubset a_u \quad & \text{for all } u < w\,, \\ a_u \sqcap a_w = \perp \quad & \text{for all } u, w \text{ with } u \nleq w \text{ and } w \nleq u\,. \end{aligned}$$

(Note that the ordering is reversed.)

Lemma 5.15. *Let $\mathfrak{L}$ be a lattice and $a \in L$. The following statements are equivalent:*

(1) $\mathrm{rk_P}(a) = \infty$.

(2) *There exists an embedding $(b_w)_{w \in 2^{<\omega}}$ of $2^{<\omega}$ into $\mathfrak{L}$ with $b_{\langle\rangle} = a$.*

(3) *There exists an embedding $(b_w)_{w \in \aleph_0^{<\omega}}$ of $\aleph_0^{<\omega}$ into $\mathfrak{L}$ with $b_{\langle\rangle} = a$.*

Proof. (3) $\Rightarrow$ (2) is trivial.

(1) $\Rightarrow$ (3) Let $\kappa := |L|^+$. We construct the family $(b_w)_w$ by induction on w such that $\mathrm{rk_P}(b_w) = \infty$. We start with $b_{\langle\rangle} = a$. If b_w has been defined then $\mathrm{rk_P}(b_w) \geq \kappa + 1$ implies that there exists an infinite partition P of b_w with $\mathrm{rk_P}(p) \geq \kappa$, for all $p \in P$. By Lemma 5.13, it follows that $\mathrm{rk_P}(p) = \infty$, for each $p \in P$. Select distinct elements $b_{wk} \in P$, for $k < \omega$. Then we have $b_{wk} \sqcap b_{wn} = \perp$ for $k \neq n$ and $\mathrm{rk_P}(b_{wi}) = \infty$, as desired.

(2) $\Rightarrow$ (1) Let $(b_w)_w$ be an embedding of $2^{<\omega}$ into $\mathfrak{L}$ with $b_{\langle\rangle} = a$. By induction on α, we prove that $\mathrm{rk_P}(b_w) \geq \alpha$, for all w. Since $b_{w0} \sqsubset b_w$ we have $b_w \neq \perp$ and $\mathrm{rk_P}(b_w) \geq 0$. For limit ordinals, the claim follows immediately from the inductive hypothesis. Hence, it remains to consider the successor step. Suppose that $\mathrm{rk_P}(b_w) \geq \alpha$, for all w. The set $\{\, b_{w0^n1} \mid n < \omega \,\}$ is an infinite partition of b_w where each element has rank at least α. Therefore, $\mathrm{rk_P}(b_w) \geq \alpha + 1$. □

In contrast to the preceding result, it turns out that we can split elements of non-infinite rank only a finite number of times into elements of the same rank.

Lemma 5.16. *Let $\mathfrak{B}$ be a boolean algebra. For every element $a \in B$ with $\mathrm{rk}_{\mathrm{P}}(a) < \infty$, there exists a finite partition P of a such that*

$$a = \sup P \quad \textit{and} \quad \mathrm{rk}_{\mathrm{P}}(p) = \mathrm{rk}_{\mathrm{P}}(a)\,, \qquad \textit{for all } p \in P\,.$$

Furthermore, if Q is any other partition of a with

$$\mathrm{rk}_{\mathrm{P}}(q) = \mathrm{rk}_{\mathrm{P}}(a)\,, \qquad \textit{for all } q \in Q\,,$$

then $|Q| \leq |P|$.

Proof. Let $\alpha := \mathrm{rk}_{\mathrm{P}}(a)$. To find P we construct a tree $T \subseteq 2^{<\omega}$ and elements $b_w \in B$, for $w \in T$, with $\mathrm{rk}_{\mathrm{P}}(b_w) = \alpha$ as follows. We start with $b_{\langle\rangle} := a$. If b_w is already defined and there is some element $c \in B$ such that $\mathrm{rk}_{\mathrm{P}}(b_w \sqcap c) = \alpha$ and $\mathrm{rk}_{\mathrm{P}}(b_w \sqcap c^*) = \alpha$, then we add $w0$ and $w1$ to T and we set $b_{w0} := b_w \sqcap c$ and $b_{w1} := b_w \sqcap c^*$. Otherwise, w becomes a leaf of T.

We claim that any such tree T is finite. For a contradiction, suppose there exists an infinite tree T as above. Since T is binary it contains an infinite path $\beta \in 2^\omega$, by Lemma 1.9. Let $w_n := \beta \upharpoonright n$ be the prefix of β of length n. For $n < \omega$, set $c_n := b_{w_n} \sqcap b^*_{w_{n+1}}$. Then we have $c_n \sqsubseteq a$ and $\mathrm{rk}_{\mathrm{P}}(c_n) = \alpha$. Furthermore, $b_{w_n} \sqsubseteq b_{w_{k+1}}$, for $k < n$, implies that

$$c_k \sqcap c_n = b_{w_k} \sqcap b^*_{w_{k+1}} \sqcap b_{w_n} \sqcap b^*_{w_{n+1}} = \bot\,.$$

Consequently, $\mathrm{rk}_{\mathrm{P}}(a) \geq \alpha$. Contradiction.

Let T be a tree as above and let $P \subseteq T$ be the set of its leaves. Set $m := |P|$ and let $p_0, \dots, p_{m-1}$ be an enumeration of P. Then $\mathrm{rk}_{\mathrm{P}}(p_n) = \alpha$, $p_k \sqcap p_n = \bot$, for $k \neq n$, and $a = p_0 \sqcup \dots \sqcup p_{m-1}$.

Let Q be another partition of a with $\mathrm{rk}_{\mathrm{P}}(q) = \alpha$, for $q \in Q$. We claim that $n \leq m$. By construction of P, there exists, for every $p \in P$, at most one $q \in Q$ with $\mathrm{rk}_{\mathrm{P}}(p \sqcap q) = \alpha$. Hence, if $n > m$ then we can find some element $q \in Q$ such that $\mathrm{rk}_{\mathrm{P}}(p \sqcap q) < \alpha$, for all $p \in P$. But

$$q = (q \sqcap p_0) \sqcup \dots \sqcup (q \sqcap p_{n-1})$$

implies, by Lemma 5.11, that $\mathrm{rk}_{\mathrm{P}}(q) < \alpha$. Contradiction. □

Definition 5.17. Let $\mathfrak{B}$ be a boolean algebra.

(a) Let $a \in B$ be an element with $\mathrm{rk_P}(a) < \infty$. The *partition degree* $\deg_P(a)$ of a is the maximal cardinality of a partition P of a with $\mathrm{rk_P}(p) = \mathrm{rk_P}(a)$, for all $p \in P$. If $\mathrm{rk_P}(a) = \infty$ then we set $\deg_P(a) := \infty$.

(b) Let $\mathfrak{u}$ be an ultrafilter of $\mathfrak{B}$. The *partition rank* of $\mathfrak{u}$ is

$$\mathrm{rk_P}(\mathfrak{u}) := \min \{ \mathrm{rk_P}(a) \mid a \in \mathfrak{u} \},$$

and its partition degree is

$$\deg_P(\mathfrak{u}) := \min \{ \deg_P(a) \mid a \in \mathfrak{u} \text{ with } \mathrm{rk_P}(a) = \mathrm{rk_P}(\mathfrak{u}) \}.$$

We say that an element $a \in \mathfrak{u}$ has *minimal rank and degree* if

$$\mathrm{rk_P}(a) = \mathrm{rk_P}(\mathfrak{u}) \quad \text{and} \quad \deg_P(a) = \deg_P(\mathfrak{u}).$$

Example. Let A be a set and $\mathfrak{F} := \langle F, \subseteq \rangle$ where

$$F := \{ X \subseteq A \mid X \text{ or } A \setminus X \text{ is finite} \}.$$

For $X \in F$, we have

$$\mathrm{rk_P}(X) = \begin{cases} 0 & \text{if } X \text{ is finite}, \\ 1 & \text{otherwise}. \end{cases}$$

and

$$\deg_P(X) = \begin{cases} |X| & \text{if } X \text{ is finite}, \\ 1 & \text{otherwise}. \end{cases}$$

For the ultrafilters

$$\mathfrak{u}_a := \Uparrow\{a\} \quad \text{and} \quad \mathfrak{u}_\infty := \{ X \subseteq A \mid A \setminus X \text{ is finite} \},$$

we have

$$\begin{aligned} \mathrm{rk_P}(\mathfrak{u}_a) &= 0 & \deg_P(\mathfrak{u}_a) &= 1, \\ \mathrm{rk_P}(\mathfrak{u}_\infty) &= 1 & \deg_P(\mathfrak{u}_\infty) &= 1. \end{aligned}$$

Remark. If P is a maximal partition of a with $\mathrm{rk_P}(p) = \mathrm{rk_P}(a)$, for all $p \in P$, then it follows that $\deg_{\mathrm{P}}(p) = 1$, for every $p \in P$. For a proof, suppose that p is an element with $\deg_{\mathrm{P}}(p) > 1$. Then there is a partition Q of p with $|Q| > 1$ and we could enlarge P by replacing p by Q.

Lemma 5.18. *Let $\mathfrak{B}$ be a boolean algebra and $0 < n < \aleph_0$. An element $a \in B$ has height n if, and only if,* $\mathrm{rk_P}(a) = 0$ *and* $\deg_{\mathrm{P}}(a) = n$.

Proof. If $\mathrm{rk_P}(a) = 0$ then $\Downarrow a$ contains only finitely many atoms since, otherwise, these would form an infinite partition of a. Hence, a has finite height.

Conversely, if $\mathrm{rk_P}(a) > 0$ then there exists an infinite partition P of a such that $\mathrm{rk_P}(p) \geq 0$, for all $p \in P$. For every $p \in P$, there is some atom in $\Downarrow p$. Since $\Downarrow p \cap \Downarrow q = \{\bot\}$, for $p \neq q$ in P, it follows that there are infinitely many atoms below a. By Lemma 5.7, it follows that $\mathrm{ht}(a) \geq \aleph_0$.

Consequently, we have $\mathrm{rk_P}(a) = 0$ if and only if $0 < \mathrm{ht}(a) < \aleph_0$. It remains to prove that $\deg_{\mathrm{P}}(a) = \mathrm{ht}(a)$, for such elements a. We proceed by induction on $n := \mathrm{ht}(a)$. If a is an atom then we have $\deg_{\mathrm{P}}(a) = 1$ since $\{a\}$ and $\varnothing$ are the only partitions of a. For the inductive step, suppose that $n > 1$. Let P be the set of atoms in $\Downarrow a$. Then $|P| = n$ and $a = \sup P$. Furthermore, by inductive hypothesis,

$$P = \{\, b \in \Downarrow a \mid \deg_{\mathrm{P}}(b) = 1 \,\}\,.$$

Let Q be a partition of a such that $|Q| = \deg_{\mathrm{P}}(a)$ and $\mathrm{rk_P}(q) = 0$, for all $q \in Q$. By maximality of $|Q|$ it follows that $\deg_{\mathrm{P}}(q) = 1$, for $q \in Q$. Hence, $Q \subseteq P$, which implies that $Q = P$ and $\deg_{\mathrm{P}}(a) = |P| = n$. ☐

Lemma 5.19. *If $\mathfrak{u}$ is an ultrafilter with* $\mathrm{rk_P}(\mathfrak{u}) < \infty$ *then* $\deg_{\mathrm{P}}(\mathfrak{u}) = 1$.

Proof. Let $a \in \mathfrak{u}$ be an element of minimal rank and degree and let P be a maximal partition of a such that $a = \sup P$ and $\mathrm{rk_P}(p) = \mathrm{rk_P}(a)$, for all $p \in P$. Since $\mathfrak{u}$ is an ultrafilter and P is finite, it follows that $\sup P \in \mathfrak{u}$ implies that $p \in \mathfrak{u}$, for some $p \in P$. By maximality of P we have $\deg_{\mathrm{P}}(p) = 1$. This implies that $\deg_{\mathrm{P}}(\mathfrak{u}) = 1$. ☐

Lemma 5.20. $\mathrm{rk}_\mathrm{P}(a \sqcap c) = \mathrm{rk}_\mathrm{P}(a) = \mathrm{rk}_\mathrm{P}(a \sqcap c^*) < \infty$ *implies that* $\deg_\mathrm{P}(a \sqcap c) < \deg_\mathrm{P}(a)$.

Exercise 5.2. Prove the preceding lemma.

Every ultrafilter of non-infinite partition rank can be characterised by any of its elements of minimal rank and degree.

Proposition 5.21. *Let* $\mathfrak{B}$ *be a boolean algebra and* $\mathfrak{u}, \mathfrak{v}$ *distinct ultrafilters of* $\mathfrak{B}$ *with* $\mathrm{rk}_\mathrm{P}(\mathfrak{u}), \mathrm{rk}_\mathrm{P}(\mathfrak{v}) < \infty$. *If* $a \in \mathfrak{u}$ *and* $b \in \mathfrak{v}$ *are elements of minimal rank and degree then* $a \neq b$.

Proof. Since $\mathfrak{u} \neq \mathfrak{v}$ there is some element $c \in \mathfrak{u} \smallsetminus \mathfrak{v}$. It follows that $a \sqcap c \in \mathfrak{u}$ and

$$\mathrm{rk}_\mathrm{P}(a \sqcap c) \leq \mathrm{rk}_\mathrm{P}(a) = \mathrm{rk}_\mathrm{P}(\mathfrak{u}) .$$

Since a is of minimal rank we therefore have

$$\mathrm{rk}_\mathrm{P}(a \sqcap c) = \mathrm{rk}_\mathrm{P}(a) .$$

Analogously, we can conclude that

$$\mathrm{rk}_\mathrm{P}(b \sqcap c^*) = \mathrm{rk}_\mathrm{P}(b) .$$

If $a = b$ then it would follow that

$$\mathrm{rk}_\mathrm{P}(a \sqcap c) = \mathrm{rk}_\mathrm{P}(a) = \mathrm{rk}_\mathrm{P}(a \sqcap c^*) .$$

This implies that $\deg_\mathrm{P}(a \sqcap c) < \deg_\mathrm{P}(a)$ in contradiction to the minimality of a. □

In particular, the number of such ultrafilters is bounded by the size of the boolean algebra.

Corollary 5.22. *Let* $\mathfrak{B}$ *be a boolean algebra. There are at most* $|B|$ *ultrafilters* $\mathfrak{u} \subseteq B$ *with* $\mathrm{rk}_\mathrm{P}(\mathfrak{u}) < \infty$.

Proof. For every ultrafilter $\mathfrak{u} \subseteq B$, choose an element $a_\mathfrak{u} \in \mathfrak{u}$ of minimal rank and degree. By Proposition 5.21, it follows that $a_\mathfrak{u} \neq a_\mathfrak{v}$, for $\mathfrak{u} \neq \mathfrak{v}$. Consequently, there are at most $|B|$ such ultrafilters. □

B3. *Universal constructions*

1. *Terms and term algebras*

We can compose the operations of a structure to build new operations. In the same way as the signature provides names for the basic operations we can associate a name with each of these derived operation. A canonical way of doing so is to name each operation by a description of how it is build up from the given operations. These canonical names are called *terms.*

Definition 1.1. (a) A *term domain* is an initial segment $T \subseteq \kappa^{<\omega}$ such that, if $\alpha < \beta < \kappa$ then $x\beta \in T$ implies $x\alpha \in T$. In particular, every term domain forms a tree.

(b) A *term* is a function $t : T \to \Lambda$ where T is a term domain and Λ a set of function symbols. The *domain* of t is the set $\operatorname{dom} t := T$. If $t(v) = \lambda$ then we say that v is *labelled* by λ.

(c) Let Σ be a signature and X a set of variables. We denote the set of all function symbols of Σ by Σ_{fun}. A Σ-*term* is a term $t : T \to \Sigma_{\text{fun}} \cup X$ satisfying the following properties:

- All inner vertices $v \in \operatorname{dom} t$ are labelled by elements of Σ_{fun}.
- If the function symbol $t(v) = f \in \Sigma_{\text{fun}}$ is of type $s_0 \dots s_{n-1} \to s'$ then v has exactly n successors $u_0, \dots, u_{n-1}$ and, for all $i < n$, either $t(u_i) \in X_{s_i}$ is a variable of type s_i or $t(u_i) = g \in \Sigma_{\text{fun}}$ is a function symbol of type $\bar{r} \to s_i$, for some $\bar{r}$.

The set of all finite Σ-terms with variables from X is denoted by $T[\Sigma, X]$. By $T_s[\Sigma, X]$ we denote the subset of all terms $t \in T[\Sigma, X]$ whose root is labelled by a function symbol of type $\bar{r} \to s$, for some $\bar{r}$.

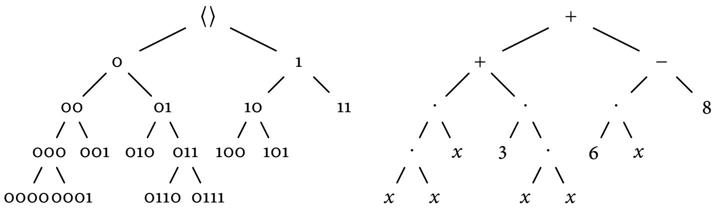


Figure 1.. Domain and labelling of t.

Remark. The difference between a general term and a Σ-term is that the symbols of the former need not to have an arity. In particular, a Σ-term is always finitely branching since, by definition, all symbols in a signature have finite arity.

Example. The polynomial

$$((x \cdot x) \cdot x + 3 \cdot (x \cdot x)) + (6 \cdot x - 8)$$

corresponds to a Σ-term $t : T \to \Sigma$ where $\Sigma = \{\cdot, +, -, 3, 6, 8\}$. (Note that we need to include the coefficients as constant symbols.) The domain T of t and its labelling are shown in Figure 1.

Definition 1.2. Let t be a term and $v \in \operatorname{dom} t$. By t_v we denote the term with domain

$$\operatorname{dom} t_v := \{\, x \mid vx \in \operatorname{dom} t \,\}$$

and labelling

$$t_v(x) := t(vx)\,.$$

A *subterm* of t is a term of the form t_v, for some $v \in \operatorname{dom} t$.

Terms as defined above are cumbersome to write down. Therefore, we represent terms $t \in T[\Sigma, X]$ by sequences $y(t) \in (\Sigma \cup X)^{<\omega}$.

Definition 1.3. We define the function $y : T[\Sigma, X] \to (\Sigma \cup X)^{<\omega}$ by

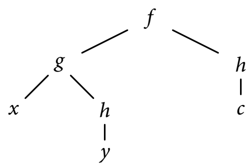
$$y(t) := t(x_0) \cdots t(x_n)$$

where $x_0 <_{\text{lex}} \cdots <_{\text{lex}} x_n$ is an enumeration of $\operatorname{dom} t$ in lexicographic order.

Remark. Equivalently, we can define $y(t)$ recursively as follows. If the root $\langle\rangle$ of t has exactly n successors $\langle 0\rangle, \dots, \langle n-1\rangle$ then we set

$$y(t) := t(\langle\rangle) \cdot y(t_{\langle 0\rangle}) \cdot \dots \cdot y(t_{\langle n-1\rangle}) .$$

Example. If t is the term



then $y(t) = fgxhyhc$.

The next lemma shows that it is save to identify t and $y(t)$. Below we will therefore not distinguish between the tree t and the sequence $y(t)$ encoding it, and we will use whatever formalism is the most convenient one at the time.

Lemma 1.4. *The function y is injective.*

Proof. Let s and t be terms and u and v arbitrary sequences. We prove by induction on $|y(s)|$ that

$$y(s)u = y(t)v \quad \text{implies} \quad s = t \text{ and } u = v .$$

For the special case that $u = \langle\rangle = v$ it follows that y is injective.

Let $f := s(\langle\rangle)$ and $g := t(\langle\rangle)$ be the function symbols at the roots of s and t, respectively. Then $y(s) = fx$ and $y(t) = gz$, for some sequences x and z. Since

$$fxu = y(s)u = y(t)v = gzv$$

it follows that $f = g$. Let n be the arity of f. If $n = 0$ then $x = \langle\rangle$ and $z = \langle\rangle$ and we have $fu = fv$ which implies $u = v$. Otherwise, let $s_i := s_{\langle i\rangle}$ and $t_i := t_{\langle i\rangle}$ be the subterms of s and t rooted at the successors of the root. By definition, we have

$$y(s) = f y(s_0)\cdots y(s_{n-1}) \quad \text{and} \quad y(t) = f y(t_0)\cdots y(t_{n-1})\,.$$

Hence, $y(s)u = y(t)v$ implies

$$y(s_0)\cdots y(s_{n-1})u = y(t_0)\cdots y(t_{n-1})v\,.$$

Since $|y(s_0)| < |y(s)|$ we can apply the inductive hypothesis and it follows that

$$s_0 = t_0 \quad \text{and} \quad y(s_1)\cdots y(s_{n-1})u = y(t_1)\cdots y(t_{n-1})v\,.$$

Applying the inductive hypothesis $n - 1$ more times we can conclude that

$$s_1 = t_1\,,\ \dots,\ s_{n-1} = t_{n-1} \quad \text{and} \quad u = v\,. \qquad \square$$

We can use the function y to obtain a simple upper bound on the number of finite Σ-terms.

Lemma 1.5. $\big|T[\Sigma, X]\big| \leq |\Sigma| \oplus |X| \oplus \aleph_0$.

Proof. Since $y : T[\Sigma, X] \to (\Sigma \cup X)^{<\omega}$ is injective we have

$$\big|T[\Sigma, X]\big| \leq \big|(\Sigma \cup X)^{<\omega}\big| = |\Sigma \cup X| \oplus \aleph_0 = |\Sigma| \oplus |X| \oplus \aleph_0\,,$$

by Lemma A4.4.31. $\square$

Remark. Note that, for finite terms $t \in T[\Sigma, X]$, we can perform proofs and definitions by induction on $|\mathrm{dom}(t)|$. Usually such proofs proceed in two steps. First, we show the desired property for all terms consisting of a single variable. Then we prove, for every n-ary function symbol, that, if the terms $t_0, \dots, t_{n-1}$ have the desired property then so does $f t_0 \dots t_{n-1}$.

We have introduced terms as names for derived operations, but we have yet to define which operation a term denotes.

Definition 1.6. Let $t \in T[\Sigma, X]$ be a Σ-term.

(a) The set of *free variables* of t is

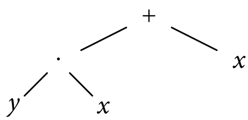
$$\operatorname{free}(t) := \operatorname{rng} t \cap X\,.$$

(b) Let $\mathfrak{A}$ be a Σ-structure, $t \in T[\Sigma, X]$ a Σ-term, and $\beta : X_0 \to A$ a function with domain $\operatorname{free}(t) \subseteq X_0 \subseteq X$. The *value* $t^{\mathfrak{A}}[\beta]$ of t in $\mathfrak{A}$ is defined inductively by the following rules.

- If $t = x \in X$ is a variable then $t^{\mathfrak{A}}[\beta] := \beta(x)$.
- If $t = f t_0 \dots t_{n-1}$ with $f \in \Sigma$ then

$$t^{\mathfrak{A}}[\beta] := f^{\mathfrak{A}}(t_0^{\mathfrak{A}}[\beta], \dots, t_{n-1}^{\mathfrak{A}}[\beta])\,.$$

Example. Consider the ring of integers $\mathfrak{Z} = \langle \mathbb{Z}, +, \cdot \rangle$ and let t be the term



If $\beta : X \to \mathbb{Z}$ maps $x \mapsto 3$ and $y \mapsto 5$ then $t^{\mathfrak{Z}}[\beta] = 18$.

A trivial induction on the size of a term t shows that its value $t^{\mathfrak{A}}[\beta]$ depends only on those variables that appear in t.

Lemma 1.7 (Coincidence Lemma). *Let $t \in T[\Sigma, X]$ be a Σ-term and $\mathfrak{A}$ a Σ-structure. If $\beta, \gamma : X \to A$ are variable assignments with*

$$\beta \restriction \operatorname{free}(t) = \gamma \restriction \operatorname{free}(t)$$

then $t^{\mathfrak{A}}[\beta] = t^{\mathfrak{A}}[\gamma]$.

Remark. We write $t(x_0, \dots, x_{n-1})$ to indicate that

$$\mathrm{free}(t) \subseteq \{x_0, \dots, x_{n-1}\}\,.$$

For such a term, we set

$$t^{\mathfrak{A}}(a_0, \dots, a_{n-1}) := t^{\mathfrak{A}}[\beta]$$

where $\beta : X \to A$ is any function with $\beta(x_i) = a_i$. By the Coincidence Lemma, this is well-defined.

The function symbols of Σ operate in a natural way on Σ-terms. A function symbol $f \in \Sigma$ of type $s_0 \dots s_{n-1} \to r$ maps terms $t_0, \dots, t_{n-1}$ of sort $s_0, \dots, s_{n-1}$, respectively, to the term $f t_0 \dots t_{n-1}$.

Definition 1.8. For an S-sorted signature Σ and a set of variables X, the *term algebra* $\mathfrak{T}[\Sigma, X]$ is the S-sorted Σ-structure defined as follows.

- The domain of sort $s \in S$ is $T_s[\Sigma, X]$.
- For each n-ary function symbol $f \in \Sigma$, we have the function $f^{\mathfrak{T}[\Sigma,X]}$ with

$$f^{\mathfrak{T}[\Sigma,X]}(t_0, \dots, t_{n-1}) := f t_0 \dots t_{n-1}\,.$$

- For each relation symbol $R \in \Sigma$, we have $R^{\mathfrak{T}[\Sigma,X]} := \varnothing$.

Example. If $\mathfrak{T} = \mathfrak{T}[\Sigma, X]$ is a term algebra and $\beta : X \to X$ the identity function then $t^{\mathfrak{T}}[\beta] = t$, for all $t \in T[\Sigma, X]$.

The term algebra $\mathfrak{T} = \mathfrak{T}[\Sigma, X]$ is also called the *free algebra* over X since the only equations $s^{\mathfrak{T}} = t^{\mathfrak{T}}$ that hold in $\mathfrak{T}$ are the trivial ones of the form $t = t$. This fact is used in the following lemma which states that $\mathfrak{T}$ is a universal object in the category of all Σ-structures.

Theorem 1.9. *Let $\mathfrak{A}$ be a Σ-structure and $\beta : X \to A$ an arbitrary function. There exists a unique homomorphism*

$$h : \mathfrak{T}[\Sigma, X] \to \mathfrak{A} \quad \textit{with} \quad h \restriction X = \beta\,.$$

The range of h is the set $\mathrm{rng}\, h = \langle\!\langle \mathrm{rng}\, \beta \rangle\!\rangle_{\mathfrak{A}}$.

Proof. We define $h(t) := t^{\mathfrak{A}}[\beta]$. For $x \in X$, it follows that

$$h(x) = x^{\mathfrak{A}}[\beta] = \beta(x) .$$

We claim that h is a homomorphism. Since all relations of $\mathfrak{T}[\Sigma, X]$ are empty we only need to verify that h commutes with functions. Let $f \in \Sigma$ be an n-ary function symbol and $t_0, \dots, t_{n-1} \in T[\Sigma, X]$. We have

$$\begin{aligned} h(ft_0 \dots t_{n-1}) &= (ft_0 \dots t_{n-1})^{\mathfrak{A}}[\beta] \\ &= f^{\mathfrak{A}}(t_0^{\mathfrak{A}}[\beta], \dots, t_{n-1}^{\mathfrak{A}}[\beta]) \\ &= f^{\mathfrak{A}}(h(t_0), \dots, h(t_{n-1})) , \end{aligned}$$

as desired.

Suppose that $g : \mathfrak{T}[\Sigma, X] \to \mathfrak{A}$ is a homomorphism with $g \restriction X = \beta$. By induction on $t \in T[\Sigma, X]$, we prove that $g(t) = h(t)$. If $x \in X$ then, by assumption, $g(x) = \beta(x) = h(x)$. For the inductive step, let $f \in \Sigma$ be an n-ary function symbol and $t_0, \dots, t_{n-1} \in T[\Sigma, X]$. We have

$$\begin{aligned} g(ft_0 \dots t_{n-1}) &= f^{\mathfrak{A}}(g(t_0), \dots, g(t_{n-1})) \\ &= f^{\mathfrak{A}}(h(t_0), \dots, h(t_{n-1})) = h(ft_0 \dots t_{n-1}) . \end{aligned}$$

Consequently, $g = h$.

It remains to prove that $\operatorname{rng} h = \langle\!\langle \operatorname{rng} \beta \rangle\!\rangle_{\mathfrak{A}}$. By Lemma B1.2.9, $\operatorname{rng} h$ induces a substructure of $\mathfrak{A}$. Since $\operatorname{rng} \beta \subseteq \operatorname{rng} h$ it follows that $\langle\!\langle \operatorname{rng} \beta \rangle\!\rangle_{\mathfrak{A}} \subseteq \operatorname{rng} h$.

To show that $\operatorname{rng} h \subseteq B := \langle\!\langle \operatorname{rng} \beta \rangle\!\rangle_{\mathfrak{A}}$ we prove, by induction on $t \in T[\Sigma, X]$, that $h(t) \in B$. For $x \in X$, we have $h(x) = \beta(x) \in \operatorname{rng} \beta \subseteq B$. Let $f \in \Sigma$ be an n-ary function symbol and $t_0, \dots, t_{n-1} \in T[\Sigma, X]$. Setting $a_i := h(t_i)$, for $i < n$, it follows that

$$h(ft_0 \dots t_{n-1}) = f^{\mathfrak{A}}(h(t_0), \dots, h(t_{n-1})) = f^{\mathfrak{A}}(a_0, \dots, a_{n-1}) .$$

By inductive hypothesis, we know that $a_0, \dots, a_{n-1} \in B$. Since B is closed under all functions of $\mathfrak{A}$ we have $f^{\mathfrak{A}}(a_0, \dots, a_{n-1}) \in B$, as desired. □

Remark. We can rephrase the theorem in the following way: For every S-sorted signature Σ and each Σ-structure $\mathfrak{A}$, there exists a bijection

$$\mathfrak{Hom}_s(\Sigma)(\mathfrak{T}[\Sigma, X], \mathfrak{A}) \to \mathfrak{Set}_S(X, A) : h \mapsto h \upharpoonright X\,,$$

where $\mathfrak{Set}_S$ is the category of S-sorted sets. In category theoretical terms this means that the term-algebra functor

$$\mathfrak{Set}_S \to \mathfrak{Hom}_s(\Sigma) : X \mapsto \mathfrak{T}[\Sigma, X]$$

and the forgetful functor

$$\mathfrak{Hom}_s(\Sigma) \to \mathfrak{Set}_S : \mathfrak{A} \mapsto A$$

form an *adjunction.*

Corollary 1.10. *Let $\mathfrak{A}$ be a Σ-structure and $X \subseteq A$ a subset. We have $\langle\!\langle X \rangle\!\rangle_{\mathfrak{A}} = \operatorname{rng} h$ where h is the unique homomorphism $h : \mathfrak{T}[\Sigma, X] \to \mathfrak{A}$ with $h \upharpoonright X = \mathrm{id}_X$.*

Corollary 1.11. *If $\mathfrak{A}$ is a Σ-structure and $X \subseteq A$ then*

$$\big|\langle\!\langle X \rangle\!\rangle_{\mathfrak{A}}\big| \leq \big|T[\Sigma, X]\big| \leq |X| \oplus |\Sigma| \oplus \aleph_0\,.$$

If s and t are terms and x a free variable of s then we can construct the term $s[x/t]$ by replacing every occurrence of x by the term t.

Definition 1.12. (a) Let Σ be an S-sorted signature and $t \in T[\Sigma, X]$ a term. If, for all $i < n$, $x_i \in X_{s_i}$ is a variable of sort s_i and $t_i \in T_{s_i}[\Sigma, X]$ a term of the same sort then we define the *substitution*

$$t[x_0/t_0, \dots, x_{n-1}/t_{n-1}] := t^{\mathfrak{T}[\Sigma, X]}[\beta]$$

where $\beta : X \to T[\Sigma, X]$ is the function with $\beta(x_i) := t_i$, for $i < n$, and $\beta(x) := x$, for all other variables $x \in X$.

(b) Similarly, if $\beta : A \to B$ is some function and a and b elements, then we denote by $\beta[a/b]$ the function $A \cup \{a\} \to B \cup \{b\}$ with

$$\beta[a/b](x) := \begin{cases} b & \text{if } x = a\,, \\ \beta(x) & \text{otherwise}\,. \end{cases}$$

The next lemma states the trivial fact that, when computing the value of a term $s[x/t]$ it does not matter whether we substitute t for x first and then evaluate the whole term, or whether we compute the value of t first and then evaluate s with the corresponding value for x. For instance, if $s = x + y$ and $t = y + y$ then $s[x/t] = (y + y) + y$ and the lemma claims that $s[x/t](1) = (1 + 1) + 1 = 3$ coincides with $s(2, 1) = 2 + 1 = 3$.

Lemma 1.13 (Substitution Lemma)**.** *Let $s, t \in T[\Sigma, X]$ be terms, $x \in X$ a variable, $\mathfrak{A}$ a Σ-structure, and $\beta : X \to A$ function. We have*

$$(s[x/t])^{\mathfrak{A}}[\beta] = s^{\mathfrak{A}}[\beta'] \qquad \textit{where} \qquad \beta' := \beta[x/t^{\mathfrak{A}}[\beta]]\,.$$

Proof. We prove the claim by induction on the term s. If $s = x$ then

$$(x[x/t])^{\mathfrak{A}}[\beta] = t^{\mathfrak{A}}[\beta] = \beta'(x) = x^{\mathfrak{A}}[\beta']\,.$$

If $s = y \neq x$ then

$$(y[x/t])^{\mathfrak{A}}[\beta] = y^{\mathfrak{A}}[\beta] = \beta(y) = \beta'(y) = y^{\mathfrak{A}}[\beta']\,.$$

Finally, if $s = fs_0 \dots s_{n-1}$ then we have by inductive hypothesis

$$\begin{aligned}(fs_0 \dots s_{n-1})[x/t]^{\mathfrak{A}}[\beta] &= f^{\mathfrak{A}}(s_0[x/t]^{\mathfrak{A}}[\beta], \dots, s_{n-1}[x/t]^{\mathfrak{A}}[\beta]) \\ &= f^{\mathfrak{A}}(s_0^{\mathfrak{A}}[\beta'], \dots, s_{n-1}^{\mathfrak{A}}[\beta']) \\ &= (fs_0 \dots s_{n-1})^{\mathfrak{A}}[\beta']\,. \end{aligned}$$ □

The operations $T[\Sigma, X]$ and $\mathfrak{T}[\Sigma, X]$ assigning to a signature Σ and a set X of variables, respectively, the set of terms and the term algebra can be seen as functors between suitable categories.

Definition 1.14. (a) Let $\mathfrak{SigVar}$ be the category consisting of all triples $\langle S, \Sigma, X\rangle$ where S is a set of sorts, Σ an S-sorted signature, and X an S-sorted set of variables. The morphisms

$$\langle \chi, \varphi, \psi\rangle : \langle S, \Sigma, X\rangle \to \langle T, \Gamma, Y\rangle$$

are triples of functions $\chi : S \to T$, $\varphi : \Sigma \to \Gamma$, and $\psi : X \to Y$ with the following properties:

- A relation symbol $R \in \Sigma$ of type $s_0 \dots s_{n-1}$ is mapped to a relation symbol $\varphi(R) \in \Gamma$ of type $\chi(s_0) \dots \chi(s_{n-1})$.
- A function symbol $f \in \Sigma$ of type $s_0 \dots s_{n-1} \to t$ is mapped to a function symbol $\varphi(f) \in \Gamma$ of type $\chi(s_0) \dots \chi(s_{n-1}) \to \chi(t)$.
- A variable $x \in X$ of type s is mapped to a variable $\psi(x) \in Y$ of type $\chi(s)$.

Since the set of sorts S is determined by the signature Σ we will usually omit it from $\langle S, \Sigma, X\rangle$ and just write $\langle \Sigma, X\rangle$.

(b) We define two subcategories of $\mathfrak{SigVar}$. The category $\mathfrak{Sig}$ consists of all triples $\langle S, \Sigma, X\rangle \in \mathfrak{SigVar}$ with $X = \varnothing$ and the category $\mathfrak{Var}$ consists of all $\langle S, \Sigma, X\rangle \in \mathfrak{SigVar}$ with $\Sigma = \varnothing$.

(c) A morphism $\alpha = \langle \chi, \varphi, \psi\rangle \in \mathfrak{SigVar}(\langle \Sigma, X\rangle, \langle \Gamma, Y\rangle)$ induces the map

$$T[\alpha] : T[\Sigma, X] \to T[\Gamma, Y]$$

which assigns to a term $t \in T_s[\Sigma, X]$ the term $T[\alpha](t) \in T_{\chi(s)}[\Gamma, Y]$ with

$$T[\alpha](t)(x) := \begin{cases} \varphi(t(x)) & \text{if } t(x) \in \Sigma\,, \\ \psi(t(x)) & \text{if } t(x) \in X\,. \end{cases}$$

Let $\mathfrak{Term}$ denote the category with objects $T[\Sigma, X]$, for all Σ, X, and morphisms

$$\mathfrak{Term}(T[\Sigma, X], T[\Gamma, Y]) := \{\, T[\alpha] \mid \alpha \in \mathfrak{SigVar}(\langle \Sigma, X\rangle, \langle \Gamma, Y\rangle) \,\}\,.$$

Example. Let $\Sigma := \{\circ, {}^{-1}, e\}$ be the signature of multiplicative groups and $\Gamma := \{+, -, 0\}$ the signature of additive groups. Since there exists an isomorphism $\Sigma \to \Gamma$ in $\mathfrak{Sig}$ these signatures are interchangeable.

Remark. It follows immediately from the definition of $\mathfrak{Term}$ that the operation

$$\langle \Sigma, X\rangle \mapsto T[\Sigma, X] \quad \text{and} \quad \alpha \mapsto T[\alpha]$$

forms a functor $T : \mathfrak{SigVar} \to \mathfrak{Term}$.

We can also define corresponding categories of structures.

Definition 1.15. (a) Let $\mu = \langle \chi, \varphi \rangle : \langle S, \Sigma \rangle \to \langle T, \Gamma \rangle$ be a morphism of $\mathfrak{Sig}$. The *μ-reduct* $\mathfrak{A}|_\mu$ of a Γ-structure $\mathfrak{A}$ is the Σ-structure $\mathfrak{B}$ where the domain of sort $s \in S$ is $B_s := A_{\chi(s)}$ and the relations and functions are defined by

$$\xi^{\mathfrak{B}} := \varphi(\xi)^{\mathfrak{A}}, \qquad \text{for } \xi \in \Gamma.$$

(b) For a signature Σ, we denote by $\mathrm{Str}[\Sigma]$ the class of all Σ-structures and by $\mathrm{Str}[\Sigma, X]$ the class of all pairs $\langle \mathfrak{A}, \beta \rangle$ where $\mathfrak{A}$ is a Σ-structure and $\beta : X \to A$ a variable assignment.

Every morphism $\mu = \langle \chi, \varphi, \psi \rangle : \langle T, \Gamma, Y \rangle \to \langle S, \Sigma, X \rangle$ of $\mathfrak{SigVar}$ induces a function

$$\mathrm{Str}[\mu] : \mathrm{Str}[\Sigma, X] \to \mathrm{Str}[\Gamma, Y] : \langle \mathfrak{A}, \beta \rangle \mapsto \langle \mathfrak{A}|_\mu, \beta \circ \psi \rangle .$$

(c) In the category $\mathfrak{StrVar}$ the objects are the classes $\mathrm{Str}[\Sigma, X]$ and the morphisms are all mappings $\mathrm{Str}[\Sigma, X] \to \mathrm{Str}[\Gamma, Y]$ induced by a morphism $\langle \Gamma, Y \rangle \to \langle \Sigma, X \rangle$ of $\mathfrak{SigVar}$. As above we define the subcategory $\mathfrak{Str}$ where the objects are those classes $\mathrm{Str}[\Sigma, X]$ with $X = \varnothing$.

(d) The *canonical functor* $\mathrm{Str} : \mathfrak{SigVar} \to \mathfrak{StrVar}$ maps a pair $\langle \Sigma, X \rangle$ to the class $\mathrm{Str}[\Sigma, X]$ and a morphism $\langle \Sigma, X \rangle \to \langle \Gamma, Y \rangle$ to the function $\mathrm{Str}[\Gamma, Y] \to \mathrm{Str}[\Sigma, X]$ it induces. By abuse of notation we denote the corresponding functor $\mathrm{Str} : \mathfrak{Sig} \to \mathfrak{Str}$ by the same symbol. Note that Str is contravariant.

Remark. Suppose that $\Sigma \subseteq \Gamma$ and let $\mathfrak{A}$ be a Γ-structure. If $\mu : \Sigma \to \Gamma$ is inclusion map then $\mathfrak{A}|_\mu = \mathfrak{A}|_\Sigma$ is the ordinary Σ-reduct of $\mathfrak{A}$.

The next lemma relates the structures $\mathfrak{A}$ and $\mathrm{Str}[\mu](\mathfrak{A})$. It follows immediately from the respective definitions.

Lemma 1.16. *Let $\mu : \langle \Sigma, X \rangle \to \langle \Gamma, Y \rangle$ be a morphism of $\mathfrak{SigVar}$. For all interpretations $\langle \mathfrak{A}, \beta \rangle \in \mathrm{Str}[\Gamma, Y]$ and terms $t \in T[\Sigma, X]$, we have*

$$\big(T[\mu](t)\big)^{\mathfrak{A}}[\beta] = t^{\mathfrak{B}}[\gamma] \quad \textit{where} \quad \langle \mathfrak{B}, \gamma \rangle = \mathrm{Str}[\mu](\mathfrak{A}, \beta) .$$

$$\begin{array}{ccc} T[\Sigma, X] & \xrightarrow{T[\mu]} & T[\Gamma, Y] \\ \downarrow & & \downarrow \\ \mathfrak{B} & \longrightarrow & \mathfrak{A} \end{array}$$

Example. Let $\Sigma = \{\circ, ^{-1}, e\}$ and $\Gamma = \{+, -, 0\}$ be signatures of groups and $X = \{x\}$ and $Y = \{y\}$ sets of variables. Consider the morphism

$$\mu = \langle \mathrm{id}, \varphi, \psi \rangle : \langle \Sigma, X \rangle \to \langle \Gamma, Y \rangle$$

with $\varphi(\circ) = +, \quad \varphi(^{-1}) = -, \quad \varphi(e) = 0, \quad$ and $\quad \psi(x) = y.$

Let $\mathfrak{Z} = \langle \mathbb{Z}, +, -, 0 \rangle$ be the additive group of the integers and $\beta : y \mapsto 3$ a variable assignment. Then $\mathrm{Str}[\mu]\langle \mathfrak{Z}, \beta \rangle = \langle \mathfrak{Z}', \gamma \rangle$ where $\mathfrak{Z}' = \langle \mathbb{Z}, \circ, ^{-1}, e \rangle$ and $\gamma : x \mapsto 3$. For the term $t(x) = x \circ e \circ x^{-1}$ the lemma states that

$$t^{\mathfrak{Z}'}[\gamma] = (x \circ e \circ x^{-1})^{\mathfrak{Z}'}[\gamma] = 3 + 0 - 3 = 0$$

equals

$$\big(T[\mu](t)\big)^{\mathfrak{Z}}[\beta] = (y + 0 + (-y))^{\mathfrak{Z}}[\beta] = 3 + 0 - 3 = 0\,.$$

2. *Direct and reduced products*

Products are a common construction in algebra since many important classes, such as groups and rings, are closed under products. In this section we will introduce products of arbitrary structures and prove some of their basic properties.

Below we will frequently deal with tuples of sequences of the form

$$\bar{a} = \big\langle (a^i_0)_{i \in I}, \dots, (a^i_{n-1})_{i \in I} \big\rangle \in (A^I)^n\,.$$

To simplify notation we define

$$\bar{a}^i := \langle a^i_0, \dots, a^i_{n-1} \rangle \in A^n \quad \text{and} \quad \bar{a}_k := (a^i_k)_{i \in I} \in A^I\,.$$

Definition 2.1. Let $(\mathfrak{A}^i)_{i\in I}$ be a sequence of Σ-structures.

(a) Their *direct product* is the Σ-structure

$$\mathfrak{B} := \prod_{i\in I} \mathfrak{A}^i\,,$$

where the domain of sort s is $B_s := \prod_{i\in I} A^i_s$, for every n-ary relation $R \in \Sigma$, we have

$$R^{\mathfrak{B}} = \left\{ \bar{a} \in B^n \;\middle|\; \bar{a}^i \in R^{\mathfrak{A}^i} \text{ for all } i \in I \right\},$$

and, for each function $f \in \Sigma$,

$$f^{\mathfrak{B}}(\bar{a}) := \big(f^{\mathfrak{A}^i}(\bar{a}^i)\big)_{i\in I}\,.$$

If $\mathfrak{A}^i = \mathfrak{A}$, for all $i \in I$, we usually write $\mathfrak{A}^I$ instead of $\prod_{i\in I} \mathfrak{A}$.

(b) Recall that the *k-th projection* is the function

$$\mathrm{pr}_k : \prod_{i\in I} \mathfrak{A}^i \to \mathfrak{A}^k : (a^i)_{i\in I} \mapsto a^k\,.$$

Example. (a) Let $\mathfrak{U} = \langle U, +, (\lambda_a)_{a\in K}\rangle$ be a K-vector space of dimension 1. Every K-vector space $\mathfrak{V} = \langle V, +, (\lambda_a)_a\rangle$ of dimension $n < \omega$ is isomorphic to $\mathfrak{U}^n$.

(b) Let $\mathfrak{B}_2 = \langle [2], \sqcup, \sqcap, 0, 1, {}^*, \leq\rangle$ be the two-element boolean algebra and $\mathfrak{A} = \langle \mathscr{P}(X), \cup, \cap, \varnothing, X, {}^*, \subseteq\rangle$ the power-set algebra of a set X. Then $\mathfrak{A} \cong \prod_{i\in X} \mathfrak{B}_2 = \mathfrak{B}_2^X$.

Analogously to products of sets we can characterise products of structures as terminal objects in a suitable category.

Lemma 2.2. *Let* $\mathrm{pr}_k : \prod_{i\in I} \mathfrak{A}^i \to \mathfrak{A}^k$ *be a projection.*

(a) pr_k *is a surjective homomorphism.*

(b) pr_k *is semi-strict if and only if, for every relation symbol* R*, the set* $\{\, i \in I \mid R^{\mathfrak{A}^i} = \varnothing \,\}$ *contains* k *or it equals either* $\varnothing$ *or* I.

Lemma 2.3. *Let* $(\mathfrak{A}^i)_{i \in I}$ *be a sequence of* Σ*-structures. For every structure* $\mathfrak{B}$ *and all homomorphisms* $h_k : \mathfrak{B} \to \mathfrak{A}^k$, $k \in I$, *there exists a unique homomorphism* $\varphi : \mathfrak{B} \to \prod_{i \in I} \mathfrak{A}^i$ *with* $h_k = \mathrm{pr}_k \circ \varphi$, *for all* k.

Exercise 2.1. Prove the preceding lemmas.

Exercise 2.2. Prove that the direct product of groups is again a group and that the direct product of rings is a ring.

Given a class $\mathcal{K}$ of structures that is closed under products one can try to classify $\mathcal{K}$ by isolating a subclass $\mathcal{K}_0 \subseteq \mathcal{K}$ such that every structure in $\mathcal{K}$ can be expressed as product of elements of $\mathcal{K}_0$. The classification of finitely generated abelian groups is of this kind. If $\mathcal{K}$ is furthermore closed under substructures then we can also try to find a subclass $\mathcal{K}_1$ such that every structure in $\mathcal{K}$ is the substructure of a product of elements of $\mathcal{K}_1$. For instance, every K-vector space of dimension κ is a substructure of K^κ. This motivates an investigation of substructures of products.

Definition 2.4. Let $(\mathfrak{A}^i)_{i \in I}$ be a sequence of Σ-structures.

(a) A Σ-structure $\mathfrak{B}$ is a *subdirect product* of $(\mathfrak{A}^i)_i$ if there exists an embedding $g : \mathfrak{B} \to \prod_{i \in I} \mathfrak{A}^i$ such that $\mathrm{pr}_k \circ g$ is surjective and semi-strict, for all $k \in I$.

(b) A structure $\mathfrak{B}$ is *subdirectly irreducible* if, for every sequence $(\mathfrak{A}^i)_i$ of which $\mathfrak{B}$ is a subdirect product, there exists an index k with $\mathfrak{B} \cong \mathfrak{A}^k$.

Lemma 2.5. *Let* $\mathfrak{B}$ *be a subdirect product of* $(\mathfrak{A}^i)_{i \in I}$ *and* $g : \mathfrak{B} \to \prod_i \mathfrak{A}^i$ *the corresponding embedding. If* $s, t \in T[\Sigma, X]$ *are terms,* $\beta : X \to B$ *a variable assignment, and* $\beta_i := \mathrm{pr}_i \circ g \circ \beta$ *then we have*

$$s^{\mathfrak{B}}[\beta] = t^{\mathfrak{B}}[\beta] \quad \text{iff} \quad s^{\mathfrak{A}^i}[\beta_i] = t^{\mathfrak{A}^i}[\beta_i], \quad \textit{for all } i \in I .$$

Proof. The lemma follows immediately if we can show that

$$g(t^{\mathfrak{B}}[\beta]) = (t^{\mathfrak{A}^i}[\beta_i])_i .$$

We proceed by induction on the size of t. For $t = x \in X$, we have

$$g(x^{\mathfrak{B}}[\beta]) = g(\beta(x)) = (\beta_i(x))_i .$$

If $t = fs_0 \dots s_{n-1}$ then

$$\begin{aligned} g\big((fs_0 \dots s_{n-1})^{\mathfrak{B}}[\beta]\big) &= g\big(f^{\mathfrak{B}}(s_0^{\mathfrak{B}}[\beta], \dots, s_{n-1}^{\mathfrak{B}}[\beta])\big) \\ &= f^{\prod_i \mathfrak{A}^i}\big(g(s_0^{\mathfrak{B}}[\beta]), \dots, g(s_{n-1}^{\mathfrak{B}}[\beta])\big) \\ &= f^{\prod_i \mathfrak{A}^i}\big((s_0^{\mathfrak{A}^i}[\beta_i])_i, \dots, (s_{n-1}^{\mathfrak{A}^i}[\beta_i])_i\big) \\ &= \big(f^{\mathfrak{A}^i}(s_0^{\mathfrak{A}^i}[\beta_i], \dots, s_{n-1}^{\mathfrak{A}^i}[\beta_i])\big)_i \\ &= \big((fs_0 \dots s_{n-1})^{\mathfrak{A}^i}[\beta_i]\big)_i . \end{aligned}$$

□

An important special case of a subdirect product are *reduced products* which are obtained from a product by factorising over a filter. To define what we mean by 'factorising over a filter' we need some preliminaries.

Definition 2.6. Let $(\mathfrak{A}^i)_{i \in I}$ be a sequence of Σ-structures and $\mathfrak{u} \subseteq \wp(I)$ a filter. Let S be the set of sorts of Σ and set

$$B := \bigcup_{\substack{s \in S \\ w \in \mathfrak{u}}} B_s^w \quad \text{where} \quad B_s^w := \prod_{i \in w} A_s^i .$$

For $\bar{a}, \bar{b} \in B_{s_0}^{w_0} \times \dots \times B_{s_{n-1}}^{w_{n-1}}$, we define

$$\begin{aligned} [\![\bar{a}^i = \bar{b}^i]\!]_i &:= \{\, i \in w_0 \cap \dots \cap w_{n-1} \mid \bar{a}^i = \bar{b}^i \,\}, \\ [\![\bar{a}^i \in R]\!]_i &:= \{\, i \in w_0 \cap \dots \cap w_{n-1} \mid \bar{a}^i \in R^{\mathfrak{A}^i} \,\}, \end{aligned}$$

and $\quad \bar{a} \sim_{\mathfrak{u}} \bar{b} \quad$:iff $\quad [\![\bar{a}^i = \bar{b}^i]\!]_i \in \mathfrak{u}$.

We denote the $\sim_{\mathfrak{u}}$-class of a tuple $\bar{a} \in B$ by $[\bar{a}]_{\mathfrak{u}}$.

Lemma 2.7. *Let $(\mathfrak{A}^i)_{i \in I}$ be a sequence of Σ-structures and $\mathfrak{u} \subseteq \wp(I)$ a filter.*

(a) $\sim_{\mathfrak{u}}$ *is an equivalence relation.*

(b) $\bar{a} \sim_{\mathfrak{u}} \bar{b}$ *implies* $[\![\bar{a}^i \in R]\!]_i \in \mathfrak{u}$ iff $[\![\bar{b}^i \in R]\!]_i \in \mathfrak{u}$.

(c) $\bar{a} \sim_{\mathfrak{u}} \bar{b}$ *implies* $f^{\mathfrak{B}}(\bar{a}) \sim_{\mathfrak{u}} f^{\mathfrak{B}}(\bar{b})$.

Proof. (a) We have $(a^i)_{i\in I} \sim_{\mathfrak{u}} (a^i)_{i\in I}$ since $I \in \mathfrak{u}$. Furthermore, since $=$ is symmetric it follows that so is $\sim_{\mathfrak{u}}$. Finally, suppose that

$$(a^i)_{i\in I} \sim_{\mathfrak{u}} (b^i)_{i\in I} \quad \text{and} \quad (b^i)_{i\in I} \sim_{\mathfrak{u}} (c^i)_{i\in I}\,.$$

Since $[\![(a^i)_i = (c^i)_i]\!]_i \supseteq [\![(a^i)_i = (b^i)_i]\!]_i \cap [\![(b^i)_i = (c^i)_i]\!]_i \in \mathfrak{u}$

it follows that $(a^i)_{i\in I} \sim_{\mathfrak{u}} (c^i)_{i\in I}$.

(b) We have $[\![\bar{a}^i = \bar{b}^i]\!]_i \in \mathfrak{u}$ and, by symmetry, we may assume that $[\![\bar{a}^i \in R]\!]_i \in \mathfrak{u}$. Hence, $[\![\bar{b}^i \in R]\!]_i \supseteq [\![\bar{a}^i \in R]\!]_i \cap [\![\bar{a}^i = \bar{b}^i]\!]_i \in \mathfrak{u}$ and it follows that $[\![\bar{b}^i \in R]\!]_i \in \mathfrak{u}$.

(c) follows immediately from $[\![f(\bar{a}^i) = f(\bar{b}^i)]\!]_i \supseteq [\![\bar{a}^i = \bar{b}^i]\!]_i \in \mathfrak{u}$. □

Definition 2.8. Let $\mathfrak{u}$ be a filter over I and $J \subseteq I$. The *restriction* of $\mathfrak{u}$ to J is the set

$$\mathfrak{u}|_J := \{\, s \cap J \mid s \in \mathfrak{u} \,\}\,.$$

Lemma 2.9. *Let $\mathfrak{u}$ be a filter over I and $S \in \mathfrak{u}$.*

(a) $\mathfrak{u}|_S$ *is a filter over S.*

(b) *If $\mathfrak{u}$ is an ultrafilter then so is $\mathfrak{u}|_S$.*

Definition 2.10. Let $(\mathfrak{A}^i)_{i\in I}$ be a sequence of Σ-structures and $\mathfrak{u} \subseteq \mathscr{P}(I)$ a filter.

(a) The *reduced product* of $(\mathfrak{A}^i)_{i\in I}$ over $\mathfrak{u}$ is the structure

$$\mathfrak{B} := \prod_{i\in I} \mathfrak{A}^i/\mathfrak{u}$$

defined as follows. For each sort s, let

$$I_s := \{\, i \in I \mid A^i_s \neq \varnothing \,\}\,.$$

The domain of sort s is

$$B_s := \begin{cases} \big(\prod_{i\in I_s} A^i_s\big)/{\sim_{\mathfrak{u}|_{I_s}}} & \text{if } I_s \in \mathfrak{u}\,, \\ \varnothing & \text{otherwise}\,. \end{cases}$$

For every n-ary relation $R \in \Sigma$, we have

$$R^{\mathfrak{B}} := \{ [\bar{a}]_{\mathfrak{u}} \in B^n \mid [\![\bar{a}^i \in R]\!]_i \in \mathfrak{u} \},$$

and, for each function $f \in \Sigma$,

$$f^{\mathfrak{B}}([\bar{a}]_{\mathfrak{u}}) := [(b_i)_i]_{\mathfrak{u}} \quad \text{where} \quad b_i := f^{\mathfrak{A}^i}(\bar{a}^i).$$

(b) If $\mathfrak{u}$ is an ultrafilter then $\prod_{i \in I} \mathfrak{A}^i / \mathfrak{u}$ is also called an *ultraproduct*. In the special case that $\mathfrak{A}^i = \mathfrak{A}$, for all i, we call $\prod_{i \in I} \mathfrak{A}/\mathfrak{u}$ the *ultrapower* of $\mathfrak{A}$ over $\mathfrak{u}$ and we simply write $\mathfrak{A}^{\mathfrak{u}}$.

Remark. Note that $\prod_{i \in I} \mathfrak{A}^i / \mathfrak{u}$ is well-defined by Lemma 2.7.

Lemma 2.11. *Let $\mathfrak{B} = \prod_{i \in I} \mathfrak{A}^i / \mathfrak{u}$. If $s, t \in T[\Sigma, X]$ are terms, $\beta : X \to B$ a variable assignment, and $\beta_i := \mathrm{pr}_i \circ \beta$ then we have*

$$s^{\mathfrak{B}}[\beta] = t^{\mathfrak{B}}[\beta] \quad \text{iff} \quad \{ i \in I \mid s^{\mathfrak{A}^i}[\beta_i] = t^{\mathfrak{A}^i}[\beta_i] \} \in \mathfrak{u}.$$

Proof. By induction on t one can show that $t^{\mathfrak{B}}[\beta] = \big[(t^{\mathfrak{A}^i}[\beta_i])_i \big]_{\mathfrak{u}}$. Consequently, the claim follows by definition of $\sim_{\mathfrak{u}}$. □

Exercise 2.3. Prove that an ultraproduct of linear orders is again a linear order and that an ultraproduct of fields is a field.

Lemma 2.12. *Let $\mathfrak{A}$ be a Σ-structure and $\mathfrak{u}$ a proper filter. There exists an embedding $h : \mathfrak{A} \to \mathfrak{A}^{\mathfrak{u}}$.*

Proof. Suppose that $\mathfrak{u}$ is a filter over I. We denote by $\bar{a}^=$ the constant sequence $(\bar{a}^i)_i$ with $\bar{a}^i := \bar{a}$, for all i. We claim that $h : a \mapsto [a^=]_{\mathfrak{u}}$ is the desired embedding.

h is injective since, if $a \neq b$ then $[\![(a^=)^i = (b^=)^i]\!]_i = \varnothing \notin \mathfrak{u}$, which implies that $h(a) \neq h(b)$. If $R \in \Sigma$ is an n-ary relation then we have

$$[\![(\bar{a}^=)^i \in R]\!]_i = \begin{cases} I \in \mathfrak{u} & \text{if } \bar{a} \in R^{\mathfrak{A}}, \\ \varnothing \notin \mathfrak{u} & \text{if } \bar{a} \notin R^{\mathfrak{A}}. \end{cases}$$

Therefore, we have $\bar{a} \in R^{\mathfrak{A}}$ iff $h(\bar{a}) \in R^{\mathfrak{A}^{\mathfrak{u}}}$. Finally, if $f \in \Sigma$ is an n-ary function then we have

$$\begin{aligned} f^{\mathfrak{A}^{\mathfrak{u}}}(h(\bar{a})) = f^{\mathfrak{A}^{\mathfrak{u}}}([\bar{a}^{=}]_{\mathfrak{u}}) &= \big[f^{\mathfrak{A}^{I}}(\bar{a}^{=})\big]_{\mathfrak{u}} \\ &= \big[(f^{\mathfrak{A}}(\bar{a}))^{=}\big]_{\mathfrak{u}} = h(f^{\mathfrak{A}}(\bar{a}))\,. \end{aligned}$$

It follows that h is the desired injective strict homomorphism. □

Example. Let $\mathfrak{R} = \langle \mathbb{R}, +, -, \cdot, 0, 1, \leq \rangle$ be the ordered field of real numbers and $\mathfrak{u}$ a non-principal ultrafilter on ω. The ultrapower $\mathfrak{R}^{\mathfrak{u}}$ is again an ordered field with $\mathfrak{R} \subseteq \mathfrak{R}^{\mathfrak{u}}$. Let $(a_i)_{i<\omega} \in \mathbb{R}^{\omega}$, be the sequence with $a_i = i$, and let $a := [(a_i)_i]_{\mathfrak{u}}$ be its $\sim_{\mathfrak{u}}$-class. It follows that $a > x$, for every real number $x \in \mathbb{R}$. Hence, $\mathfrak{R}^{\mathfrak{u}}$ contains an infinite number a. The element a^{-1} is positive but smaller than every positive real number. Thus, we have constructed an extension of $\mathfrak{R}$ containing infinite and infinitesimal elements.

In the definition of a reduced product we have neglected those factors with empty domains. This choice is motivated by the following observation which is an immediate consequence of Lemma **??** below. For simplicity, we only treat the case that all domains are nonempty.

Lemma 2.13. *Let $(\mathfrak{A}^i)_{i \in I}$ be a family of Σ-structures whose domains are all nonempty and let $\mathfrak{u}$ be a filter over I. For every $J \in \mathfrak{u}$, we have*

$$\prod_{i \in I} \mathfrak{A}^i/\mathfrak{u} \cong \prod_{j \in J} \mathfrak{A}^j/\mathfrak{u}|_J\,.$$

Proof. To simplify notation set $\mathfrak{v} := \mathfrak{u}|_J$ and define

$$\mathfrak{A}_I := \textstyle\prod_{i \in I} \mathfrak{A}^i\,, \qquad \mathfrak{A}_I/\mathfrak{u} := \textstyle\prod_{i \in I} \mathfrak{A}^i/\mathfrak{u}\,,$$

and $$\mathfrak{A}_J := \textstyle\prod_{j \in J} \mathfrak{A}^j\,, \qquad \mathfrak{A}_J/\mathfrak{v} := \textstyle\prod_{j \in J} \mathfrak{A}^j/\mathfrak{v}\,.$$

For sequences $(\bar{a}^i)_{i \in I}$ set $\bar{a} \restriction J := (\bar{a}^j)_{j \in J}$. Let

$$\begin{aligned} &\varphi : \mathfrak{A}_I \to \mathfrak{A}_I/\mathfrak{u} : (a^i)_i \mapsto [(a^i)_i]_{\mathfrak{u}} \\ &\psi : \mathfrak{A}_J \to \mathfrak{A}_J/\mathfrak{v} : (a^j)_j \mapsto [(a^j)_j]_{\mathfrak{v}} \\ &\pi : \mathfrak{A}_I \to \mathfrak{A}_j : \bar{a} \mapsto \bar{a} \restriction J \end{aligned}$$

$$\begin{array}{ccc} \mathfrak{A}_I & \xrightarrow{\ \pi\ } & \mathfrak{A}_J \\ {\scriptstyle\varphi}\downarrow & & \downarrow{\scriptstyle\psi} \\ \mathfrak{A}_I/\mathfrak{u} & \xrightarrow[\ \eta\ ]{} & \mathfrak{A}_J/\mathfrak{v} \end{array}$$

be the canonical homomorphisms. For sequences $(a^i)_{i \in I}$ and $(b^i)_{i \in I}$, we have

$$\begin{aligned} \big\langle (a^i)_i, (b^i)_i \big\rangle \in \ker \varphi \quad &\text{iff} \quad [\![a^i = b^i]\!]_i \in \mathfrak{u} \\ &\text{iff} \quad [\![a^i = b^i]\!]_i \cap J \in \mathfrak{v} \\ &\text{iff} \quad \big\langle (a^i)_{i \in I}, (b^i)_{i \in I} \big\rangle \in \ker(\psi \circ \pi) \,. \end{aligned}$$

By the Factorisation Lemma, it follows that there exists a unique bijection $\eta : \varphi(\mathfrak{A}_I) \to (\psi \circ \pi)(\mathfrak{A}_I)$ with $\psi \circ \pi = \eta \circ \varphi$, i.e.,

$$\eta([\bar{a}]_{\mathfrak{u}}) = [\bar{a} \restriction J]_{\mathfrak{v}} \,.$$

It remains to prove that this function is an isomorphism. (Note that, if φ and ψ are semi-strict then we can apply Corollary B1.2.7.)

For a function symbol f, we have

$$\begin{aligned} \eta\big(f^{\mathfrak{A}_I/\mathfrak{u}}([\bar{a}]_{\mathfrak{u}})\big) &= \eta\big([f^{\mathfrak{A}_I}(\bar{a})]_{\mathfrak{u}}\big) \\ &= \big[f^{\mathfrak{A}_J}(\bar{a} \restriction J)\big]_{\mathfrak{v}} \\ &= f^{\mathfrak{A}_J/\mathfrak{v}}([\bar{a} \restriction J]_{\mathfrak{v}}) = f^{\mathfrak{A}_J/\mathfrak{v}}\big(\eta([\bar{a}]_{\mathfrak{u}})\big) \,, \end{aligned}$$

and, for a relation symbol R, we have

$$\begin{aligned} [\bar{a}]_{\mathfrak{u}} \in R^{\mathfrak{A}_I/\mathfrak{u}} \quad &\text{iff} \quad [\![\bar{a}^i \in R]\!]_i \in \mathfrak{u} \\ &\text{iff} \quad [\![\bar{a}^i \in R]\!]_i \cap J \in \mathfrak{u} \\ &\text{iff} \quad \eta([\bar{a}]_{\mathfrak{u}}) = [\bar{a} \restriction J]_{\mathfrak{v}} \in R^{\mathfrak{A}_J/\mathfrak{v}}. \end{aligned}$$

□

Corollary 2.14. *Let $(\mathfrak{A}^i)_{i \in I}$ be a family of Σ-structures. If $\mathfrak{u} = \Uparrow J$ is a principal filter over I then*

$$\prod_{i \in I} \mathfrak{A}^i / \mathfrak{u} \cong \prod_{j \in J} \mathfrak{A}^j .$$

In particular, if $J = \{j\}$ then $\prod_{i \in I} \mathfrak{A}^i / \mathfrak{u} \cong \mathfrak{A}^j$.

3. Directed limits and colimits

With each structure $\mathfrak{A}$ we can associate the family of its finitely generated substructures, ordered by inclusion. Conversely, given such a partially ordered family of structures, we can try to assemble them into a single structure. This leads to the notion of a *directed colimit.* Not every family of structures arises from a superstructure $\mathfrak{A}$. Before introducing directed colimits, we therefore isolate the key property of those families that do.

Definition 3.1. Let κ be a cardinal. We call a partial order $\mathfrak{I} = \langle I, \leq \rangle$ *κ-directed* if every subset $X \subseteq I$ of size $|X| < \kappa$ has an upper bound. For $\kappa = \aleph_0$, we simply speak of *directed* sets.

Example. (a) Every ideal is directed.

(b) An infinite cardinal κ is regular if, and only if, the linear order $\langle \kappa, \leq \rangle$ is κ-directed.

(c) Let A be a set, κ a regular cardinal, and $F := \{ X \subseteq A \mid |X| < \kappa \}$. The order $\langle F, \subseteq \rangle$ is κ-directed.

(d) Let $\mathfrak{A}$ be a Σ-structure and $\mathcal{S}$ the class of all substructures of $\mathfrak{A}$ that are generated by a set of size less than κ. If κ is regular, the order $(\mathcal{S}, \subseteq)$ is κ-directed.

Let us show that, if we partition a directed set into finitely many parts, at least one of them is again directed.

Definition 3.2. Let $\langle I, \leq \rangle$ be a directed partial order. A subset $D \subseteq I$ is *dense* if $\Uparrow i \cap D \neq \emptyset$, for all $i \in I$.

Lemma 3.3. *Let $\langle I, \leq \rangle$ be a κ-directed partial order. If $D \subseteq I$ is dense then $\langle D, \leq \rangle$ is κ-directed.*

Proof. Let $X \subseteq D$ be a set of size $|X| < \kappa$. Since I is κ-directed, it contains an upper bound l of X. As D is dense we can find an element $m \in {\Uparrow} l \cap D$. Hence, D contains an upper bound m of X. □

If we partition a κ-directed set into less than κ pieces, one of them is dense and, hence, κ-directed.

Proposition 3.4. *Let $\langle I, \leq \rangle$ be a κ-directed partial order. If $(J_\alpha)_{\alpha<\lambda}$ is a family of subsets $J_\alpha \subseteq I$ of size $\lambda < \kappa$ such that $\bigcup_{\alpha<\lambda} J_\alpha = I$, then at least one set J_α is dense.*

Proof. For $i \in I$, set

$$A_i := \{\, \alpha < \lambda \mid {\Uparrow} i \cap J_\alpha \neq \varnothing \,\}\,,$$
$$U_i := \{\, \alpha < \lambda \mid \alpha \in A_l\,,\ \text{for all } l \geq i \,\}\,.$$

Clearly, if there is some index $\alpha < \lambda$ such that $\alpha \in U_i$, for every i, then the set J_α is dense in I.

To find such an index we first prove that $U_i \neq \varnothing$, for all i. For a contradiction, suppose that there is some $i \in I$ with $U_i = \varnothing$. Then we can find, for every $\alpha < \lambda$, an element $l_\alpha \geq i$ such that ${\Uparrow} l_\alpha \cap J_\alpha = \varnothing$. Let m be an upper bound of $\{\, l_\alpha \mid \alpha < \lambda \,\}$ in I. Then $m \notin J_\alpha$, for all α. A contradiction.

To conclude the proof it is sufficient to show that $U_i = U_j$, for all $i, j \in I$. Fix some $l \geq i, j$. Then we have

$$U_i = \bigcap_{m \in {\Uparrow} i} A_m \subseteq \bigcap_{m \in {\Uparrow} l} A_m = U_l\,.$$

Conversely, suppose that there were an element $\alpha \in U_l \setminus U_i$. Then we could find some $m \geq i$ such that ${\Uparrow} m \cap J_\alpha = \varnothing$. For $s \geq m, l$, this would imply that $\alpha \notin A_s \supseteq U_l$. A contradiction. Hence, we have $U_i = U_l = U_j$, as desired. □

Directed sets can be regarded as generalisations of chains. Surprisingly in many cases it suffices to consider chains even if the use of a directed set might be more convenient. Before giving examples, let us present two technical results. The first one allows us to extend an arbitrary set to a directed one. In Section B4.4 below we will generalise this lemma to κ-directed sets, where the situation is more complicated.

Lemma 3.5. *Let $\langle I, \leq \rangle$ be a directed partial order. For every $X \subseteq I$ there exists a directed subset $D \subseteq I$ with $X \subseteq D$ and $|D| \leq |X| \oplus \aleph_0$.*

Proof. Set

$$F := \{ s \subseteq X \mid s \neq \varnothing \text{ finite} \} .$$

For every $s \in F$, we choose elements $a_s \in I$, by induction on $|s|$, as follows. Let

$$u_s := s \cup \{ a_v \mid v \subset s \} .$$

If u_s has a greatest element b then we set $a_s := b$. Otherwise, since u_s is finite and I is directed we can find an element $a_s \in I$ with $u_s \subseteq \Downarrow a_s$.

After having defined the elements a_s we can set

$$D := X \cup \{ a_s \mid s \in F \} . \qquad \square$$

Proposition 3.6. *Let $\mathfrak{I}$ be an infinite directed set of cardinality $\kappa := |I|$. There exists a chain $(H_\alpha)_{\alpha<\kappa}$ of directed subsets $H_\alpha \subseteq I$ of size $|H_\alpha| < \kappa$ such that $I = \bigcup_{\alpha<\kappa} H_\alpha$.*

Proof. Fix an enumeration $(i_\alpha)_{\alpha<\kappa}$ of I. We define H_α by induction on α. Set $H_0 := \varnothing$ and $H_\delta := \bigcup_{\alpha<\delta} H_\alpha$, for limit ordinals δ. For the successor step, we use Lemma 3.5 to choose a directed set $H_{\alpha+1} \supseteq H_\alpha \cup \{ i_\alpha \}$ of size $|H_{\alpha+1}| \leq |H_\alpha| \oplus \aleph_0$.

Each set H_α is directed. Furthermore, $i_\alpha \in H_{\alpha+1}$ implies $\bigcup_\alpha H_\alpha = I$. It remains to show that $|H_\alpha| < \kappa$. By induction on α, we prove the stronger claim that $|H_\alpha| \leq |\alpha|$, for every infinite ordinal α.

For $\alpha = \omega$, we have

$$|H_\omega| = \sup\{\, |H_n| \mid n < \omega \,\} \leq \aleph_0 \,.$$

Analogously, for limit ordinals δ,

$$|H_\delta| = \sup\{\, |H_\alpha| \mid \alpha < \delta \,\} \leq |\delta| \,.$$

Finally, we have $|H_{\alpha+1}| \leq |H_\alpha| \oplus \aleph_0 \leq |\alpha| \oplus \aleph_0 = |\alpha+1|$, for $\omega \leq \alpha < \kappa$. □

We will give several examples of how to use Proposition 3.6 to replace directed sets by chains.

Proposition 3.7. *Let $\langle A, \leq \rangle$ be a partial order. The following statements are equivalent:*

(1) *A is inductively ordered.*

(2) *Every nonempty directed set $I \subseteq A$ has a supremum.*

Proof. The direction (2) ⇒ (1) is trivial since every chain is directed. We prove the converse by induction on $\kappa := |I|$. Since every finite directed set has a greatest element we may assume that I is infinite. Let $(H_\alpha)_\alpha$ be the sequence of directed sets from Proposition 3.6. By inductive hypothesis, the suprema $a_\alpha := \sup H_\alpha$ exist. Since $(a_\alpha)_{\alpha<\kappa}$ is a chain it follows that $\sup I = \sup_\alpha a_\alpha$ exists as well. □

Lemma 3.8. *Let c be a closure operator on A. The following statements are equivalent:*

(1) *c has finite character.*

(2) *$c(\bigcup C) = \bigcup C$, for every chain $C \subseteq \operatorname{fix} c$.*

(3) *$c(\bigcup I) = \bigcup I$, for every directed set $I \subseteq \operatorname{fix} c$.*

Proof. (1) ⇒ (2) was proved in Lemma A2.4.6.

(2) ⇒ (3) We prove the claim by induction on $\kappa := |I|$. If I is finite then $\bigcup I = X$, for some $X \in I$, and we are done. Hence, we may assume that I is infinite. Let $(H_\alpha)_\alpha$ be the sequence of directed sets from Proposition 3.6.

By inductive hypothesis, we know that $X_\alpha := \bigcup H_\alpha \in \mathrm{fix}\, c$. Since $(X_\alpha)_{\alpha<\kappa}$ is a chain it follows that $\bigcup I = \bigcup_\alpha X_\alpha \in \mathrm{fix}\, c$, as desired.

(3) ⇒ (1) Let $X \subseteq A$ and set $I := \{\, c(X_0) \mid X_0 \subseteq X \text{ is finite} \,\}$. We have to show that $c(X) = \bigcup I$. For one direction, note that $X_0 \subseteq X$ implies that $c(X_0) \subseteq c(X)$. Consequently, we have $\bigcup I \subseteq c(X)$.

For the converse, note that I is directed since $c(X_0), c(X_1) \in I$ implies that $c(X_0 \cup X_1) \in I$ and we have $c(X_i) \subseteq c(X_0 \cup X_1)$. By (3), it follows that $\bigcup I \in \mathrm{fix}\, c$. Therefore,

$$\begin{aligned} X &= \bigcup \{\, X_0 \mid X_0 \subseteq X \text{ is finite} \,\} \\ &\subseteq \bigcup \{\, c(X_0) \mid X_0 \subseteq X \text{ is finite} \,\} = \bigcup I \end{aligned}$$

implies that $c(X) \subseteq c(\bigcup I) = \bigcup I$. □

Lemma 3.9. *Let $f : A \to B$ a function between partial orders where A is complete. The following statements are equivalent:*

(1) *f is continuous.*

(2) $\sup f[I] = f(\sup I)$, *for every directed set $I \subseteq A$.*

Proof. Again the direction (2) ⇒ (1) is trivial. We prove the converse by induction on $\kappa := |I|$. Since every finite directed set has a greatest element we may assume that I is infinite. Let $(H_\alpha)_\alpha$ be the sequence of directed sets from Proposition 3.6. The set

$$C := \{\, \sup H_\alpha \mid \alpha < \kappa \,\}$$

is a chain with $\sup C = \sup I$. Since f is continuous it follows that

$$\sup f[I] = \sup f[C] = f(\sup C) = f(\sup I)\,. \qquad \square$$

Having defined directed sets, we can introduce directed colimits. The systems we want to map to their colimit consist of a directed partial order of Σ-structures where each inclusion is labelled by a homomorphism specifying how the smaller structure is included in the larger one. Although we will mainly be interested in Σ-structures, we give the definition in a general category-theoretic setting.

Definition 3.10. Let $\mathcal{I}$ be a small category and $\mathcal{C}$ an arbitrary category. A *diagram* over $\mathcal{I}$ is a functor $D : \mathcal{I} \to \mathcal{C}$. If $\mathcal{I}$ is a κ-directed partial order, we call D a *κ-directed diagram*. The *size* of D is the cardinal $|\mathcal{I}^{\mathrm{mor}}|$.

Remark. In the case where the index category $\mathcal{I}$ is a partial order, a diagram $D : \mathcal{I} \to \mathcal{C}$ consists of objects $D(i) \in \mathcal{C}$, for $i \in I$, and morphisms

$$D(i,k) : D(i) \to D(k)\,, \quad \text{for } i \leq k\,,$$

such that

$$D(i,i) = \mathrm{id}_{D(i)} \quad \text{and} \quad D(k,l) \circ D(i,k) = D(i,l)\,,$$

for all $i \leq k \leq l$.

Before giving the general category-theoretic definition of a κ-directed colimit, let us present the special case of Σ-structures.

Definition 3.11. Let $D : \mathfrak{I} \to \mathfrak{Hom}(\Sigma)$ be a directed diagram. The *directed colimit* of D is the Σ-structure

$$\varinjlim D$$

where the domain of sort s is the set $(\sum_i D(i)_s)/\sim$ obtained from the disjoint union of the domains $D(i)_s$ by factorising by the relation

$$\langle i, a\rangle \sim \langle j, b\rangle \quad :\text{iff} \quad D(i,k)(a) = D(j,k)(b) \quad \text{for some } k \geq i, j\,.$$

That is, we identify $a \in D(i)$ and $b \in D(j)$ iff they are mapped to the same element in some $D(k)$.

We denote by $[i, a]$ the $\sim$-class of $\langle i, a\rangle$. The relations and functions are defined by

$$R := \big\{ \langle [i, a_0], \dots, [i, a_{n-1}]\rangle \bigm| \langle a_0, \dots, a_{n-1}\rangle \in R^{D(i)} \big\}\,,$$

$$\text{and} \quad f([i, a_0], \dots, [i, a_{n-1}]) := \big[i, f^{D(i)}(a_0, \dots, a_{n-1})\big]\,.$$

(Note that is it sufficient to consider elements $[i_0, a_0], \dots, [i_{n-1}, a_{n-1}]$ where $i_0 = \dots = i_{n-1}$.)

Remark. Directed colimits are also called *direct limits* in the literature. We will not use this term to avoid confusion with directed limits, which we will introduce below.

Example. Let $\mathfrak{Z} := \langle \mathbb{Z}, + \rangle$ be the group of integers.

(a) We define a directed diagram $D : \omega \to \mathfrak{Hom}(+)$ by $D(n) := \mathfrak{Z}$, for all n, and

$$D(k,n) : \mathfrak{Z} \to \mathfrak{Z} : z \mapsto 2^{n-k} z\,, \quad \text{for } k \leq n\,.$$

Its colimit is the structure $\varinjlim D = \langle \mathbb{Q}_2, + \rangle$ where

$$\mathbb{Q}_2 := \{\, m/2^k \mid m \in \mathbb{Z}\,,\ k \in \mathbb{N} \,\}$$

is the set of dyadic numbers.

(b) If, instead, we use the homomorphisms

$$D(k,n) : \mathfrak{Z} \to \mathfrak{Z} : z \mapsto \frac{n!}{k!} z\,, \quad \text{for } k \leq n\,,$$

then the colimit $\varinjlim D = \langle \mathbb{Q}, + \rangle$ is the group of rationals.

Remark. If the directed set $\mathfrak{I}$ has a greatest element k, then we have $\varinjlim D \cong D(k)$.

Exercise 3.1. Let $D : \mathfrak{I} \to \mathfrak{Hom}(\Sigma)$ be a directed diagram and $S \subseteq I$ dense. Prove that

$$\varinjlim D \cong \varinjlim\,(D \restriction S)\,,$$

where $D \restriction S : \mathfrak{I}|_S \to \mathfrak{Hom}(\Sigma)$ is the restriction of D to S.

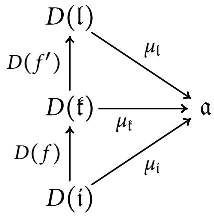
Directed colimits can also be characterised in category-theoretical terms via so-called limiting cocones. We use this property to define directed colimits in an arbitrary category.

Definition 3.12. Let $D : \mathcal{I} \to \mathcal{C}$ be a diagram.

(a) A *cocone* from D to an object $\mathfrak{a} \in \mathcal{C}$ is a family $\mu = (\mu_\mathfrak{i})_{\mathfrak{i} \in \mathcal{I}^{\text{obj}}}$ of morphisms $\mu_\mathfrak{i} : D(\mathfrak{i}) \to \mathfrak{a}$ such that

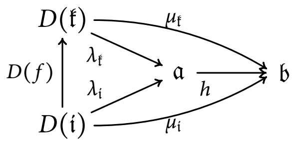
$$\mu_\mathfrak{k} \circ D(f) = \mu_\mathfrak{i}\,,$$

for all $f : \mathfrak{i} \to \mathfrak{k}$ in $\mathcal{I}^{\text{mor}}$.



(b) A cocone λ from D to $\mathfrak{a}$ is *limiting* if, for every cocone μ from D to some object $\mathfrak{b}$, there exists a unique morphism $h : \mathfrak{a} \to \mathfrak{b}$ with

$$\mu_\mathfrak{i} = h \circ \lambda_\mathfrak{i}\,, \quad \text{for all } \mathfrak{i} \in \mathcal{I}\,.$$



(Thus, limiting cocones are precisely the initial objects in the category of all cocones of D.)

(c) An object $\mathfrak{a} \in \mathcal{C}$ is a *colimit* of D if there exists a limiting cocone from D to $\mathfrak{a}$. We denote the colimit of D by $\varinjlim D$.

(d) We say that a category $\mathcal{C}$ *has κ-directed colimits* if all κ-directed diagrams $D : \mathfrak{J} \to \mathcal{C}$ have a colimit.

Example. Let $\mathfrak{L}$ be a partial order and $D : \mathcal{I} \to \mathfrak{L}$ a diagram.

(a) There exists a cocone from D to an element $a \in L$ if, and only if, a is an upper bound of $\operatorname{rng} D$.

(b) An element $a \in L$ is a colimit of D if, and only if, $a = \sup \operatorname{rng} D$.

Remark. (a) Equivalently, we could define a cocone from D to $\mathfrak{a}$ to be a natural transformation μ from D to the *diagonal functor* $\Delta(\mathfrak{a}) : \mathcal{I} \to \mathcal{C}$ with

$$\Delta(\mathfrak{a})(\mathfrak{i}) = \mathfrak{a}\,, \qquad \text{for all } \mathfrak{i} \in \mathcal{I}^{\text{obj}},$$
and $$\Delta(\mathfrak{a})(f) = \mathrm{id}_{\mathfrak{a}}\,, \qquad \text{for all } f \in \mathcal{I}^{\text{mor}}.$$

$$\begin{array}{ccc} D(\mathfrak{l}) & \xrightarrow{h_{\mathfrak{l}}} & \mathfrak{a} \\ {\scriptstyle D(f')}\uparrow & & \uparrow{\scriptstyle \mathrm{id}_{\mathfrak{a}}} \\ D(\mathfrak{k}) & \xrightarrow{h_{\mathfrak{k}}} & \mathfrak{a} \\ {\scriptstyle D(f)}\uparrow & & \uparrow{\scriptstyle \mathrm{id}_{\mathfrak{a}}} \\ D(\mathfrak{i}) & \xrightarrow{h_{\mathfrak{i}}} & \mathfrak{a} \end{array}$$

(b) Not that, by the uniqueness of h in the definition of a limiting cocone, colimits are unique up to isomorphism. As limiting cocones are initial objects in the category of all cocones, this also follows directly from Lemma B1.3.7.

According to the next lemma, the colimit $\varinjlim D$ of a directed diagram $D : \mathfrak{I} \to \mathfrak{Hom}(\Sigma)$ of Σ-structures coincides with the category-theoretical notion of a colimit.

Lemma 3.13. *Every κ-directed diagram $D : \mathfrak{I} \to \mathfrak{Hom}(\Sigma)$ has a limiting cocone λ from D to $\varinjlim D$.*

Proof. Let $\mathfrak{A} := \varinjlim D$ and $[i, a]$ be the $\sim$-class of $\langle i, a\rangle$. We claim that the functions

$$\lambda_i : D(i) \to \mathfrak{A} : a \mapsto [i, a]\,, \qquad \text{for } i \in I\,,$$

form a limiting cocone. Let $a \in D(i)$ and $j \geq i$. By definition, we have $\langle j, D(i,j)(a)\rangle \sim \langle i, a\rangle$. Hence,

$$\lambda_i(a) = [i, a] = [j, D(i,j)(a)] = \lambda_j(D(i,j)(a))\,,$$

and $(\lambda_i)_{i \in I}$ is a cocone.

To show that it is limiting, suppose that μ is a cocone from D to $\mathfrak{B}$. We define the desired homomorphism $h : \mathfrak{A} \to \mathfrak{B}$ by

$$h[i, a] := \mu_i(a)\,.$$

h is obviously the unique function such that $h \circ \lambda_i = \mu_i$. Therefore, it remains to show that h is well-defined. Suppose that $\langle i, a \rangle \sim \langle j, b \rangle$. Then there is some $k \geq i, j$ with $D(i,k)(a) = D(j,k)(b)$. Hence, we have

$$\begin{aligned} h[i,a] = \mu_i(a) &= (\mu_k \circ D(i,k))(a) \\ &= (\mu_k \circ D(j,k))(b) = \mu_j(b) = h[j,b]. \end{aligned}$$ □

Corollary 3.14. *$\mathfrak{Hom}(\Sigma)$ has κ-directed colimits, for all infinite cardinals κ.*

Exercise 3.2. Prove that the functions λ_i and h defined in the proof above are homomorphisms.

Let us give several applications of the notion of a directed colimit.

Definition 3.15. Let $\mathfrak{A}$ be a structure and κ a cardinal. A substructure $\mathfrak{A}_0 \subseteq \mathfrak{A}$ is *κ-generated* if $\mathfrak{A}_0 = \langle\!\langle X \rangle\!\rangle_{\mathfrak{A}}$, for some set X of size $|X| < \kappa$.

Proposition 3.16. *Let κ be a regular cardinal. Every structure $\mathfrak{A}$ is the κ-directed colimit of its κ-generated substructures.*

Proof. Let $I := \{ \langle\!\langle X \rangle\!\rangle_{\mathfrak{A}} \mid |X| < \kappa \}$ be the set of all κ-generated substructures of $\mathfrak{A}$. If $(\langle\!\langle X_i \rangle\!\rangle_{\mathfrak{A}})_{i \leq \alpha} \in I^\alpha$, for $\alpha < \kappa$, then $\langle\!\langle \bigcup_i X_i \rangle\!\rangle_{\mathfrak{A}} \in I$ since κ is regular. Consequently, $\langle I, \subseteq \rangle$ is κ-directed.

For $\mathfrak{C} \in I$, set $D(\mathfrak{C}) := \mathfrak{C}$ and let $D(\mathfrak{B}, \mathfrak{C}) : \mathfrak{B} \to \mathfrak{C}$, for $\mathfrak{B} \subseteq \mathfrak{C}$ in I, be the inclusion map. Then

$$\mathfrak{A} \cong \varinjlim D.$$ □

Lemma 3.17. *Every reduced product $\prod_{i \in I} \mathfrak{A}^i / \mathfrak{u}$ is the directed colimit of products $\prod_{i \in s} \mathfrak{A}^i$ with $s \in \mathfrak{u}$.*

Proof. For $s \in \mathfrak{u}$, set $D(s) := \prod_{i \in s} \mathfrak{A}^i$. We order $\mathfrak{u}$ by inverse inclusion. For $s \supseteq t$ in $\mathfrak{u}$, let

$$D(s,t) : D(s) \to D(t) : (a^i)_{i \in s} \mapsto (a^i)_{i \in t}$$

by the canonical projection. We claim that

$$\varinjlim D \cong \prod_{i \in I} \mathfrak{A}^i / \mathfrak{u} .$$

Note that, if $(a^i)_{i \in I} \in \prod_{i \in I} A^i$ and $s, t \in \mathfrak{u}$ then we have

$$[s, (a^i)_{i \in s}] = [t, (a^i)_{i \in s}]$$

since $(a^i)_{i \in s \cap t} = (a^i)_{i \in s \cap t}$ and $s \cap t \in \mathfrak{u}$. Consequently, we can define a function $\varphi : \prod_i \mathfrak{A}^i / \mathfrak{u} \to \varinjlim D$ by

$$\varphi\big([(a^i)_i]_{\mathfrak{u}}\big) := [s, (a^i)_{i \in s}] , \quad \text{for some/all } s \in \mathfrak{u} .$$

It is easy to check that φ is the desired isomorphism. □

The dual notion to a directed colimit is a directed limit.

Definition 3.18. Let $\mathfrak{J}$ be a directed partial order.

(a) An *inverse diagram* over $\mathfrak{J}$ is a functor $D : \mathfrak{J}^{\mathrm{op}} \to \mathcal{C}$.

(b) The *directed limit* of an inverse diagram $D : \mathfrak{J}^{\mathrm{op}} \to \mathfrak{Hom}(\Sigma)$ is the Σ-structure

$$\varprojlim D := \big(\prod_i \mathfrak{A}^i\big)\big|_U$$

obtained from the product of the $\mathfrak{A}^i$ by restriction to the set

$$U := \big\{ (a_i)_i \in \prod_i A^i \bigm| a_i = D(i, j)(a_j) \text{ for all } i \leq j \big\} .$$

Remark. Directed limits are also called *inverse limits.*

Example. (a) Let $D : \mathfrak{J} \to \mathfrak{Hom}(\Sigma)$ be a chain. If we reverse the order of the index set I, this chain becomes an inverse diagram whose limit is isomorphic to the intersection of the $D(i)$, that is,

$$\varprojlim D \cong D(k)|_C$$

where $C := \bigcap_i D(i)$ and $k \in I$ is arbitrary.

(b) Let $\mathfrak{K}$ be a field and $D(n) := \mathfrak{K}[x]/(x^n)$, for $n < \omega$, the ring of polynomials over $\mathfrak{K}$ of degree less than n. The directed limit $\varprojlim D \cong \mathfrak{K}[[x]]$ is isomorphic to the ring of formal power series over $\mathfrak{K}$.

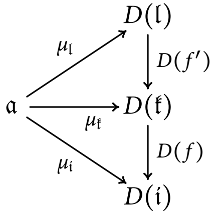
As above we can characterise inverse limits in category-theoretical terms.

Definition 3.19. Let $D : \mathcal{I}^{\mathrm{op}} \to \mathcal{C}$ be an inverse diagram.

(a) A *cone* from an object $\mathfrak{a} \in \mathcal{C}$ to D is a family $\mu = (\mu_{\mathfrak{i}})_{\mathfrak{i} \in \mathcal{I}^{\mathrm{obj}}}$ of morphisms $\mu_{\mathfrak{i}} : \mathfrak{a} \to D(\mathfrak{i})$ such that

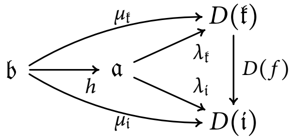
$$D(f) \circ \mu_{\mathfrak{k}} = \mu_{\mathfrak{i}} ,$$

for all $f : \mathfrak{i} \to \mathfrak{k}$ in $\mathcal{I}^{\mathrm{mor}}$.



(b) A cone λ to $\mathfrak{a}$ is *limiting* if, for every cone μ from some object $\mathfrak{b}$ to D, there exists a unique morphism $h : \mathfrak{b} \to \mathfrak{a}$ with

$$\mu_{\mathfrak{i}} = \lambda_{\mathfrak{i}} \circ h , \quad \text{for all } \mathfrak{i} \in \mathcal{I} .$$



(Thus, limiting cones are precisely the terminal objects in the category of all cones of D.)

(c) An object $\mathfrak{a} \in \mathcal{C}$ is a *limit* of D if there exists a limiting cone from $\mathfrak{a}$ to D.

Lemma 3.20. *Every κ-directed inverse diagram $D : \mathfrak{J}^{\mathrm{op}} \to \mathfrak{Hom}(\Sigma)$ has a limiting cone from $\varprojlim D$ to D.*

Exercise 3.3. Prove Lemma 3.20.

Exercise 3.4. Let $\mathcal{I}$ be a category where the only morphisms are the identity morphisms. Show that the limit of a diagram $D : \mathcal{I} \to \mathfrak{Hom}(\Sigma)$ is isomorphic to the direct product

$$\prod_{i \in \mathcal{I}} D(i) .$$

4. *Equivalent diagrams*

In this section we study the question of when two diagrams have the same colimit. Our aim is, given a diagram $D : \mathcal{I} \to \mathcal{C}$ to find a diagram $E : \mathcal{J} \to \mathcal{C}$ with the same colimit where the index category $\mathcal{J}$ is simpler in one way or another. We start by developing methods to prove that two diagrams have the same colimit. These methods are based on the notion of a cocone functor.

Definition 4.1. Let $\mathcal{C}$ be a category.

(a) Let μ be a cocone from $D : \mathcal{I} \to \mathcal{C}$ to some object $\mathfrak{a}$. For a morphism $f : \mathfrak{a} \to \mathfrak{b}$, we define

$$f * \mu := (f \circ \mu_i)_{i \in \mathcal{I}} .$$

(b) The *cocone functor* $\mathrm{Cone}(D, -) : \mathcal{C} \to \mathfrak{Set}$ associated with a diagram $D : \mathcal{I} \to \mathcal{C}$ maps

- objects $\mathfrak{a}$ to the set $\mathrm{Cone}(D, \mathfrak{a})$ of all cocones from D to $\mathfrak{a}$, and
- morphisms $f : \mathfrak{a} \to \mathfrak{b}$ to the function

$$\mathrm{Cone}(D, f) : \mathrm{Cone}(D, \mathfrak{a}) \to \mathrm{Cone}(D, \mathfrak{b}) : \mu \mapsto f * \mu .$$

(c) The *covariant hom-functor* associated with an object $\mathfrak{a} \in \mathcal{C}$ is the functor

$$\mathcal{C}(\mathfrak{a}, -) : \mathcal{C} \to \mathfrak{Set}$$

mapping an object $\mathfrak{b} \in \mathcal{C}$ to the set $\mathcal{C}(\mathfrak{a}, \mathfrak{b})$ of all morphisms from $\mathfrak{a}$ to $\mathfrak{b}$ and mapping a morphism $f : \mathfrak{b} \to \mathfrak{c}$ to the function

$$\mathcal{C}(\mathfrak{a}, f) : \mathcal{C}(\mathfrak{a}, \mathfrak{b}) \to \mathcal{C}(\mathfrak{a}, \mathfrak{c}) : g \mapsto f \circ g .$$

Given a functor $F : \mathcal{C} \to \mathcal{D}$ and an object $\mathfrak{b} \in \mathcal{D}$, we will abbreviate $\mathcal{D}(\mathfrak{b}, -) \circ F$ by $\mathcal{D}(\mathfrak{b}, F-)$.

Remark. In this terminology a limiting cocone of D is an element $\lambda \in \mathrm{Cone}(D, \mathfrak{a})$ such that, for every $\mu \in \mathrm{Cone}(D, \mathfrak{b})$, there exists a unique morphism $f : \mathfrak{a} \to \mathfrak{b}$ with $\mu = f * \lambda$.

We start with a characterisation of limiting cocones in terms of the cocone functor.

Lemma 4.2. *Let $D : \mathcal{I} \to \mathcal{C}$ be a diagram. A cocone $\lambda \in \mathrm{Cone}(D, \mathfrak{a})$ is limiting if, and only if, the family $\eta = (\eta_{\mathfrak{b}})_{\mathfrak{b} \in \mathcal{C}}$ of morphisms defined by*

$$\eta_{\mathfrak{b}} : \mathcal{C}(\mathfrak{a}, \mathfrak{b}) \to \mathrm{Cone}(D, \mathfrak{b}) : f \mapsto f * \lambda$$

is a natural isomorphism $\eta : \mathcal{C}(\mathfrak{a}, -) \cong \mathrm{Cone}(D, -)$.

Proof. ($\Leftarrow$) Suppose that η is a natural isomorphism. To show that λ is limiting, consider a cocone $\mu \in \mathrm{Cone}(D, \mathfrak{b})$. Setting $h := \eta_{\mathfrak{b}}^{-1}(\mu)$, we obtain the desired equation

$$\mu = \eta_{\mathfrak{b}}(h) = h * \lambda\,.$$

To conclude the proof, let $h' : \mathfrak{a} \to \mathfrak{b}$ be a second morphism with $\mu = h' * \lambda$. Then $\eta_{\mathfrak{b}}(h') = \mu = \eta_{\mathfrak{b}}(h)$ implies, by injectivity of $\eta_{\mathfrak{b}}$, that $h' = h$.

($\Rightarrow$) We start by showing that η is a natural transformation. Let $f : \mathfrak{a} \to \mathfrak{b}$ and $g : \mathfrak{b} \to \mathfrak{c}$ be morphisms. Then

$$\begin{aligned}\eta_{\mathfrak{b}}(\mathcal{C}(\mathfrak{a}, g)(f)) &= \eta_{\mathfrak{b}}(g \circ f) \\ &= (g \circ f) * \lambda \\ &= g * (f * \lambda) = \mathrm{Cone}(D, g)(\eta_{\mathfrak{b}}(f))\,.\end{aligned}$$

Now, suppose that λ is limiting. We claim that $\eta_{\mathfrak{b}}$ is bijective. For surjectivity, let $\mu \in \mathrm{Cone}(D, \mathfrak{b})$. As λ is limiting, there exists a unique morphism $h : \mathfrak{a} \to \mathfrak{b}$ such that $\mu = h * \lambda$. Hence, $\mu = \eta_{\mathfrak{b}}(h) \in \mathrm{rng}\, \eta_{\mathfrak{b}}$.

For injectivity, let $f, f' : \mathfrak{a} \to \mathfrak{b}$ be morphisms with $\eta_{\mathfrak{b}}(f) = \eta_{\mathfrak{b}}(f')$. We set $\mu := \eta_{\mathfrak{b}}(f)$. Since λ is limiting, there exists a unique morphism $h : \mathfrak{a} \to \mathfrak{b}$ such that $\mu = h * \lambda$. As

$$f * \lambda = \eta_{\mathfrak{b}}(f) = \mu = \eta_{\mathfrak{b}}(f') = f' * \lambda\,,$$

it follows by uniqueness of h that $f = h = f'$. □

The following lemma is our main tool to prove that two diagrams have the same colimit.

Lemma 4.3. *Let $D : \mathcal{I} \to \mathcal{C}$ and $E : \mathcal{J} \to \mathcal{C}$ be diagrams. Every natural isomorphism $\eta : \mathrm{Cone}(D, -) \cong \mathrm{Cone}(E, -)$ maps limiting cocones of D to limiting cocones of E.*

Proof. Let $\lambda \in \mathrm{Cone}(D, \mathfrak{a})$ be a limiting cocone of D. Then $\eta_{\mathfrak{a}}(\lambda) \in \mathrm{Cone}(E, \mathfrak{a})$ is a cocone from E to $\mathfrak{a}$. It remains to prove that it is limiting. Given an arbitrary cocone $\mu \in \mathrm{Cone}(E, \mathfrak{b})$, the preimage $\eta_{\mathfrak{b}}^{-1}(\mu)$ is a cocone from D to $\mathfrak{b}$. As λ is limiting, there exists a unique morphism $h : \mathfrak{a} \to \mathfrak{b}$ such that

$$\eta_{\mathfrak{b}}^{-1}(\mu) = h * \lambda = \mathrm{Cone}(D, h)(\lambda)\,.$$

Applying $\eta_{\mathfrak{b}}$ to this equation, we obtain

$$\mu = \eta_{\mathfrak{b}}\big(\mathrm{Cone}(D, h)(\lambda)\big) = \mathrm{Cone}(E, h)(\eta_{\mathfrak{a}}(\lambda)) = h * \eta_{\mathfrak{a}}(\lambda)\,,$$

as desired. Furthermore, if $h' : \mathfrak{a} \to \mathfrak{b}$ is another morphism satisfying $\mu = h' * \eta_{\mathfrak{a}}(\lambda)$, then

$$\eta_{\mathfrak{b}}^{-1}(\mu) = \eta_{\mathfrak{b}}^{-1}\big(\mathrm{Cone}(E, h')(\eta_{\mathfrak{a}}(\lambda))\big) = \mathrm{Cone}(D, h')(\lambda) = h' * \lambda\,,$$

and it follows by uniqueness of h that $h' = h$. □

Below we will frequently simplify a diagram $D : \mathcal{I} \to \mathcal{C}$ by finding a functor $F : \mathcal{J} \to \mathcal{I}$ such that $D \circ F$ has the same colimit as D and the index category $\mathcal{J}$ is simpler than $\mathcal{I}$. To study the colimit of such a composition $D \circ F$, we introduce two natural transformations $\pi_{D,F}$ and $\tau_{D,F}$.

Definition 4.4. Let $D : \mathcal{I} \to \mathcal{C}$ be a diagram.

(a) The *projection* $\pi_{D,F}$ along a functor $F : \mathcal{J} \to \mathcal{I}$ is the function mapping a cocone μ of D to the family $(\mu_{F(\mathfrak{j})})_{\mathfrak{j} \in \mathcal{J}}$.

(b) The *translation* $\tau_{G,D}$ by a functor $G : \mathcal{C} \to \mathcal{D}$ is the function mapping a cocone μ of D to the family $G[\mu] := (G(\mu_{\mathfrak{i}}))_{\mathfrak{i} \in \mathcal{I}}$.

Lemma 4.5. *Let $D : \mathcal{I} \to \mathcal{C}$ be a diagram.*

(a) *The projection along a functor $F : \mathcal{J} \to \mathcal{I}$ is a natural transformation*

$$\pi_{D,F} : \mathrm{Cone}(D, -) \to \mathrm{Cone}(D \circ F, -)\,.$$

(b) *The translation by a functor $G : \mathcal{C} \to \mathcal{D}$ is a natural transformation*

$$\tau_{G,D} : \mathrm{Cone}(D, -) \to \mathrm{Cone}(G \circ D, G-)\,.$$

(c) *For diagrams $F : \mathcal{J} \to \mathcal{I}$ and $G : \mathcal{K} \to \mathcal{J}$,*

$$\pi_{D,F\circ G} = \pi_{D\circ F,G} \circ \pi_{D,F}\,.$$

Proof. (a) Given a cocone μ from D to $\mathfrak{a}$, the image $\pi_{D,F}(\mu)$ is clearly a cocone from $D \circ F$ to $\mathfrak{a}$. Hence, it remains to prove that $\pi_{D,F}$ is natural. Let $f : \mathfrak{a} \to \mathfrak{b}$ be a morphism of $\mathcal{C}$ and $\mu \in \mathrm{Cone}(D, \mathfrak{a})$ a cocone. Then

$$\begin{aligned}\pi_{D,F}(\mathrm{Cone}(D, f)(\mu)) &= \pi_{D,F}((f \circ \mu_{\mathfrak{i}})_{\mathfrak{i}\in\mathcal{I}}) \\ &= (f \circ \mu_{F(\mathfrak{j})})_{\mathfrak{j}\in\mathcal{J}} \\ &= \mathrm{Cone}(D \circ F, f)(\pi_{D,F}(\mu))\,.\end{aligned}$$

(b) Given a cocone μ from D to $\mathfrak{a}$, the image $\tau_{G,D}(\mu)$ is clearly a cocone from $G \circ D$ to $G(\mathfrak{a})$. Hence, it remains to prove that $\tau_{G,D}$ is natural. Let $f : \mathfrak{a} \to \mathfrak{b}$ be a morphism of $\mathcal{C}$ and $\mu \in \mathrm{Cone}(D, \mathfrak{a})$ a cocone. Then

$$\begin{aligned}\tau_{G,D}(\mathrm{Cone}(D, f)(\mu)) &= \tau_{G,D}((f \circ \mu_{\mathfrak{i}})_{\mathfrak{i}\in\mathcal{I}}) \\ &= \big(G(f) \circ G(\mu_{\mathfrak{i}})\big)_{\mathfrak{i}\in\mathcal{I}} \\ &= G(f) * G[\mu] \\ &= \mathrm{Cone}(G \circ D, G(f))(\tau_{G,D}(\mu))\,.\end{aligned}$$

(c) For $\mu \in \mathrm{Cone}(D, \mathfrak{a})$, we have

$$\begin{aligned}\pi_{D\circ F,G}(\pi_{D,F}(\mu)) &= \pi_{D\circ F,G}\big((\mu_{F(\mathfrak{i})})_{\mathfrak{i}\in\mathcal{I}}\big) \\ &= (\mu_{F(G(\mathfrak{k}))})_{\mathfrak{k}\in\mathcal{K}} = \pi_{D,F\circ G}(\mu)\,.\end{aligned}$$

∎

We extend the terminology of Definition B1.3.9 as follows.

Definition 4.6. Let $F : \mathcal{C} \to \mathcal{D}$ be a functor and let $\mathcal{P}$ be a class of diagrams.

(a) We say that F *preserves $\mathcal{P}$-colimits* if, whenever λ is a limiting cocone of a diagram $D \in \mathcal{P}$, then $F[\lambda]$ is a limiting cocone of $F \circ D$.

(b) We say that F *reflects $\mathcal{P}$-colimits* if, whenever λ is a cocone of a diagram $D \in \mathcal{P}$ such that $F[\lambda]$ is limiting, then λ is also limiting.

(c) Analogously, we define when F preserves or reflects $\mathcal{P}$-limits.

Lemma 4.7. *Let $F : \mathcal{C} \to \mathcal{D}$ be full and faithful.*

(a) *For every diagram $D : \mathcal{I} \to \mathcal{C}$,*

$$\tau_{F,D} : \mathrm{Cone}(D, -) \to \mathrm{Cone}(F \circ D, F-)$$

is a natural isomorphism.

(b) *F reflects all limits and colimits.*

Proof. (a) For injectivity, suppose that $\mu, \mu' \in \mathrm{Cone}(D, \mathfrak{a})$ are cocones with $F[\mu] = F[\mu']$. As F is faithful, $F(\mu_\mathfrak{i}) = F(\mu'_\mathfrak{i})$ implies that $\mu_\mathfrak{i} = \mu'_\mathfrak{i}$, for all $\mathfrak{i} \in \mathcal{I}$.

For surjectivity, let $\mu \in \mathrm{Cone}(F \circ D, F(\mathfrak{a}))$. As F is full, we can find morphisms $\lambda_\mathfrak{i} : D(\mathfrak{i}) \to \mathfrak{a}$, for every $\mathfrak{i} \in \mathcal{I}$, such that $F(\lambda_\mathfrak{i}) = \mu_\mathfrak{i}$. Then $F[\lambda] = \mu$ where $\lambda := (\lambda_\mathfrak{i})_{\mathfrak{i} \in \mathcal{I}}$. Hence, it remains to prove that λ is a cocone of D. Let $f : \mathfrak{i} \to \mathfrak{j}$ be a morphism of $\mathcal{I}$. Then

$$F(\lambda_\mathfrak{j} \circ D(f)) = F(\lambda_\mathfrak{j}) \circ F(D(f)) = \mu_\mathfrak{j} \circ F(D(f)) = \mu_\mathfrak{i} = F(\lambda_\mathfrak{i})$$

implies, by faithfulness of F, that $\lambda_\mathfrak{j} \circ D(f) = \lambda_\mathfrak{i}$.

(b) Let $D : \mathcal{I} \to \mathcal{C}$ be a diagram and $\lambda \in \mathrm{Cone}(D, \mathfrak{a})$ a cocone such that $F[\lambda]$ is limiting. Let

$$\eta : \mathcal{D}(F(\mathfrak{a}), -) \cong \mathrm{Cone}(F \circ D, -) : f \mapsto f * F[\lambda]$$

be the natural isomorphism of Lemma 4.2. As F is full and faithful, the natural transformation

$$\zeta : \mathcal{C}(\mathfrak{a}, -) \to \mathcal{D}(F(\mathfrak{a}), F-) : f \mapsto F(f)$$

is also a natural isomorphism. By (a), it follows that the composition

$$\tau_{F,D}^{-1} \circ \eta \circ \zeta : \mathcal{C}(\mathfrak{a}, -) \to \mathrm{Cone}(D, -)$$

is a natural isomorphism that maps a morphism $f : \mathfrak{a} \to \mathfrak{b}$ to

$$\begin{aligned}(\tau_{F,D}^{-1} \circ \eta \circ \zeta)(f) &= (\tau_{F,D}^{-1} \circ \eta)(F(f)) \\ &= \tau_{F,D}^{-1}(F(f) * F[\lambda]) \\ &= \tau_{F,D}^{-1}(F[f * \lambda]) = f * \lambda\,.\end{aligned}$$

Consequently, it follows by Lemma 4.2 that λ is limiting. □

Equivalences and skeletons

As a first application we show that isomorphic and equivalent diagrams have the same colimit.

Lemma 4.8. *Every natural isomorphism $\eta : D \cong E$ between two diagrams $D, E : \mathcal{I} \to \mathcal{J}$, induces a natural isomorphism*

$$\zeta : \mathrm{Cone}(D, -) \cong \mathrm{Cone}(E, -) : \mu \mapsto (\mu_{\mathfrak{i}} \circ \eta_{\mathfrak{i}}^{-1})_{\mathfrak{i} \in \mathcal{I}}\,.$$

Proof. We define ζ and its inverse ξ by

$$\begin{aligned}\zeta(\mu) &:= (\mu_{\mathfrak{i}} \circ \eta_{\mathfrak{i}}^{-1})_{\mathfrak{i} \in \mathcal{I}}\,, &&\text{for } \mu \in \mathrm{Cone}(D, \mathfrak{a})\,, \\ \xi(\mu) &:= (\mu_{\mathfrak{i}} \circ \eta_{\mathfrak{i}})_{\mathfrak{i} \in \mathcal{I}}\,, &&\text{for } \mu \in \mathrm{Cone}(E, \mathfrak{a})\,.\end{aligned}$$

To show that ζ and ξ are well-defined, let $\mu \in \mathrm{Cone}(D, \mathfrak{a})$ and let $f : \mathfrak{i} \to \mathfrak{j}$ be a morphism of $\mathcal{I}$. Then

$$\begin{aligned}\zeta(\mu)_{\mathfrak{j}} \circ E(f) &= \mu_{\mathfrak{j}} \circ \eta_{\mathfrak{j}}^{-1} \circ E(f) \\ &= \mu_{\mathfrak{j}} \circ D(f) \circ \eta_{\mathfrak{i}}^{-1} = \mu_{\mathfrak{i}} \circ \eta_{\mathfrak{i}}^{-1} = \zeta(\mu)_{\mathfrak{i}}\,.\end{aligned}$$

Hence, $\zeta(\mu)$ is a cocone of E. In the same way, one can check that

$$\xi(\mu)_{\mathfrak{j}} \circ D(f) = \xi(\mu)_{\mathfrak{i}}\,, \quad \text{for } \mu \in \mathrm{Cone}(E, \mathfrak{a}) \text{ and } f : \mathfrak{i} \to \mathfrak{j}\,.$$

Furthermore, ζ is a natural transformation since, for $\mu \in \mathrm{Cone}(D, \mathfrak{a})$ and $f : \mathfrak{a} \to \mathfrak{b}$,

$$\begin{aligned}\zeta(\mathrm{Cone}(D, f)(\mu)) &= \zeta((f \circ \mu_{\mathrm{i}})_{\mathrm{i} \in \mathcal{I}}) \\ &= (f \circ \mu_{\mathrm{i}} \circ \eta_{\mathrm{i}}^{-1})_{\mathrm{i} \in \mathcal{I}} \\ &= \mathrm{Cone}(E, f)((\mu_{\mathrm{i}} \circ \eta_{\mathrm{i}}^{-1})_{\mathrm{i} \in \mathcal{I}}) \\ &= \mathrm{Cone}(E, f)(\zeta(\mu))\,.\end{aligned}$$

Finally, note that

$$\xi(\zeta(\mu)) = \xi((\mu_{\mathrm{i}} \circ \eta_{\mathrm{i}}^{-1})_{\mathrm{i} \in \mathcal{I}}) = (\mu_{\mathrm{i}} \circ \eta_{\mathrm{i}}^{-1} \circ \eta_{\mathrm{i}})_{\mathrm{i} \in \mathcal{I}} = \mu\,,$$

and, similarly, $\zeta(\xi(\mu)) = \mu$. □

Proposition 4.9. *Let $F : \mathcal{I} \to \mathcal{J}$ be an equivalence between two small categories $\mathcal{I}$ and $\mathcal{J}$ and let $D : \mathcal{J} \to \mathcal{C}$ be a diagram. The projection*

$$\pi_{D,F} : \mathrm{Cone}(D, -) \to \mathrm{Cone}(D \circ F, -)$$

along F is a natural isomorphism.

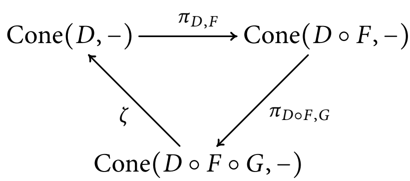
Proof. By Theorem B1.3.14, there exist a functor $G : \mathcal{J} \to \mathcal{I}$ and natural isomorphisms $\rho : G \circ F \cong \mathrm{id}_{\mathcal{I}}$ and $\eta : \mathrm{id}_{\mathcal{J}} \cong F \circ G$ such that

$$F(\rho_{\mathrm{i}}) = \eta_{F(\mathrm{i})}^{-1} \quad \text{and} \quad G(\eta_{\mathrm{j}}) = \rho_{G(\mathrm{j})}^{-1}.$$

It follows that $D[\eta^{-1}]$ is a natural isomorphism $D \circ F \circ G \cong D$ which, by Lemma 4.8, induces a natural isomorphism

$$\zeta : \mathrm{Cone}(D \circ F \circ G, -) \to \mathrm{Cone}(D, -) : \mu \mapsto (\mu_{\mathrm{j}} \circ D(\eta_{\mathrm{j}}))_{\mathrm{j} \in \mathcal{J}}\,.$$

We claim that $\zeta \circ \pi_{D \circ F, G}$ is an inverse of $\pi_{D,F}$.



For $\mu \in \mathrm{Cone}(D, \mathfrak{a})$, $\mu_{F(G(\mathrm{j}))} \circ D(\eta_\mathrm{j}) = \mu_\mathrm{j}$ implies that

$$\begin{aligned}(\zeta \circ \pi_{D\circ F,G} \circ \pi_{D,F})(\mu) &= (\zeta \circ \pi_{D\circ F,G})\big((\mu_{F(\mathrm{i})})_{\mathrm{i}\in\mathcal{I}}\big) \\ &= \zeta\big((\mu_{F(G(\mathrm{j}))})_{\mathrm{j}\in\mathcal{J}}\big) \\ &= \big(\mu_{F(G(\mathrm{j}))} \circ D(\eta_\mathrm{j})\big)_{\mathrm{j}\in\mathcal{J}} = (\mu_\mathrm{j})_{\mathrm{j}\in\mathcal{J}}\,.\end{aligned}$$

Similarly, let $\mu \in \mathrm{Cone}(D \circ F, \mathfrak{a})$. Then $\mu_\mathrm{i} \circ D(F(\rho_\mathrm{i})) = \mu_{G(F(\mathrm{i}))}$ implies that

$$\begin{aligned}(\pi_{D,F} \circ \zeta \circ \pi_{D\circ F,G})(\mu) &= (\pi_{D,F} \circ \zeta)\big((\mu_{G(\mathrm{j})})_{\mathrm{j}\in\mathcal{J}}\big) \\ &= \pi_{D,F}\big((\mu_{G(\mathrm{j})} \circ D(\eta_\mathrm{j}))_{\mathrm{j}\in\mathcal{J}}\big) \\ &= \big(\mu_{G(F(\mathrm{i}))} \circ D(\eta_{F(\mathrm{i})})\big)_{\mathrm{i}\in\mathcal{I}} \\ &= \big(\mu_{G(F(\mathrm{i}))} \circ D(F(\rho_\mathrm{i})^{-1})\big)_{\mathrm{i}\in\mathcal{I}} \\ &= (\mu_\mathrm{i})_{\mathrm{i}\in\mathcal{I}}\,.\end{aligned}$$

□

Corollary 4.10. *Let $F : \mathcal{I} \to \mathcal{J}$ be an equivalence between two small categories $\mathcal{I}$ and $\mathcal{J}$. Then*

$$\varinjlim\,(D \circ F) = \varinjlim D\,, \quad \textit{for every diagram } D : \mathcal{J} \to \mathcal{C}\,.$$

As an application of this corollary, we show how to get rid of isomorphic copies in the index category of a diagram.

Definition 4.11. A *skeleton* of a category $\mathcal{C}$ is a full subcategory $\mathcal{C}_0 \subseteq \mathcal{C}$ such that

- every object of $\mathcal{C}$ is isomorphic to some object of $\mathcal{C}_0$,
- no two objects of $\mathcal{C}_0$ are isomorphic.

Example. A skeleton of $\mathfrak{Set}$ is given by the full subcategory induced by the class Cn of all cardinals.

We will prove in the next two lemmas that skeletons are unique up to isomorphism, and that they are equivalent to the original category. Consequently, given a diagram $D : \mathcal{I} \to \mathcal{C}$, we can replace the index category $\mathcal{I}$ by its skeleton without changing the colimit.

Lemma 4.12. *If $\mathcal{C}_0$ and $\mathcal{C}_1$ are skeletons of $\mathcal{C}$, there exists an isomorphism $\mathcal{C}_0 \cong \mathcal{C}_1$.*

Proof. We define functors $F_i : \mathcal{C}_i \to \mathcal{C}_{1-i}$, for $i < 2$, as follows. For $\mathfrak{a} \in \mathcal{C}_i$, let $\mathfrak{a}^{(1-i)}$ be the unique element of $\mathcal{C}_{1-i}$ isomorphic to $\mathfrak{a}$. We fix isomorphisms $\pi^0_\mathfrak{a} : \mathfrak{a} \to \mathfrak{a}^{(1)}$, for $\mathfrak{a} \in \mathcal{C}_0^{\mathrm{obj}}$, and we set $\pi^1_\mathfrak{a} := (\pi^0_{\mathfrak{a}^{(0)}})^{-1}$. We define

$$\begin{aligned} F^i(\mathfrak{a}) &:= \mathfrak{a}^{(1-i)}, && \text{for } \mathfrak{a} \in \mathcal{C}_i^{\mathrm{obj}}, \\ F^i(f) &:= \pi^i_\mathfrak{b} \circ f \circ (\pi^i_\mathfrak{a})^{-1}, && \text{for } f : \mathfrak{a} \to \mathfrak{b} \text{ in } \mathcal{C}_i^{\mathrm{mor}}. \end{aligned}$$

We claim that $F^{1-i} \circ F^i = \mathrm{id}$. For $\mathfrak{a} \in \mathcal{C}_i^{\mathrm{obj}}$, we have

$$F^{1-i}(F^i(\mathfrak{a})) = F^{1-i}(\mathfrak{a}^{(1-i)}) = (\mathfrak{a}^{(1-i)})^{(i)} = \mathfrak{a}\,.$$

For $f : \mathfrak{a} \to \mathfrak{b}$ in $\mathcal{C}_i^{\mathrm{mor}}$, we have

$$\begin{aligned} F^{1-i}(F^i(f)) &= F^{1-i}(\pi^i_\mathfrak{b} \circ f \circ (\pi^i_\mathfrak{a})^{-1}) \\ &= \pi^{1-i}_{\mathfrak{b}^{(1-i)}} \circ \pi^i_\mathfrak{b} \circ f \circ (\pi^i_\mathfrak{a})^{-1} \circ (\pi^{1-i}_{\mathfrak{a}^{(1-i)}})^{-1} \\ &= (\pi^i_\mathfrak{b})^{-1} \circ \pi^i_\mathfrak{b} \circ f \circ (\pi^i_\mathfrak{a})^{-1} \circ \pi^i_\mathfrak{a} \\ &= f\,. \end{aligned}$$

□

Lemma 4.13. *Every skeleton $\mathcal{C}_0$ of a category $\mathcal{C}$ is equivalent to $\mathcal{C}$.*

Proof. Let $I : \mathcal{C}_0 \to \mathcal{C}$ be the inclusion functor. We define a functor $Q : \mathcal{C} \to \mathcal{C}_0$ as follows. For each $\mathfrak{a} \in \mathcal{C}^{\mathrm{obj}}$, let $\mathfrak{a}^!$ be the unique element of $\mathcal{C}_0$ isomorphic to $\mathfrak{a}$ and let $\pi_\mathfrak{a} : \mathfrak{a} \to \mathfrak{a}^!$ be an isomorphism. We set

$$\begin{aligned} Q(\mathfrak{a}) &:= \mathfrak{a}^!, && \text{for } \mathfrak{a} \in \mathcal{C}^{\mathrm{obj}}, \\ Q(f) &:= \pi_\mathfrak{b} \circ f \circ \pi_\mathfrak{a}^{-1}, && \text{for } f : \mathfrak{a} \to \mathfrak{b} \text{ in } \mathcal{C}^{\mathrm{mor}}. \end{aligned}$$

We claim that the families $\eta := (\pi_\mathfrak{a})_{\mathfrak{a} \in \mathcal{C}_0}$ and $\rho := (\pi_\mathfrak{a})_{\mathfrak{a} \in \mathcal{C}}$ are natural isomorphisms $\eta : Q \circ I \cong \mathrm{id}$ and $\rho : I \circ Q \cong \mathrm{id}$. Since each component of η and ρ is an isomorphism, it is sufficient to prove that η and ρ are natural transformations. For η, let $f : \mathfrak{a} \to \mathfrak{b}$ be a morphism of $\mathcal{C}_o$. Then

$$Q(I(f)) \circ \eta_{\mathfrak{a}} = \pi_{\mathfrak{b}} \circ f \circ \pi_{\mathfrak{a}}^{-1} \circ \pi_{\mathfrak{a}} = \eta_{\mathfrak{a}} \circ f .$$

For ρ, let $f : \mathfrak{a} \to \mathfrak{b}$ be a morphism of $\mathcal{C}$. Then

$$I(Q(f)) \circ \rho_{\mathfrak{a}} = \pi_{\mathfrak{b}} \circ f \circ \pi_{\mathfrak{a}}^{-1} \circ \pi_{\mathfrak{a}} = \rho_{\mathfrak{a}} \circ f . \qquad \square$$

By Corollary 4.10, we obtain the following result.

Corollary 4.14. *Let $\mathcal{I}_o \subseteq \mathcal{I}$ be a skeleton of $\mathcal{I}$ and $F : \mathcal{I}_o \to \mathcal{I}$ the inclusion functor. Then*

$$\varinjlim D = \varinjlim (D \circ F) , \quad \textit{for every diagram } D : \mathcal{I} \to \mathcal{C} .$$

Chains

As a second application we show how to reduce directed diagrams to diagrams where the index category is a linear order.

Definition 4.15. A diagram $D : \mathcal{I} \to \mathcal{C}$ is a *chain* if $\mathcal{I}$ is a linear order.

Proposition 4.16. *Let $\mathcal{C}$ be a category with directed colimits, $D : \mathfrak{I} \to \mathcal{C}$ a directed diagram, and set $\kappa := |I|$. There exists a chain $C : \kappa \to \mathcal{C}$ such that*

$$\varinjlim C = \varinjlim D$$

and, for every $\alpha < \kappa$,

$$C(\alpha) = \varinjlim (D \restriction H_\alpha) , \quad \textit{for some directed subset } H_\alpha \subseteq I \textit{ of size } |H_\alpha| < |I| .$$

Proof. By Proposition 3.6, there exists a chain $(H_\alpha)_{\alpha<\kappa}$ of directed subsets $H_\alpha \subseteq I$ of size $|H_\alpha| < \kappa$ such that $I = \bigcup_{\alpha<\kappa} H_\alpha$. For $\alpha < \beta < \kappa$, let λ^α be a limiting cocone of $D \restriction H_\alpha$ and let

$$\pi_\alpha : \mathrm{Cone}(D, -) \to \mathrm{Cone}(D \restriction H_\alpha, -) ,$$
$$\pi_{\alpha,\beta} : \mathrm{Cone}(D \restriction H_\beta, -) \to \mathrm{Cone}(D \restriction H_\alpha, -) ,$$

be the projections along the inclusion functors $H_\alpha \to I$ and $H_\alpha \to H_\beta$, respectively. We define C^{obj} by

$$C(\alpha) := \varinjlim (D \restriction H_\alpha), \quad \text{for } \alpha < \kappa .$$

To define C^{mor}, let $\alpha < \beta$. Since λ^α is limiting and $\pi_{\alpha,\beta}(\lambda^\beta)$ is a cocone of $D \restriction H_\alpha$, there exists a unique morphism

$$C(\alpha,\beta) : \varinjlim (D \restriction H_\alpha) \to \varinjlim (D \restriction H_\beta),$$

such that

$$\pi_{\alpha,\beta}(\lambda^\beta) = C(\alpha,\beta) * \lambda^\alpha .$$

To prove that C is the desired chain, it is sufficient, by Lemma 4.3, to find a natural isomorphism

$$\eta : \mathrm{Cone}(D,-) \cong \mathrm{Cone}(C,-) .$$

By Lemma 4.2, there are natural isomorphisms

$$\tau_\alpha : \mathrm{Cone}(D \restriction H_\alpha, -) \cong \mathcal{C}(C(\alpha),-), \quad \text{for } \alpha < \kappa ,$$

such that

$$\begin{aligned} \mu &= \tau_\alpha(\mu) * \lambda^\alpha , && \text{for cocones } \mu \text{ of } D \restriction H_\alpha , \\ f &= \tau_\alpha(f * \lambda^\alpha), && \text{for all } f : C(\alpha) \to \mathfrak{a} . \end{aligned}$$

For a cocone μ of D, we set

$$\eta(\mu) := (\tau_\alpha(\pi_\alpha(\mu)))_{\alpha<\kappa} .$$

First, let us show that $\eta(\mu)$ is indeed a cocone of C. For indices $\alpha < \beta$, Lemma 4.5 (c) implies that

$$\begin{aligned} \tau_\alpha(\pi_\alpha(\mu)) &= \tau_\alpha\big(\pi_{\alpha,\beta}(\pi_\beta(\mu))\big) \\ &= \tau_\alpha\big(\pi_{\alpha,\beta}\big(\tau_\beta(\pi_\beta(\mu)) * \lambda^\beta\big)\big) \\ &= \tau_\alpha\big(\tau_\beta(\pi_\beta(\mu)) * \pi_{\alpha,\beta}(\lambda^\beta)\big) \\ &= \tau_\alpha\big((\tau_\beta(\pi_\beta(\mu)) \circ C(\alpha,\beta)) * \lambda^\alpha\big) \\ &= \tau_\beta(\pi_\beta(\mu)) \circ C(\alpha,\beta) . \end{aligned}$$

Hence, $(\tau_\alpha(\pi_\alpha(\mu)))_{\alpha<\kappa}$ is a cocone from C to $\mathfrak{a}$.

To see that η is a natural transformation, let $\mu \in \mathrm{Cone}(D, \mathfrak{a})$ and $f : \mathfrak{a} \to \mathfrak{b}$. Then

$$\begin{aligned}
\eta_{\mathfrak{b}}(\mathrm{Cone}(D,f)(\mu)) &= \big(\tau_\alpha(\pi_\alpha(f * \mu))\big)_{\alpha<\kappa} \\
&= \big(\tau_\alpha(f * \pi_\alpha(\mu))\big)_{\alpha<\kappa} \\
&= \big(\mathcal{C}(C(\alpha), f)(\tau_\alpha(\pi_\alpha(\mu)))\big)_{\alpha<\kappa} \\
&= f * (\tau_\alpha(\pi_\alpha(\mu)))_{\alpha<\kappa} \\
&= \mathrm{Cone}(C,f)(\eta_{\mathfrak{a}}(\mu))\,.
\end{aligned}$$

It remains to show that η is a natural isomorphism. We define an inverse ζ of η as follows. Given $\mu \in \mathrm{Cone}(D, \mathfrak{a})$ and $i \in I$, we set

$$(\zeta(\mu))_i := \mu_\alpha \circ \lambda^\alpha_i\,, \quad \text{for some } \alpha < \kappa \text{ such that } i \in H_\alpha\,.$$

First, we have to show that the value of $\zeta(\mu)$ does not depend on the choice of the ordinals α. For $i \in H_\alpha$ and $\alpha < \beta$,

$$\pi_{\alpha,\beta}(\lambda^\beta) = C(\alpha,\beta) * \lambda^\alpha$$

implies that

$$\mu_\alpha \circ \lambda^\alpha_i = \mu_\beta \circ C(\alpha,\beta) \circ \lambda^\alpha_i = \mu_\beta \circ \lambda^\beta_i\,.$$

To show that ζ is an inverse of η, we fix, for every $i \in I$, some ordinal $\alpha_i < \kappa$ with $i \in H_{\alpha_i}$. For $\mu \in \mathrm{Cone}(D, \mathfrak{a})$, it follows that

$$\begin{aligned}
\zeta(\eta(\mu)) &= \zeta\big((\tau_\alpha(\pi_\alpha(\mu)))_{\alpha<\kappa}\big) \\
&= \big(\tau_{\alpha_i}(\pi_{\alpha_i}(\mu)) \circ \lambda^{\alpha_i}_i\big)_{i\in I} \\
&= \big((\tau_{\alpha_i}(\pi_{\alpha_i}(\mu)) * \lambda^{\alpha_i})_i\big)_{i\in I} \\
&= (\pi_{\alpha_i}(\mu)_i)_{i\in I} \\
&= (\mu_i)_{i\in I}\,.
\end{aligned}$$

Conversely, for $\mu \in \mathrm{Cone}(C, \mathfrak{a})$, we have

$$\begin{aligned}
\eta(\zeta(\mu)) &= \eta\big((\mu_{\alpha_i} \circ \lambda_i^{\alpha_i})_{i \in I}\big) \\
&= \big(\tau_\beta\big(\pi_\beta\big((\mu_{\alpha_i} \circ \lambda_i^{\alpha_i})_{i \in I}\big)\big)\big)_{\beta < \kappa} \\
&= \big(\tau_\beta\big((\mu_{\alpha_i} \circ \lambda_i^{\alpha_i})_{i \in H_\beta}\big)\big)_{\beta < \kappa} \\
&= \big(\tau_\beta\big((\mu_\beta \circ \lambda_i^{\beta})_{i \in H_\beta}\big)\big)_{\beta < \kappa} \\
&= \big(\tau_\beta(\mu_\beta * \lambda^\beta)\big)_{\beta < \kappa} = (\mu_\beta)_{\beta < \kappa}\,.
\end{aligned}$$
□

Proposition 4.17. *Let $\mathcal{C}$ be a category with directed colimits. A class $\mathcal{K} \subseteq \mathcal{C}$ is closed under arbitrary directed colimits if, and only if, it is closed under colimits of chains.*

Proof. $(\Rightarrow)$ is trivial since every chain is directed. For $(\Leftarrow)$, suppose that $\mathcal{K}$ is closed under colimits of chains. Let $D : \mathcal{I} \to \mathcal{C}$ be a directed diagram such that $D(i) \in \mathcal{K}$, for all i. We prove by induction on $|I|$ that $\varinjlim D \in \mathcal{K}$. If I is finite then $\varinjlim D = D(k) \in \mathcal{K}$, for some k. Hence, we may suppose that I is infinite. Let $C : \kappa \to \mathcal{C}$ be the chain from Proposition 4.16. By inductive hypothesis, it follows that $C(\alpha) \in \mathcal{K}$, for every $\alpha < \kappa$. Since C is a chain, it follows $\varinjlim D = \varinjlim C \in \mathcal{K}$. □

5. *Links and dense functors*

There is a large class of cases where the projection $\pi_{D,F}$ along a functor F is a natural isomorphism. As we have seen, this implies that $D \circ F$ has the same colimit as D.

Alternating paths

Before introducing this class of functors, we develop several technical results to compare two functors. We start with the notion of an alternating path.

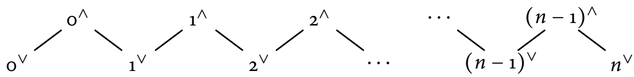
Definition 5.1. Let $\mathcal{C}$ be a category.

(a) For $n < \omega$, we denote by $\mathfrak{Z}_n = \langle Z_n, \leq \rangle$ the partial order on the elements

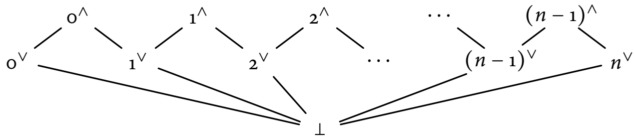
$$Z_n := \{0^\vee, \dots, n^\vee, 0^\wedge, \dots, (n-1)^\wedge\}$$

that is defined by

$$x < y \quad :\text{iff} \quad x = i^\vee \text{ and } y = k^\wedge \text{ for } k \leq i \leq k+1\,.$$



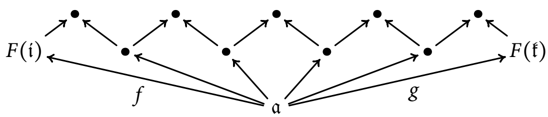
And we write $\mathfrak{Z}_n^\perp$ for the extension of $\mathfrak{Z}_n$ by a bottom element.



(b) A *alternating path* from $\mathfrak{a} \in \mathcal{C}$ to $\mathfrak{b} \in \mathcal{C}$ is a diagram $P : \mathfrak{Z}_n \to \mathcal{C}$, for some n, such that $P(0^\vee) = \mathfrak{a}$ and $P(n^\vee) = \mathfrak{b}$.

(c) We say that $\mathcal{C}$ is *connected* if, for every pair of objects $\mathfrak{a}, \mathfrak{b} \in \mathcal{C}$, there exists an alternating path from $\mathfrak{a}$ to $\mathfrak{b}$.

Remark. We will frequently be interested in alternating paths in comma categories $(\mathfrak{a} \downarrow F)$. In this case, an alternating path $P : \mathfrak{Z}_n \to (\mathfrak{a} \downarrow F)$ from $f : \mathfrak{a} \to F(\mathfrak{i})$ to $g : \mathfrak{a} \to F(\mathfrak{k})$ corresponds to a diagram $P^\perp : \mathfrak{Z}_n^\perp \to \mathcal{C}$ with $P^\perp(\perp, 0^\vee) = f$ and $P^\perp(\perp, n^\vee) = g$.



Definition 5.2. Let $F : \mathcal{I} \to \mathcal{C}$ a functor.

(a) For two morphisms $f, g \in (\mathfrak{a} \downarrow F)$, we write

$$f \curlywedge_F g \quad :\text{iff} \quad (\mathfrak{a} \downarrow F) \text{ contains an alternating path from } f \text{ to } g.$$

If $f \curlywedge_F g$, we call f and g *alternating-path equivalent,* or *a.p.-equivalent* for short. We denote the a.p.-equivalence class of f by $[f]_F^{\curlywedge}$.

(b) For families $f = (f_i)_{i\in I}$ and $g = (g_i)_{i\in I}$ of morphisms, we set

$$f \curlywedge_F g \quad :\text{iff} \quad f_i \curlywedge_F g_i \quad \text{for all } i \in I\,.$$

Again, we denote the a.p.-equivalence class of f by $[f]_F^{\curlywedge}$.

The following lemma collects the basic properties of the relation $\curlywedge_F$.

Lemma 5.3. *Let $F : \mathcal{I} \to \mathcal{C}$ be a functor and $f, g \in (\mathfrak{a} \downarrow F)$.*

(a) *$\curlywedge_F$ is an equivalence relation.*

(b) *For every morphism $h : \mathfrak{b} \to \mathfrak{a}$,*

$$f \curlywedge_F g \quad \textit{implies} \quad f \circ h \curlywedge_F g \circ h\,.$$

(c) *For all functors $D : \mathcal{C} \to \mathcal{D}$,*

$$f \curlywedge_F g \quad \textit{implies} \quad D(f) \curlywedge_{D\circ F} D(g)\,.$$

(d) *For all functors $G : \mathcal{J} \to \mathcal{I}$ and morphisms $h, h' \in \mathcal{I}^{\mathrm{mor}}$,*

$$F(h) \circ f \curlywedge_{F\circ G} F(h') \circ g \quad \textit{implies} \quad f \curlywedge_F g\,.$$

Proof. (a) $\curlywedge_F$ is reflexive since, for every morphism $f : \mathfrak{a} \to F(\mathfrak{i})$, there is an alternating path $P : \mathfrak{Z}_0 \to (\mathfrak{a} \downarrow F)$ of length 0 with $P(0^\vee) = f$. For symmetry, note that, if there is an alternating path from f to g, we can reverse it to obtain one from g to f. For transitivity, suppose that $f \curlywedge_F g$ and $g \curlywedge_F h$. Then we can find alternating paths $P : \mathfrak{Z}_m \to (\mathfrak{a} \downarrow F)$ and $Q : \mathfrak{Z}_n \to (\mathfrak{a}\downarrow F)$ from f to g and from g to h, respectively. Concatenating these paths, we obtain the desired alternating path $\mathfrak{Z}_{m+n} \to (\mathfrak{a}{\downarrow}F)$ from f to h.

(b) Let $P : \mathfrak{Z}_n \to (\mathfrak{a} \downarrow F)$ be an alternating path from f to g. We obtain an alternating path $Q : \mathfrak{Z}_n \to (\mathfrak{b} \downarrow F)$ from $f \circ h$ to $g \circ h$ by setting

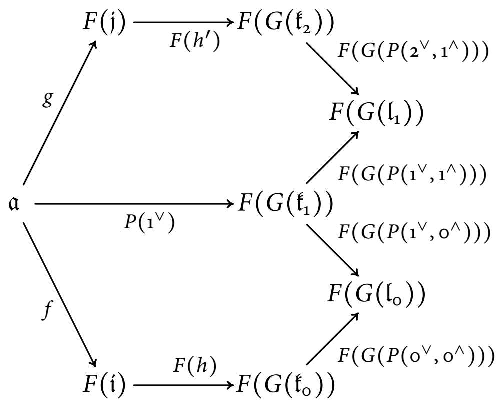
$$Q(x) := P(x) \circ h \quad \text{and} \quad Q(x,y) := P(x,y)\,, \quad \text{for } x, y \in Z_n\,.$$

(c) If $P : \mathfrak{Z}_n \to (\mathfrak{a} \downarrow F)$ is an alternating path from f to g, then $D \circ P : \mathfrak{Z}_n \to (D(\mathfrak{a}) \downarrow D \circ F)$ is an alternating path from $D(f)$ to $D(g)$.

(d) Let $P : \mathfrak{Z}_n \to (\mathfrak{a} \downarrow F \circ G)$ be an alternating path from $F(h) \circ f$ to $F(h') \circ g$. We can define an alternating path $Q : \mathfrak{Z}_n \to (\mathfrak{a} \downarrow F)$ from f to g by

$$Q(x) := \begin{cases} f & \text{if } x = 0^\vee, \\ g & \text{if } x = n^\vee, \\ P(x) & \text{otherwise}. \end{cases}$$

$$Q(i^\vee, k^\wedge) := \begin{cases} G(P(0^\vee, 0^\wedge)) \circ h & \text{if } (i,k) = (0,0)\,, \\ G(P(n^\vee, (n-1)^\wedge)) \circ h' & \text{if } (i,k) = (n, n-1)\,, \\ G(P(i^\vee, k^\wedge)) & \text{otherwise}. \end{cases}$$



□

The main reason why we are interested in alternating paths is the next lemma.

Lemma 5.4. *Let $D : \mathcal{I} \to \mathcal{C}$ be a diagram and $f : \mathfrak{a} \to D(\mathfrak{i})$, $g : \mathfrak{a} \to D(\mathfrak{j})$ morphisms. Then*

$$f \wedge\!\!\wedge_D g \quad \textit{implies} \quad \mu_{\mathfrak{i}} \circ f = \mu_{\mathfrak{j}} \circ g\,, \quad \textit{for all cocones } \mu \textit{ of } D\,.$$

Proof. Let $P : \mathfrak{Z}_n \to (\mathfrak{a} \downarrow D)$ be an alternating path from f to g. We prove the claim by induction on its length n.

For $n = 0$, we have $f = g$ and there is nothing to do. If $n > 1$, we can use the inductive hypothesis twice to obtain

$$\mu_{\mathfrak{i}} \circ f = \mu_{\mathfrak{k}} \circ P(1^{\vee}) = \mu_{\mathfrak{j}} \circ g\,,$$

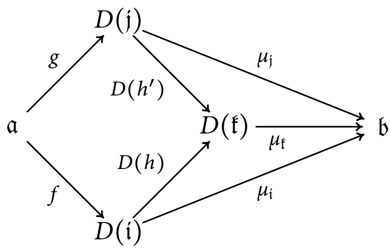
where $\mathfrak{k} \in I$ is the index such that $P(1^{\vee}) : \mathfrak{a} \to D(\mathfrak{k})$.

Hence, it remains to prove the case where $n = 1$. Let $h : \mathfrak{i} \to \mathfrak{k}$ and $h' : \mathfrak{j} \to \mathfrak{k}$ be morphisms of $\mathcal{I}$ such that

$$P(0^{\vee}, 0^{\wedge}) = D(h) \quad \text{and} \quad P(1^{\vee}, 0^{\wedge}) = D(h')\,.$$

It follows that

$$\begin{aligned}\mu_{\mathfrak{i}} \circ f = \mu_{\mathfrak{i}} \circ P(0^{\vee}) &= \mu_{\mathfrak{k}} \circ D(h) \circ P(0^{\vee}) \\ &= \mu_{\mathfrak{k}} \circ D(h') \circ P(1^{\vee}) = \mu_{\mathfrak{j}} \circ P(1^{\vee}) = \mu_{\mathfrak{j}} \circ g\,.\end{aligned}$$



□

Links

The second technical notion we introduce is that of a *link*, which generalises the notion of a natural transformation.

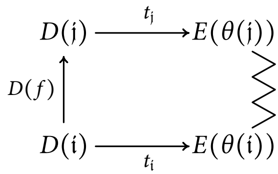
Definition 5.5. Let $D : \mathcal{I} \to \mathcal{C}$ and $E : \mathcal{J} \to \mathcal{C}$ be diagrams. A *link* from D to E is a family $t = (t_{\mathfrak{i}})_{\mathfrak{i} \in \mathcal{I}^{\mathrm{obj}}}$ of morphisms

$$t_{\mathfrak{i}} : D(\mathfrak{i}) \to E(\theta(\mathfrak{i}))\,, \quad \text{for some function } \theta : \mathcal{I}^{\mathrm{obj}} \to \mathcal{J}^{\mathrm{obj}},$$

satisfying

$$t_{\mathfrak{i}} \mathrel{⩕}_E t_{\mathfrak{j}} \circ D(f)\,,$$

for all $f : \mathfrak{i} \to \mathfrak{j}$ in $\mathcal{I}$.



We call θ the *index map* of the link.

Example. (a) Every natural transformation $\eta : D \to E$ is a link from D to E with index map $\theta(\mathfrak{i}) := \mathfrak{i}$.

(b) Every cocone $\mu \in \mathrm{Cone}(D, \mathfrak{a})$ is a link from D to the singleton functor $[1] \to \mathcal{C}$ mapping the unique object $0 \in [1]$ to $\mathfrak{a}$. The index map is $\theta(\mathfrak{i}) := 0$. Alternatively, we can regard μ as a link from D to the identity functor $\mathrm{id}_{\mathcal{C}} : \mathcal{C} \to \mathcal{C}$ with index map $\theta(\mathfrak{i}) := \mathfrak{a}$.

(c) Every morphism $f : \mathfrak{a} \to \mathfrak{b}$ can be regarded as a link from the functor $[1] \to \mathcal{C} : 0 \mapsto \mathfrak{a}$ to the functor $[1] \to \mathcal{C} : 0 \mapsto \mathfrak{b}$.

We extend the componentwise composition operation $*$ and the projection transformation from cocones to links as follows.

Definition 5.6. Let $D : \mathcal{I} \to \mathcal{C}$, $E : \mathcal{J} \to \mathcal{C}$, and $F : \mathcal{K} \to \mathcal{C}$ be diagrams, s a link from E to F, t a link from D to E.

(a) The *composition* of s and t is the family

$$s * t := (s_{\theta(\mathfrak{i})} \circ t_{\mathfrak{i}})_{\mathfrak{i} \in \mathcal{I}}\,,$$

where θ is the index map of t.

(b) The *projection* along t is the function π_t mapping a cocone μ of E to $\mu * t$.

(c) The *inclusion link* associated with D is the family

$$\mathrm{in}_D := (\mathrm{id}_{D(\mathfrak{i})})_{\mathfrak{i}\in\mathcal{I}}\,.$$

Lemma 5.7. *Let $D : \mathcal{I} \to \mathcal{C}$, $E : \mathcal{J} \to \mathcal{C}$, and $F : \mathcal{K} \to \mathcal{C}$ be diagrams, s, s' links from E to F, and t, t' links from D to E.*

(a) *$s * t$ is a link from D to F.*

(b) *If $s \mathrel{⩕}_E s'$ and $t \mathrel{⩕}_F t'$, then $s * t \mathrel{⩕}_F s' * t'$.*

(c) *For morphisms $f : \mathfrak{a} \to D(\mathfrak{i})$ and $g : \mathfrak{a} \to D(\mathfrak{j})$,*

$$f \mathrel{⩕}_D g \quad \textit{implies} \quad t_{\mathfrak{i}} \circ f \mathrel{⩕}_E t_{\mathfrak{j}} \circ g\,.$$

(d) *The inclusion link in_E associated with E is a link from E to the identity functor $\mathrm{id}_{\mathcal{C}} : \mathcal{C} \to \mathcal{C}$ such that*

$$\mathrm{in}_E * t = t \quad \textit{and} \quad s * \mathrm{in}_E = s\,.$$

Proof. We start with (c), which generalises Lemma 5.4. Choose an alternating path $P : \mathfrak{Z}_n \to (\mathfrak{a} \downarrow D)$ from f to g, and suppose that

$$P(k^\vee, k^\wedge) = h_k : \mathfrak{m}_k \to \mathfrak{n}_k$$

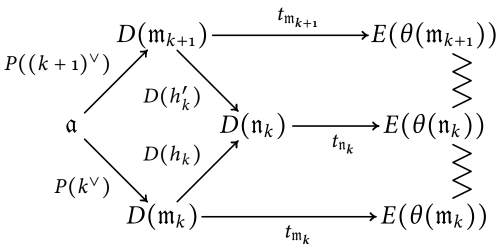
and $\quad P((k+1)^\vee, k^\wedge) = h'_k : \mathfrak{m}_{k+1} \to \mathfrak{n}_k\,.$

As t is a link, we have

$$t_{\mathfrak{m}_k} \mathrel{⩕}_E t_{\mathfrak{n}_k} \circ D(h_k) \quad \text{and} \quad t_{\mathfrak{m}_{k+1}} \mathrel{⩕}_E t_{\mathfrak{n}_k} \circ D(h'_k)\,,$$

which implies that

$$\begin{aligned} t_{\mathfrak{m}_k} \circ P(k^\vee) &\mathrel{⩕}_E t_{\mathfrak{n}_k} \circ D(h_k) \circ P(k^\vee) \\ &= t_{\mathfrak{n}_k} \circ D(h'_k) \circ P((k+1)^\vee) \mathrel{⩕}_E t_{\mathfrak{m}_{k+1}} \circ P((k+1)^\vee)\,. \end{aligned}$$



Consequently, it follows by transitivity that

$$t_{\mathfrak{i}} \circ f = t_{\mathfrak{m}_0} \circ P(0^\vee) \wedge\!\!\!\wedge_E t_{\mathfrak{m}_n} \circ P(n^\vee) = t_{\mathfrak{j}} \circ g\,.$$

(a) Let $f : \mathfrak{i} \to \mathfrak{j}$ be a morphism of $\mathcal{I}$. Since t is a link, we have

$$t_{\mathfrak{i}} \wedge\!\!\!\wedge_E t_{\mathfrak{j}} \circ D(f)\,,$$

which, by (c), implies that

$$s_{\theta(\mathfrak{i})} \circ t_{\mathfrak{i}} \wedge\!\!\!\wedge_F s_{\theta(\mathfrak{j})} \circ t_{\mathfrak{j}} \circ D(f)\,.$$

Hence, $s * t$ is a link from D to F.

(b) Let θ and θ' be the index maps of t and t', respectively. For every $\mathfrak{i} \in \mathcal{I}$, it follows by (c) that

$$t_{\mathfrak{i}} \wedge\!\!\!\wedge_E t'_{\mathfrak{i}} \quad \text{implies} \quad s_{\theta(\mathfrak{i})} \circ t_{\mathfrak{i}} \wedge\!\!\!\wedge_E s_{\theta'(\mathfrak{i})} \circ t'_{\mathfrak{i}}\,.$$

Furthermore,

$$s_{\theta'(\mathfrak{i})} \wedge\!\!\!\wedge_F s'_{\theta'(\mathfrak{i})} \quad \text{implies} \quad s_{\theta'(\mathfrak{i})} \circ t'_{\mathfrak{i}} \wedge\!\!\!\wedge_F s'_{\theta'(\mathfrak{i})} \circ t'_{\mathfrak{i}}\,.$$

By transitivity, it follows that

$$s_{\theta(\mathfrak{i})} \circ t_{\mathfrak{i}} \wedge\!\!\!\wedge_F s'_{\theta'(\mathfrak{i})} \circ t'_{\mathfrak{i}}\,.$$

(d) For every morphism $f : \mathfrak{i} \to \mathfrak{j}$ of $\mathcal{I}$, we have

$$E(f) \circ \mathrm{id}_{E(\mathfrak{i})} = E(f) = \mathrm{id}_{E(\mathfrak{j})} \circ \mathrm{id}_{E(\mathfrak{j})} \circ E(f)\,.$$

Hence, the morphisms $E(f)$ and $\mathrm{id}_{E(\mathfrak{j})}$ form an alternating path from $\mathrm{id}_{E(\mathfrak{i})}$ to $\mathrm{id}_{E(\mathfrak{j})} \circ E(f)$ in $(E(\mathfrak{i}) \downarrow \mathrm{id}_{\mathcal{C}})$. Furthermore,

$$\begin{aligned} &\mathrm{in}_E * t = (\mathrm{id}_{E(\theta(\mathfrak{i}))} \circ t_{\mathfrak{i}})_{\mathfrak{i}\in\mathcal{I}} = (t_{\mathfrak{i}})_{\mathfrak{i}\in\mathcal{I}} = t \\ \text{and}\quad &s * \mathrm{in}_E = (s_{\mathfrak{j}} \circ \mathrm{id}_{E(\mathfrak{j})})_{\mathfrak{j}\in\mathcal{J}} = (s_{\mathfrak{j}})_{\mathfrak{j}\in\mathcal{J}} = s\,. \end{aligned}$$ □

The concept of a link being quite weak, we cannot prove many statements about links in general. Their main property is the fact that they allow us to transfer cocones of E to cocones of D. In light of Lemma 5.9 below, the following lemma is a generalisation of Lemma 4.5 (a).

Lemma 5.8. *Let t be a link from $D : \mathcal{I} \to \mathcal{C}$ to $E : \mathcal{J} \to \mathcal{C}$.*

(a) *The projection π_t along t is a natural transformation*

$$\pi_t : \mathrm{Cone}(E, -) \to \mathrm{Cone}(D, -)\,.$$

(b) *$s \mathrel{\wedge\!\!\!\wedge}_E t$ implies $\pi_s = \pi_t$, for every link s from D to E.*

(c) *$\pi_{\mathrm{in}_E} = \mathrm{id}$ and $\pi_{t*s} = \pi_s \circ \pi_t$, for every link s from some diagram F to D.*

Proof. (a) We start by showing that π_t maps cocones of E to cocones of D. Let θ be the index map of t, $\mu \in \mathrm{Cone}(E, \mathfrak{a})$, and let $g : \mathfrak{i} \to \mathfrak{j}$ be a morphism of $\mathcal{I}$. As t is a link, we have

$$t_{\mathfrak{i}} \mathrel{\wedge\!\!\!\wedge}_E t_{\mathfrak{j}} \circ D(g)\,,$$

which, by Lemma 5.4, implies that

$$\mu_{\theta(\mathfrak{i})} \circ t_{\mathfrak{i}} = \mu_{\theta(\mathfrak{j})} \circ t_{\mathfrak{j}} \circ D(g)\,.$$

Hence, $\pi_t(\mu) = \mu * t$ is a cocone of D.

To show that π_t is a natural transformation, let $\mu \in \mathrm{Cone}(E, \mathfrak{a})$ and $f : \mathfrak{a} \to \mathfrak{b}$. Then

$$\begin{aligned} \pi_t(\mathrm{Cone}(E, f)(\mu)) &= (f * \mu) * t \\ &= f * (\mu * t) = \mathrm{Cone}(D, f)(\pi_t(\mu))\,. \end{aligned}$$

(b) Let ρ and θ be the index maps of, respectively, s and t. Consider a cocone $\mu \in \mathrm{Cone}(E, \mathfrak{a})$ and an index $\mathfrak{i} \in \mathcal{I}$. Since $s_\mathfrak{i} \curlywedge_E t_\mathfrak{i}$, it follows by Lemma 5.4 that

$$\mu_{\rho(\mathfrak{i})} \circ s_\mathfrak{i} = \mu_{\theta(\mathfrak{i})} \circ t_\mathfrak{i} \,.$$

Hence, $\pi_s(\mu) = \mu * s = \mu * t = \pi_t(\mu)$.

(c) For every cocone μ of E,

$$\pi_{\mathrm{in}_E}(\mu) = \mu * \mathrm{in}_E = \mu \,,$$

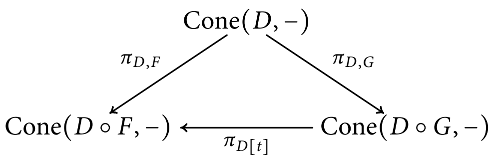
and $$\pi_{t*s}(\mu) = \mu * t * s = \pi_s(\pi_t(\mu)) \,.$$ □

Let us also make a remark about the behaviour of links when composed with a functor.

Lemma 5.9. *Let $D : \mathcal{I} \to \mathcal{C}$ be a diagram and t a link from $F : \mathcal{J} \to \mathcal{I}$ to $G : \mathcal{K} \to \mathcal{I}$.*

(a) $D[t] := (D(t_\mathfrak{j}))_{\mathfrak{j} \in \mathcal{J}}$ *is a link from $D \circ F$ to $D \circ G$.*

(b) $\pi_{D,F} = \pi_{D[t]} \circ \pi_{D,G}$.



(c) $\pi_{D,F} = \pi_{D[\mathrm{in}_F]}$.

Proof. (a) Let $g : \mathfrak{i} \to \mathfrak{j}$ be a morphism of $\mathcal{J}$. As t is a link, we have

$$t_\mathfrak{j} \circ F(g) \curlywedge_G t_\mathfrak{i} \,,$$

which, by Lemma 5.3 (c), implies that

$$D(t_\mathfrak{j}) \circ D(F(g)) \curlywedge_{D \circ G} D(t_\mathfrak{i}) \,.$$

Hence, $D[t]$ is a link from $D \circ F$ to $D \circ G$.

(b) Let $\mu \in \mathrm{Cone}(D, \mathfrak{a})$. Then

$$\begin{aligned}\pi_{D[t]}(\pi_{D,G}(\mu)) &= \pi_{D[t]}((\mu_{G(\mathfrak{k})})_{\mathfrak{k}\in\mathcal{K}}) \\ &= \big(\mu_{G(\theta(\mathfrak{j}))} \circ D(t_{\mathfrak{j}})\big)_{\mathfrak{j}\in\mathcal{J}} \\ &= (\mu_{F(\mathfrak{j})})_{\mathfrak{j}\in\mathcal{J}} = \pi_{D,F}(\mu)\,,\end{aligned}$$

where the third step follows from the fact that μ is a cocone of D.

(c) For a cocone μ of D,

$$\begin{aligned}\pi_{D[\mathrm{in}_F]}(\mu) &= \mu * D[\mathrm{in}_F] \\ &= (\mu_{F(\mathfrak{j})} \circ D(\mathrm{id}_{F(\mathfrak{j})}))_{\mathfrak{j}\in\mathcal{J}} = (\mu_{F(\mathfrak{j})})_{\mathfrak{j}\in\mathcal{J}} = \pi_{D,F}(\mu)\,. \quad \square\end{aligned}$$

We have seen in Lemma 5.7 that a.p-equivalence of links is a congruence with respect to composition. Consequently, we can define a category of a.p.-equivalence classes of links between diagrams.

Definition 5.10. Let $\mathcal{C}$ be a category and $\mathcal{P}$ a class of small categories. The *inductive $\mathcal{P}$-completion* of $\mathcal{C}$ is the category $\mathrm{Ind}_{\mathcal{P}}(\mathcal{C})$ whose objects are all diagrams $D : \mathcal{I} \to \mathcal{C}$ with $\mathcal{I} \in \mathcal{P}$. A morphism $D \to E$ between two diagrams D and E is an a.p.-equivalence class $[t]_E^{\curlywedge}$ of a link t from D to E. We write $\mathrm{Ind}_{\mathrm{all}}(\mathcal{C})$ if $\mathcal{P}$ is the class of all small categories.

Let us conclude this section with the following remarks.

Proposition 5.11. *Two diagrams $D : \mathcal{I} \to \mathcal{C}$ and $E : \mathcal{J} \to \mathcal{C}$ that are isomorphic in* $\mathrm{Ind}_{\mathrm{all}}(\mathcal{C})$ *have the same colimits.*

Proof. Let $[s]_E^{\curlywedge} : D \to E$ be an isomorphism with inverse $[t]_D^{\curlywedge} : E \to D$. By Lemma 5.8,

$$\begin{aligned} & t * s \curlywedge_D \mathrm{in}_D \quad \text{implies} \quad \pi_s \circ \pi_t = \pi_{t*s} = \pi_{\mathrm{in}_D} = \mathrm{id}\,, \\ \text{and} \quad & s * t \curlywedge_E \mathrm{in}_E \quad \text{implies} \quad \pi_t \circ \pi_s = \pi_{s*t} = \pi_{\mathrm{in}_E} = \mathrm{id}\,.\end{aligned}$$

Hence, $\pi_s : \mathrm{Cone}(E, -) \to \mathrm{Cone}(D, -)$ is a natural isomorphism and the claim follows by Lemma 4.3. $\square$

The following exercise presents an alternative, more abstract definition of the morphisms of $\mathrm{Ind}_{\mathrm{all}}(\mathcal{C})$.

Exercise 5.11. Let $D : \mathcal{I} \to \mathcal{C}$ and $E : \mathcal{J} \to \mathcal{C}$ be diagrams.

(a) Prove that, for every object $\mathfrak{a} \in \mathcal{C}$, there exists a bijection between $\varinjlim \mathcal{C}(\mathfrak{a}, E-)$ and the set

$$\{ [f]_E^{\bowtie} \mid f : \mathfrak{a} \to E(\mathfrak{j}) \text{ for some } \mathfrak{j} \in \mathcal{J} \}.$$

(b) Prove that there exists a bijection

$$\mathrm{Ind}_{\mathrm{all}}(\mathcal{C})(D, E) \to \varprojlim_D \varinjlim_E \mathcal{C}(D-, E-),$$

where $\varprojlim_D \varinjlim_E \mathcal{C}(D-, E-)$ denotes the limit of the functor

$$\mathfrak{a} \mapsto \varinjlim \mathcal{C}(D(\mathfrak{a}), E-).$$

Dense functors

After these preliminaries, we can define the class of functors preserving colimits that we mentioned above.

Definition 5.12. Let $\mathcal{C}$ be a category. A functor $F : \mathcal{I} \to \mathcal{C}$ is *dense* if, for every object $\mathfrak{a} \in \mathcal{C}$, the comma category $(\mathfrak{a} \downarrow F)$ is (D1) non-empty and (D2) connected.

Lemma 5.13. *Let $F : \mathcal{I} \to \mathcal{J}$ and $G : \mathcal{J} \to \mathcal{C}$ be dense functors. Then $G \circ F$ is also dense.*

We can characterise dense functors in terms of links.

Lemma 5.14. *Let $F : \mathcal{I} \to \mathcal{C}$ be a diagram into a small category $\mathcal{C}$ and let in_F be the inclusion link associated with F. Then F is dense if, and only if, the morphism $[\mathrm{in}_F]_{\mathrm{id}_{\mathcal{C}}}^{\bowtie} : F \to \mathrm{id}_{\mathcal{C}}$ of $\mathrm{Ind}_{\mathrm{all}}(\mathcal{C})$ has a left inverse.*

Proof. ($\Rightarrow$) Let F be dense. We use (D1) to select, for each $\mathfrak{a} \in \mathcal{C}$, a morphism $t_\mathfrak{a} : \mathfrak{a} \to F(\theta(\mathfrak{a})) \in (\mathfrak{a} \downarrow F)$. We claim that $t := (t_\mathfrak{a})_{\mathfrak{a} \in \mathcal{C}}$ is a link such that $[t]^{⩕}_F \circ [\mathrm{in}_F]^{⩕}_{\mathrm{id}_\mathcal{C}} = \mathrm{id}$.

To check that t is a link, let $f : \mathfrak{a} \to \mathfrak{b}$ be a morphism of $\mathcal{C}$. Then we can use (D2) to find the desired alternating path from $t_\mathfrak{a} \in (\mathfrak{a} \downarrow F)$ to $t_\mathfrak{b} \circ f \in (\mathfrak{a} \downarrow F)$. To show that t is a left inverse of in_F, let $\mathfrak{i} \in \mathcal{I}$. By (D2), there exists an alternating path from $t_{F(\mathfrak{i})}$ to $\mathrm{id}_{F(\mathfrak{i})}$. Hence, $t_{F(\mathfrak{i})} \circ \mathrm{id}_{F(\mathfrak{i})} ⩕_F \mathrm{id}_{F(\mathfrak{i})}$.

($\Leftarrow$) Let $[t]^{⩕}_F$ be a left inverse of $[\mathrm{in}_F]^{⩕}_{\mathrm{id}_\mathcal{C}}$. Then the morphisms $t_\mathfrak{a} \in (\mathfrak{a} \downarrow F)$ witness (D1). To check (D2), consider two morphisms $f : \mathfrak{a} \to F(\mathfrak{i})$ and $g : \mathfrak{a} \to F(\mathfrak{k})$. Since $[t]^{⩕}_F \circ [\mathrm{in}_F]^{⩕}_{\mathrm{id}_\mathcal{C}} = \mathrm{id}$, we have

$$t_{F(\mathfrak{i})} = t_{F(\mathfrak{i})} \circ \mathrm{id}_{F(\mathfrak{i})} ⩕_F \mathrm{id}_{F(\mathfrak{i})} ,$$
$$t_{F(\mathfrak{k})} = t_{F(\mathfrak{k})} \circ \mathrm{id}_{F(\mathfrak{k})} ⩕_F \mathrm{id}_{F(\mathfrak{k})} ,$$

which implies that

$$t_{F(\mathfrak{i})} \circ f ⩕_F \mathrm{id}_{F(\mathfrak{i})} \circ f = f ,$$
$$t_{F(\mathfrak{k})} \circ g ⩕_F \mathrm{id}_{F(\mathfrak{k})} \circ g = g .$$

As t is a link from $\mathrm{id}_\mathcal{C}$ to F, it follows that

$$f ⩕_F t_{F(\mathfrak{i})} \circ f ⩕_F t_\mathfrak{a} ⩕_F t_{F(\mathfrak{k})} \circ g ⩕_F g . \qquad \square$$

Let us finally prove that the projection along a dense functor preserves colimits.

Proposition 5.15. *Let $\mathcal{C}$ be a category and $D : \mathcal{I} \to \mathcal{C}$ a diagram. The projection*

$$\pi_{D,F} : \mathrm{Cone}(D, -) \to \mathrm{Cone}(D \circ F, -)$$

along a dense functor $F : \mathcal{S} \to \mathcal{I}$ is a natural isomorphism.

Proof. We have already seen in Lemma 4.5 (a) that $\pi_{D,F}$ is a natural transformation. To show that it is a natural isomorphism, we construct an inverse of $\pi_{D,F}$.

By Lemma 5.14, $[\mathrm{in}_F]^{\wedge\!\!\!\wedge}_{\mathrm{id}_{\mathcal{I}}} : F \to \mathrm{id}_{\mathcal{I}}$ has a left inverse $[t]^{\wedge\!\!\!\wedge}_F : \mathrm{id}_{\mathcal{I}} \to F$. According to Lemma 5.9, its image $D[t]$ under D is a link from D to $D \circ F$ satisfying

$$\pi_{D[t]} \circ \pi_{D,F} = \pi_{D,\mathrm{id}} = \mathrm{id}\,.$$

Hence, $\pi_{D[t]}$ is a left inverse of $\pi_{D,F}$. To show that it is also a right inverse, note that, by choice of t as left inverse to in_F, we have

$$t_{F(\mathfrak{i})} = t_{F(\mathfrak{i})} \circ \mathrm{id}_{F(\mathfrak{i})} \wedge\!\!\!\wedge_F \mathrm{id}_{F(\mathfrak{i})}\,,$$

which implies, by Lemma 5.3 (c), that

$$D(t_{F(\mathfrak{i})}) \wedge\!\!\!\wedge_{D \circ F} D(\mathrm{id}_{F(\mathfrak{i})})\,.$$

For $\mu \in \mathrm{Cone}(D \circ F, \mathfrak{a})$, it therefore follows by Lemma 5.4 that

$$\begin{aligned}
\pi_{D,F}(\pi_{D[t]}(\mu)) &= \pi_{D,F}\big((\mu_{\theta(\mathfrak{i})} \circ D(t_{\mathfrak{i}}))_{\mathfrak{i}\in\mathcal{I}}\big) \\
&= \big(\mu_{\theta(F(\mathfrak{i}))} \circ D(t_{F(\mathfrak{i})})\big)_{\mathfrak{i}\in\mathcal{S}} \\
&= \big(\mu_{\mathfrak{i}} \circ D(\mathrm{id}_{F(\mathfrak{i})})\big)_{\mathfrak{i}\in\mathcal{S}} \\
&= \mu\,.
\end{aligned}$$

□

Corollary 5.16. *Let $D : \mathcal{I} \to \mathcal{C}$ be a diagram with a colimit. If $F : \mathcal{J} \to \mathcal{I}$ is dense, then $\varinjlim (D \circ F) = \varinjlim D$.*

B4. *Accessible categories*

1. *Filtered limits and inductive completions*

Recall that every partial order can be considered as a category where there is at most one morphism between any two objects. Using this correspondence, we can generalise the notion of being κ-directed from partial orders to arbitrary categories where there may be several morphisms between two objects.

Definition 1.1. (a) A category $\mathcal{C}$ is *κ-filtered* if

(F1) for every set $X \subseteq \mathcal{C}^{\mathrm{obj}}$ of size $|X| < \kappa$, there exist an object $\mathfrak{b} \in \mathcal{C}$ and morphisms $\mathfrak{a} \to \mathfrak{b}$, for each $\mathfrak{a} \in X$;

(F2) for every pair of objects $\mathfrak{a}, \mathfrak{b} \in \mathcal{C}$ and every set $X \subseteq \mathcal{C}(\mathfrak{a}, \mathfrak{b})$ of size $|X| < \kappa$, there exist an object $\mathfrak{c} \in \mathcal{C}$ and a morphism $g : \mathfrak{b} \to \mathfrak{c}$ such that

$$g \circ f = g \circ f' , \quad \text{for all } f, f' \in X .$$

For $\kappa = \aleph_0$, we call $\mathcal{C}$ simply *filtered.*

(b) A *κ-filtered diagram* is a diagram $D : \mathcal{I} \to \mathcal{C}$ where the index category $\mathcal{I}$ is κ-filtered. The colimit of such a diagram is called a *κ-filtered colimit.*

Conditions (F1) and (F2) state that certain diagrams have a cocone. It turns out that both conditions together imply that every sufficiently small diagram has a cocone.

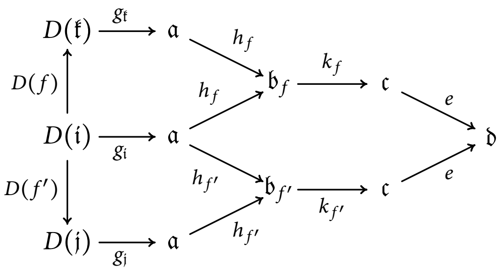
Lemma 1.2. *A category $\mathcal{C}$ is κ-filtered if, and only if, there is a cocone for every diagram $D : \mathcal{I} \to \mathcal{C}$ of size less than κ.*

Proof. ($\Leftarrow$) is obvious. For ($\Rightarrow$), let $D : \mathcal{I} \to \mathcal{C}$ be a diagram of size less than κ. By (F1), there exist an object $\mathfrak{a}$ and morphisms $g_\mathfrak{i} : D(\mathfrak{i}) \to \mathfrak{a}$, for $\mathfrak{i} \in \mathcal{I}$. By (F2), we can find, for every morphism $f : \mathfrak{i} \to \mathfrak{k}$ of $\mathcal{I}$, an object $\mathfrak{b}_f \in \mathcal{C}$ and a morphism $h_f : \mathfrak{a} \to \mathfrak{b}_f$ such that

$$h_f \circ g_\mathfrak{i} = h_f \circ g_\mathfrak{k} \circ D(f)\,.$$

By (F1), there exist an object $\mathfrak{c} \in \mathcal{C}$ and morphisms $k_f : \mathfrak{b}_f \to \mathfrak{c}$, for $f \in \mathcal{I}^{\mathrm{mor}}$. By (F2), we can find an object $\mathfrak{d} \in \mathcal{C}$ and a morphism $e : \mathfrak{c} \to \mathfrak{d}$ such that

$$e \circ k_f \circ h_f = e \circ k_{f'} \circ h_{f'}\,, \quad \text{for all } f, f' \in \mathcal{I}^{\mathrm{mor}}.$$



Set $\varphi := e \circ k_f \circ h_f$, for an arbitrary $f \in \mathcal{I}^{\mathrm{mor}}$. Then $\varphi * g$ is the desired cocone since, for every $f : \mathfrak{i} \to \mathfrak{k}$ in $\mathcal{I}^{\mathrm{mor}}$,

$$\begin{aligned} \varphi \circ g_\mathfrak{k} \circ D(f) &= e \circ k_f \circ h_f \circ g_\mathfrak{k} \circ D(f) \\ &= e \circ k_f \circ h_f \circ g_\mathfrak{i} \\ &= \varphi \circ g_\mathfrak{i}\,. \end{aligned}$$ □

It follows that a.p.-equivalence is especially simple for filtered diagrams.

Corollary 1.3. *Let $D : \mathcal{I} \to \mathcal{C}$ be a filtered diagram and $f : \mathfrak{a} \to D(\mathfrak{i})$ and $g : \mathfrak{a} \to D(\mathfrak{j})$ morphisms. Then*

$$f \mathrel{\approx}_D g \quad \text{iff} \quad \textit{there are } h : \mathfrak{i} \to \mathfrak{k} \textit{ and } h' : \mathfrak{j} \to \mathfrak{k} \textit{ in } \mathcal{I} \textit{ such that } D(h) \circ f = D(h') \circ g\,.$$

Proof. (⇐) If $D(h) \circ f = D(h') \circ g$ then h and h' form an alternating path $P : \mathfrak{Z}_1 \to (\mathfrak{a} \downarrow D)$ of length 1 from f to g.

(⇒) Fix an alternating path $P : \mathfrak{Z}_n \to (\mathfrak{a} \downarrow D)$ from f to g and let $Q : (\mathfrak{a} \downarrow D) \to \mathcal{I}$ be the projection defined by

$$Q(g) := \mathfrak{k}\,, \quad \text{for objects } g : \mathfrak{a} \to D(\mathfrak{k})\,,$$
$$Q(h) := h\,, \quad \text{for morphisms } h : g \to g'\,.$$

Then $Q \circ P : \mathfrak{Z}_n \to \mathcal{I}$ is an alternating path in $\mathcal{I}$ and Lemma 1.2 provides a cocone μ from $Q \circ P$ to some object $\mathfrak{m} \in \mathcal{I}$. By Lemma B3.4.5 (b), it follows that $D[\mu]$ is a cocone from $D \circ Q \circ P$ to $D(\mathfrak{m})$. Since all morphisms of P are in the range of $D \circ Q \circ P$, it follows that P factorises as $P = I \circ P_0$, where $P_0 : \mathfrak{Z}_n \to (\mathfrak{a} \downarrow D \circ Q \circ P)$ is an alternating path from f to g and $I : (\mathfrak{a} \downarrow D \circ Q \circ P) \to (\mathfrak{a} \downarrow D)$ is the inclusion functor. Hence, $f \mathrel{\wedge\!\!\!\wedge}_{D \circ Q \circ P} g$ and, applying Lemma B3.5.4 to the diagram $D \circ Q \circ P$, we obtain

$$D(\mu_0) \circ f = D(\mu_n) \circ g\,. \qquad \square$$

When considering κ-filtered categories, we will frequently restrict our attention to the case where κ is regular. This practice is justified by the following lemma.

Lemma 1.4. *Let κ be a singular cardinal. Every κ-filtered category $\mathcal{C}$ is κ^+-filtered.*

Proof. Let $\mathcal{C}$ be κ-filtered. To show that it is κ^+-filtered, we have to check two conditions.

(F1) Let $X \subseteq \mathcal{C}^{\mathrm{obj}}$ be a set of size $|X| \leq \kappa$. As κ is singular, we can write X as a union $\bigcup_{\alpha<\lambda} X_\alpha$ of $\lambda < \kappa$ sets of size $|X_\alpha| < \kappa$. Since $\mathcal{C}$ is κ-filtered, it follows that, for every $\alpha < \lambda$, there exist an object $\mathfrak{a}_\alpha \in \mathcal{C}$ and morphisms $f^\alpha_{\mathfrak{b}} : \mathfrak{b} \to \mathfrak{a}_\alpha$, for $\mathfrak{b} \in X_\alpha$. Similarly, we can find an object $\mathfrak{c} \in \mathcal{C}$ and morphisms $g_\alpha : \mathfrak{a}_\alpha \to \mathfrak{c}$, for $\alpha < \lambda$. For each $\mathfrak{b} \in X$, fix an ordinal $\alpha(\mathfrak{b})$ such that $\mathfrak{b} \in X_{\alpha(\mathfrak{b})}$. It follows that the family

$$g_{\alpha(\mathfrak{b})} \circ f^{\alpha(\mathfrak{b})}_{\mathfrak{b}} : \mathfrak{b} \to \mathfrak{c}\,, \quad \text{for } \mathfrak{b} \in X\,,$$

witnesses (F1).

(F2) Let $X \subseteq \mathcal{C}(\mathfrak{a}, \mathfrak{b})$ be a set of size $|X| \leq \kappa$. We write X as the union $\bigcup_{\alpha<\lambda} X_\alpha$ of an increasing sequence $(X_\alpha)_{\alpha<\lambda}$ of $\lambda < \kappa$ sets of size $|X_\alpha| < \kappa$. Since $\mathcal{C}$ is κ-filtered, it follows that, for every $\alpha < \lambda$, there exist an object $\mathfrak{c}_\alpha \in \mathcal{C}$ and a morphism $g_\alpha : \mathfrak{b} \to \mathfrak{c}_\alpha$ such that

$$g_\alpha \circ f = g_\alpha \circ f' , \quad \text{for all } f, f' \in X_\alpha .$$

By Lemma 1.2, we can find an object $\mathfrak{d}$ and morphisms $h_\alpha : \mathfrak{c}_\alpha \to \mathfrak{d}$ and $h' : \mathfrak{b} \to \mathfrak{d}$ such that

$$h_\alpha \circ g_\alpha = h' , \quad \text{for all } \alpha < \lambda .$$

We claim that h' is the desired morphism. Let $f, f' \in X$. Then $f \in X_\alpha$ and $f' \in X_\beta$, for some $\alpha, \beta < \lambda$. Setting $\gamma := \max\{\alpha, \beta\}$, it follows that $f, f' \in X_\gamma$ and

$$h' \circ f = h_\gamma \circ g_\gamma \circ f = h_\gamma \circ g_\gamma \circ f' = h' \circ f' . \qquad \square$$

Reducing filtered to directed colimits

We will show below that every κ-filtered colimit can also be obtained as colimit of a κ-directed diagram. Hence, in terms of colimits this generalisation does not provide more expressive power. We start with some technical lemmas.

Lemma 1.5. *Let $\mathcal{I}$ and $\mathcal{J}$ be κ-filtered categories.*

(a) *$\mathcal{I} \times \mathcal{J}$ is κ-filtered.*

(b) *The projection functor $P : \mathcal{I} \times \mathcal{J} \to \mathcal{I}$ is dense.*

Proof. (a) (F1) Let $\langle \mathfrak{a}_i, \mathfrak{b}_i \rangle_{i<\gamma}$ be a family of objects of size $\gamma < \kappa$. Since $\mathcal{I}$ and $\mathcal{J}$ are κ-filtered, we can find objects $\mathfrak{c} \in \mathcal{I}$ and $\mathfrak{d} \in \mathcal{J}$ and morphisms $f_i : \mathfrak{a}_i \to \mathfrak{c}$ and $g_i : \mathfrak{b}_i \to \mathfrak{d}$, for $i < \gamma$. Consequently, we obtain morphisms $\langle f_i, g_i \rangle : \langle \mathfrak{a}_i, \mathfrak{b}_i \rangle \to \langle \mathfrak{c}, \mathfrak{d} \rangle$, for $i < \gamma$.

(F2) Consider a family of morphisms

$$\langle f_i, g_i \rangle : \langle \mathfrak{a}, \mathfrak{b} \rangle \to \langle \mathfrak{c}, \mathfrak{d} \rangle , \quad i < \gamma ,$$

of size $\gamma < \kappa$. Since $\mathcal{I}$ and $\mathcal{J}$ are κ-filtered, we can find morphisms $h : \mathfrak{c} \to \mathfrak{e}$ in $\mathcal{I}$ and $k : \mathfrak{d} \to \mathfrak{f}$ in $\mathcal{J}$ such that

$$h \circ f_i = h \circ f_j \quad \text{and} \quad k \circ g_i = k \circ g_j , \quad \text{for all } i, j < \gamma .$$

Consequently,

$$\langle h, k \rangle \circ \langle f_i, g_i \rangle = \langle h, k \rangle \circ \langle f_j, g_j \rangle , \quad \text{for all } i, j < \gamma .$$

(b) (D1) We can use (F1) with $X = \varnothing$ to find some object $\mathfrak{b} \in \mathcal{J}$. It follows that, for every $\mathfrak{a} \in \mathcal{I}$, we have a morphism $\mathrm{id}_{\mathfrak{a}} : \mathfrak{a} \to P(\langle \mathfrak{a}, \mathfrak{b} \rangle)$.

(D2) Let $f : \mathfrak{a} \to P(\langle \mathfrak{b}, \mathfrak{c} \rangle)$ and $f' : \mathfrak{a} \to P(\langle \mathfrak{b}', \mathfrak{c}' \rangle)$ be morphisms of $\mathcal{I}$. By Lemma 1.2, there exist morphisms $g : \mathfrak{b} \to \mathfrak{d}$, $g' : \mathfrak{b}' \to \mathfrak{d}$, and $g'' : \mathfrak{a} \to \mathfrak{d}$ such that $g \circ f = g'' = g' \circ f'$. As $\mathcal{J}$ is κ-filtered, there exist an object $\mathfrak{e} \in \mathcal{J}$ and morphisms $h : \mathfrak{c} \to \mathfrak{e}$ and $h' : \mathfrak{c}' \to \mathfrak{e}$. Consequently, we obtain morphisms $\langle g, h \rangle : \langle \mathfrak{b}, \mathfrak{c} \rangle \to \langle \mathfrak{d}, \mathfrak{e} \rangle$ and $\langle g', h' \rangle : \langle \mathfrak{b}', \mathfrak{c}' \rangle \to \langle \mathfrak{d}, \mathfrak{e} \rangle$ such that

$$P(\langle g, h \rangle) \circ f = P(\langle g', h' \rangle) \circ f' .$$

These two morphisms form an alternating path from f to f'. ☐

Lemma 1.6. *Let $\mathcal{I}$ be a κ-filtered category and $\mathfrak{K}$ a κ-directed partial order without maximal elements. Every subcategory $\mathcal{A} \subseteq \mathcal{I} \times \mathfrak{K}$ with $|\mathcal{A}^{\mathrm{mor}}| < \kappa$ can be extended to a subcategory $\mathcal{A} \subseteq \mathcal{A}_+ \subseteq \mathcal{I} \times \mathfrak{K}$ such that $|\mathcal{A}_+^{\mathrm{mor}}| < \kappa$ and $\mathcal{A}_+$ has a unique terminal object.*

Proof. Let $\mathcal{A} \subseteq \mathcal{I} \times \mathfrak{K}$ be a subcategory with less than κ morphisms. According to Lemma 1.5, the product $\mathcal{I} \times \mathfrak{K}$ is κ-filtered. Therefore, we can use Lemma 1.2 to find a cocone μ from the inclusion functor $\mathcal{A} \to \mathcal{I} \times \mathfrak{K}$ to some object $\langle \mathfrak{b}, k \rangle \in \mathcal{I} \times \mathfrak{K}$. Since $\mathfrak{K}$ has no maximal element, there exists some $l \in K$ with $l > k$. Let $h := \langle \mathrm{id}_{\mathfrak{b}}, h' \rangle : \langle \mathfrak{b}, k \rangle \to \langle \mathfrak{b}, l \rangle$ be the morphisms whose second component is the unique morphism $h' : k \to l$ of $\mathfrak{K}$. Let $\mathcal{A}_+$ be the category obtained from $\mathcal{A}$ by adding the object $\langle \mathfrak{b}, l \rangle$, the identity morphism $\mathrm{id}_{\langle \mathfrak{b}, l \rangle}$, and the morphisms

$$h \circ \mu_{\langle \mathfrak{a}, i \rangle} : \langle \mathfrak{a}, i \rangle \to \langle \mathfrak{b}, l \rangle , \quad \text{for all } \langle \mathfrak{a}, i \rangle \in \mathcal{A} .$$

(Note that these morphisms are closed under composition since $h * \mu$ is a cocone.) Then $\langle \mathfrak{b}, l \rangle$ is the unique terminal object of $\mathcal{A}_+$. □

Theorem 1.7. *Let κ be a regular cardinal. For every small κ-filtered category $\mathcal{C}$, there exist a dense κ-directed diagram $D : \mathfrak{J} \to \mathcal{C}$.*

Proof. Set $\mathcal{J} := \mathcal{C} \times \kappa$ and let $P : \mathcal{J} \to \mathcal{C}$ be the projection functor. By Lemma 1.5, $\mathcal{J}$ is κ-filtered and P is dense. It is therefore sufficient to find a dense κ-directed diagram $D : \mathfrak{J} \to \mathcal{J}$. Then the composition $P \circ D$ is the desired dense κ-directed diagram.

As index set we use the partial order $\mathfrak{J} := \langle \mathcal{I}, \subseteq \rangle$ where $\mathcal{I}$ is the set of all subcategories $\mathcal{A} \subseteq \mathcal{J}$ with $|\mathcal{A}^{\mathrm{mor}}| < \kappa$ such that $\mathcal{A}$ has a unique terminal object. To show that $\mathfrak{J}$ is κ-directed, consider a set $X \subseteq \mathcal{I}$ of size $|X| < \kappa$. Let $\mathcal{A}$ be the subcategory of $\mathcal{J}$ generated by the morphisms in

$$\bigcup_{\mathcal{B} \in X} \mathcal{B}^{\mathrm{mor}} .$$

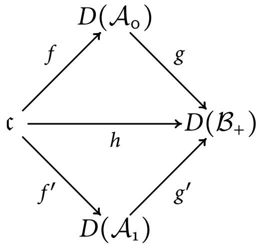
Since κ is regular, $\mathcal{A}$ still has less than κ morphisms. By Lemma 1.6, there exists a subcategory $\mathcal{A} \subseteq \mathcal{A}_+ \subseteq \mathcal{J}$ with a unique terminal object. Hence, $\mathcal{A}_+ \in \mathcal{I}$ is an upper bound of X.

Let $D : \mathfrak{J} \to \mathcal{J}$ be the functor mapping a subcategory $\mathcal{A} \in \mathfrak{J}$ to its terminal object and mapping a pair $\mathcal{A} \subseteq \mathcal{B}$ of subcategories to the unique morphism from the terminal object of $\mathcal{A}$ to the terminal object of $\mathcal{B}$. We claim that D is dense in $\mathcal{J}$.

For (D1), let $\mathfrak{c} \in \mathcal{J}$. The subcategory $\mathcal{A}$ of $\mathcal{J}$ consisting just of the object $\mathfrak{c}$ and its identity morphism has a unique terminal object. Hence, $\mathcal{A} \in \mathfrak{J}$ and $D(\mathcal{A}) = \mathfrak{c}$. Consequently, the identity morphism $\mathrm{id}_{\mathfrak{c}} : \mathfrak{c} \to D(\mathcal{A})$ has the desired properties.

For (D2), let $f : \mathfrak{c} \to D(\mathcal{A}_0)$ and $f' : \mathfrak{c} \to D(\mathcal{A}_1)$ be morphisms of $\mathcal{J}$. Let $\mathcal{B}$ be a subcategory of $\mathcal{J}$ of size $|\mathcal{B}^{\mathrm{mor}}| < \kappa$ containing f, f' and every morphism of $\mathcal{A}_0^{\mathrm{mor}} \cup \mathcal{A}_1^{\mathrm{mor}}$. By Lemma 1.6, there exists a subcategory $\mathcal{B}_+ \in \mathcal{I}$ containing $\mathcal{B}$. Since $D(\mathcal{B}_+)$ is a terminal object, $\mathcal{B}_+$ contains unique morphisms

$$h : \mathfrak{c} \to D(\mathcal{B}_+)\,,$$
$$g : D(\mathcal{A}_0) \to D(\mathcal{B}_+)\,,$$
$$g' : D(\mathcal{A}_1) \to D(\mathcal{B}_+)\,.$$



By uniqueness, it follows that $g \circ f = h = g' \circ f'$. Hence, g and g' from an alternating path from f to f' □

Corollary 1.8. *Let κ be a regular cardinal. For every κ-filtered diagram $D : \mathcal{I} \to \mathcal{C}$ with a colimit, there exists a κ-directed diagram $F : \mathfrak{K} \to \mathcal{I}$ such that $\varinjlim (D \circ F) = \varinjlim D$.*

Corollary 1.9. *Let κ be a regular cardinal. A functor $F : \mathcal{C} \to \mathcal{D}$ preserves κ-filtered colimits if, and only if, it preserves κ-directed ones.*

Inductive completions

There is a general way to construct the closure of a category under κ-filtered colimits.

Definition 1.10. Let $\mathcal{C}$ be a category, κ an infinite cardinal, and λ either an infinite cardinal or $\lambda = \infty$.

(a) The *inductive (κ, λ)-completion* of $\mathcal{C}$ is the category

$$\mathrm{Ind}_\kappa^\lambda(\mathcal{C}) := \mathrm{Ind}_{\mathcal{P}_\kappa^\lambda}(\mathcal{C})\,,$$

where $\mathcal{P}_\kappa^\lambda$ is the class of all small κ-filtered categories of size less than λ. For $\kappa = \aleph_0$ and $\lambda = \infty$, we drop the indices and simply write $\mathrm{Ind}(\mathcal{C})$.

(b) Let $\mathcal{P}$ be a class of small categories containing the singleton category $[1]$. The *inclusion functor* $I : \mathcal{C} \to \mathrm{Ind}_{\mathcal{P}}(\mathcal{C})$ sends an object $\mathfrak{a} \in \mathcal{C}$ to the singleton diagram $C_\mathfrak{a} : [1] \to \mathcal{C} : 0 \mapsto \mathfrak{a}$ and a morphism $f : \mathfrak{a} \to \mathfrak{b}$ to the link $t = (t_i)_{i \in [1]}$ from $C_\mathfrak{a}$ to $C_\mathfrak{b}$ that consists of the morphism $t_0 := f$.

We will show below that $\mathrm{Ind}_\kappa^\lambda(\mathcal{C})$ is the closure of $\mathcal{C}$ under κ-filtered colimits of size less than λ. We start by determining the colimit of a κ-filtered diagram $D : \mathcal{I} \to \mathrm{Ind}_\kappa^\lambda(\mathcal{C})$. This colimit consists of a large diagram U that is built up from the diagrams $D(\mathfrak{i})$, for $\mathfrak{i} \in \mathcal{I}$.

Definition 1.11. Let $D : \mathcal{I} \to \mathrm{Ind}_\kappa^\lambda(\mathcal{C})$ be a diagram and, for $\mathfrak{i} \in \mathcal{I}$, let $\mathcal{K}(\mathfrak{i})$ be the index category of the diagram $D(\mathfrak{i}) : \mathcal{K}(\mathfrak{i}) \to \mathcal{C}$.

(a) A *union* of D is a diagram $U : \mathcal{J} \to \mathcal{C}$ of the following form. For each morphism $f : \mathfrak{i} \to \mathfrak{j}$ of $\mathcal{I}$, fix a link $t(f)$ from $D(\mathfrak{i})$ to $D(\mathfrak{j})$ such that $D(f) = [t(f)]^{\approx}_{D(\mathfrak{j})}$. Let $\mathcal{S}$ be the subcategory of $\mathcal{C}$ generated by all morphisms in

$$\bigcup_{\mathfrak{i} \in \mathcal{I}^{\mathrm{obj}}} \operatorname{rng} D(\mathfrak{i})^{\mathrm{mor}} \cup \bigcup_{f \in \mathcal{I}^{\mathrm{mor}}} t(f)\,.$$

The index category $\mathcal{J}$ has the objects

$$\mathcal{J}^{\mathrm{obj}} := \dot{\bigcup_{\mathfrak{i} \in \mathcal{I}^{\mathrm{obj}}}} \mathcal{K}(\mathfrak{i})^{\mathrm{obj}} = \{\, \langle \mathfrak{i}, \mathfrak{k} \rangle \mid \mathfrak{i} \in \mathcal{I},\ \mathfrak{k} \in \mathcal{K}(\mathfrak{i}) \,\}\,,$$

and the morphisms

$$\mathcal{J}\big(\langle \mathfrak{i}, \mathfrak{k} \rangle, \langle \mathfrak{j}, \mathfrak{l} \rangle\big) := \mathcal{S}\big(D(\mathfrak{i})(\mathfrak{k}), D(\mathfrak{j})(\mathfrak{l})\big)\,.$$

The functor $U : \mathcal{J} \to \mathcal{C}$ is defined by

$$\begin{aligned} U(\langle \mathfrak{i}, \mathfrak{k} \rangle) &:= D(\mathfrak{i})(\mathfrak{k})\,, && \text{for } \langle \mathfrak{i}, \mathfrak{k} \rangle \in \mathcal{J}^{\mathrm{obj}}, \\ U(f) &:= f\,, && \text{for } f \in \mathcal{J}^{\mathrm{mor}}. \end{aligned}$$

(b) Let μ be a cocone from D to some object $E \in \mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$ and, for $\mathfrak{i} \in \mathcal{I}$, let $t^{\mathfrak{i}} = (t^{\mathfrak{i}}_{\mathfrak{f}})_{\mathfrak{f} \in \mathcal{K}(\mathfrak{i})}$ be a link such that $\mu_{\mathfrak{i}} = [t^{\mathfrak{i}}]^{\curlywedge}_{E}$. The *union* of μ is the a.p.-equivalence class $[t]^{\curlywedge}_{E}$ of the family

$$t := (t^{\mathfrak{i}}_{\mathfrak{f}})_{\langle \mathfrak{i}, \mathfrak{f} \rangle \in \mathcal{J}} .$$

Remark. Note that, due to the choice of the links $t(f)$, a diagram D might have several unions. It will follow from Proposition 1.13 below that they are all isomorphic.

To prove that the union of a diagram is its colimit, we start with a lemma collecting several technical properties of the union operation.

Lemma 1.12. *Let $U : \mathcal{J} \to \mathcal{C}$ be a union of the diagram $D : \mathcal{I} \to \mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$, and let $E \in \mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$.*

(a) *Every cocone $\mu \in \mathrm{Cone}(D, E)$ has a unique union.*

(b) *The union $[u]^{\curlywedge}_{E}$ of $\mu \in \mathrm{Cone}(D, E)$ is a morphism $[u]^{\curlywedge}_{E} : U \to E$ of $\mathrm{Ind}_{\mathrm{all}}(\mathcal{C})$.*

(c) *The function $\eta_E : \mathrm{Cone}(D, E) \to \mathrm{Ind}_{\mathrm{all}}(\mathcal{C})(U, E)$ that maps a cocone to its union is bijective.*

(d) *For $\mathfrak{i} \in \mathcal{I}$, the inclusion link $\mathrm{in}_{D(\mathfrak{i})}$ is a link from $D(\mathfrak{i})$ to U.*

Proof. Let $\mathcal{K}(\mathfrak{i})$ be the index category of $D(\mathfrak{i})$ and, for $f \in \mathcal{I}^{\mathrm{mor}}$, let $t(f)$ be the representative of $D(f)$ used to construct the union U.

(a) We have to show that the union of μ is independent of the choice of the links. For each $\mathfrak{i} \in \mathcal{I}$, suppose that $u^{\mathfrak{i}}$ and $w^{\mathfrak{i}}$ are a.p.-equivalent links from $D(\mathfrak{i})$ to E such that

$$[u^{\mathfrak{i}}]^{\curlywedge}_{E} = \mu_{\mathfrak{i}} = [w^{\mathfrak{i}}]^{\curlywedge}_{E} .$$

Then $[u^{\mathfrak{i}}_{\mathfrak{f}}]^{\curlywedge}_{E} = [w^{\mathfrak{i}}_{\mathfrak{f}}]^{\curlywedge}_{E}$, for all $\langle \mathfrak{i}, \mathfrak{f} \rangle \in \mathcal{J}$, which implies that the corresponding links $u = (u^{\mathfrak{i}}_{\mathfrak{f}})_{\langle \mathfrak{i}, \mathfrak{f} \rangle \in \mathcal{J}}$ and $w = (w^{\mathfrak{i}}_{\mathfrak{f}})_{\langle \mathfrak{i}, \mathfrak{f} \rangle \in \mathcal{J}}$ are a.p.-equivalent and induce the same value $[u]^{\curlywedge}_{E} = [w]^{\curlywedge}_{E}$.

(b) Let $\mu \in \mathrm{Cone}(D, E)$ be a cocone where $\mu_{\mathfrak{i}} = [u^{\mathfrak{i}}]^{\curlywedge}_{E}$, and let $[u]^{\curlywedge}_{E}$ be the union of μ. We have to show that u is a link from U to E. As every morphism of $\mathcal{J}$ is a finite composition of morphisms of the form $t(f)_{\mathfrak{k}}$ and $D(\mathfrak{i})(g)$, it is sufficient to prove the equivalence

$$u^{\mathfrak{j}}_{\mathfrak{l}} \circ U(h) \approx_E u^{\mathfrak{i}}_{\mathfrak{k}}$$

for morphisms $h : \langle \mathfrak{i}, \mathfrak{k} \rangle \to \langle \mathfrak{j}, \mathfrak{l} \rangle$ of this form.

For $h = D(\mathfrak{i})(g)$ with $g : \mathfrak{k} \to \mathfrak{l}$ in $\mathcal{K}(\mathfrak{i})$, note that $u^{\mathfrak{i}}$ is a link from $D(\mathfrak{i})$ to E. Hence,

$$u^{\mathfrak{i}}_{\mathfrak{l}} \circ D(\mathfrak{i})(g) \approx_E u^{\mathfrak{i}}_{\mathfrak{k}} .$$

For $h = t(f)_{\mathfrak{k}}$ with $f : \mathfrak{i} \to \mathfrak{j}$ in $\mathcal{I}$ and $\mathfrak{k} \in \mathcal{K}(\mathfrak{i})$, the fact that μ is a cocone of D implies that $[u^{\mathfrak{j}}]^{\approx}_E \circ [t(f)]^{\approx}_{D(\mathfrak{j})} = [u^{\mathfrak{i}}]^{\approx}_E$. Hence,

$$u^{\mathfrak{j}}_{\theta(\mathfrak{k})} \circ t(f)_{\mathfrak{k}} \approx_E u^{\mathfrak{i}}_{\mathfrak{k}} ,$$

where θ is the index map of $t(f)$.

(c) We have seen in (b) that η_E maps cocones from D to E to morphisms in $\mathrm{Ind}_{\mathrm{all}}(\mathcal{C})(U, E)$. Hence, it remains to prove that η_E is bijective.

For injectivity, consider two cocones $\mu, \mu' \in \mathrm{Cone}(D, E)$ such that $\eta_E(\mu) = \eta_E(\mu')$. Fix links $u^{\mathfrak{i}}$, $w^{\mathfrak{i}}$, and $t = (t_{\mathfrak{i},\mathfrak{k}})_{\langle \mathfrak{i},\mathfrak{k} \rangle \in \mathcal{J}}$ such that

$$\mu_{\mathfrak{i}} = [u^{\mathfrak{i}}]^{\approx}_E , \quad \mu'_{\mathfrak{i}} = [w^{\mathfrak{i}}]^{\approx}_E , \quad \text{and} \quad \eta_E(\mu) = [t]^{\approx}_E .$$

Then $[u^{\mathfrak{i}}_{\mathfrak{k}}]^{\approx}_E = [t_{\langle \mathfrak{i},\mathfrak{k} \rangle}]^{\approx}_E = [w^{\mathfrak{i}}_{\mathfrak{k}}]^{\approx}_E$ for all indices $\mathfrak{i}, \mathfrak{k}$. Consequently,

$$\mu_{\mathfrak{i}} = [u^{\mathfrak{i}}]^{\approx}_E = [w^{\mathfrak{i}}]^{\approx}_E = \mu'_{\mathfrak{i}} , \quad \text{for all } \mathfrak{i} \in \mathcal{I} ,$$

which implies that $\mu = \mu'$.

For surjectivity, let $s = (s_{\mathfrak{i},\mathfrak{k}})_{\langle \mathfrak{i},\mathfrak{k} \rangle \in \mathcal{J}}$ be a link from U to E. For $\mathfrak{i} \in \mathcal{I}$, we set $s^{\mathfrak{i}} := (s_{\mathfrak{i},\mathfrak{k}})_{\mathfrak{k} \in \mathcal{K}(\mathfrak{i})}$ and $\mu := ([s^{\mathfrak{i}}]^{\approx}_E)_{\mathfrak{i} \in \mathcal{I}}$. As $\eta_E(\mu) = [s]^{\approx}_E$ it is sufficient to prove that μ is a cocone from D to E.

We start by showing that each family $s^{\mathfrak{i}}$ is a link from $D(\mathfrak{i})$ to E. Let $g : \mathfrak{k} \to \mathfrak{l}$ be a morphism of $\mathcal{K}(\mathfrak{i})$. As s is a link from U to E, we have $s_{\mathfrak{j},\mathfrak{l}} \circ D(\mathfrak{i})(g) \approx_E s_{\mathfrak{i},\mathfrak{k}}$, as desired.

It remains to show that μ is a cocone. Let $f : \mathfrak{i} \to \mathfrak{j}$ be a morphism of $\mathcal{I}$ and let θ be the index map of $t(f)$. Since s is a link from U to E,

$$s_{\mathfrak{j},\theta(\mathfrak{k})} \circ U(t(f)_{\mathfrak{k}}) \mathrel{\wedge\!\!\!\wedge}_E s_{\mathfrak{i},\mathfrak{k}}\,, \quad \text{for every } \mathfrak{k} \in \mathcal{K}(\mathfrak{i})\,.$$

Consequently,

$$\mu_{\mathfrak{j}} \circ D(f) = [s^{\mathfrak{j}}]_E^{\wedge\!\!\!\wedge} \circ [t(f)]_{D(\mathfrak{j})}^{\wedge\!\!\!\wedge} = [s^{\mathfrak{i}}]_E^{\wedge\!\!\!\wedge} = \mu_{\mathfrak{i}}\,.$$

(d) Consider a morphism $g : \mathfrak{k} \to \mathfrak{l}$ of $\mathcal{K}(\mathfrak{i})$ and set $f := D(\mathfrak{i})(g)$. Then $f : \langle \mathfrak{i}, \mathfrak{k} \rangle \to \langle \mathfrak{i}, \mathfrak{l} \rangle$ in $\mathcal{J}$ and

$$U(\mathrm{id}_{\langle \mathfrak{i},\mathfrak{l}\rangle}) \circ \mathrm{id}_{D(\mathfrak{i})(\mathfrak{l})} \circ D(\mathfrak{i})(g) = f = U(f) = U(f) \circ \mathrm{id}_{D(\mathfrak{i})(\mathfrak{k})}\,.$$

Hence, $\mathrm{id}_{\langle \mathfrak{i},\mathfrak{l}\rangle}$ and f form an alternating path from $\mathrm{id}_{D(\mathfrak{i})(\mathfrak{l})} \circ D(\mathfrak{i})(g)$ to $\mathrm{id}_{D(\mathfrak{i})(\mathfrak{k})}$ in $(D(\mathfrak{i})(\mathfrak{k}) \downarrow U)$. □

After these preparations we can prove that a union is a colimit.

Proposition 1.13. *Let $\mathcal{C}$ be a category, κ, λ regular cardinals (or $\lambda = \infty$), and let $D : \mathcal{I} \to \mathrm{Ind}_\kappa^\lambda(\mathcal{C})$ be a κ-filtered diagram of size less than λ with union U.*

(a) $U \in \mathrm{Ind}_\kappa^\lambda(\mathcal{C})$.

(b) *$U = \varinjlim D$ and a limiting cocone $\mu = (\mu_{\mathfrak{i}})_{\mathfrak{i}\in\mathcal{I}}$ from D to U is given by*

$$\mu_{\mathfrak{i}} = [\mathrm{in}_{D(\mathfrak{i})}]_U^{\wedge\!\!\!\wedge} : D(\mathfrak{i}) \to U\,.$$

Proof. Let $\mathcal{K}(\mathfrak{i})$ be the index category of $D(\mathfrak{i})$ and, for $f \in \mathcal{I}^{\mathrm{mor}}$, let $t(f)$ be the representative of $D(f)$ used to construct the union U.

(a) Since λ is regular, we have

$$|\mathcal{J}^{\mathrm{mor}}| \leq \sum_{\mathfrak{i}\in\mathcal{I}} |\mathcal{K}(\mathfrak{i})^{\mathrm{mor}}| < \lambda\,.$$

Hence, it remains to prove that U is κ-filtered.

(F1) Let $X \subseteq \mathcal{I}^{\mathrm{obj}}$ be a set of size $|X| < \kappa$. Since $\mathcal{I}$ is κ-filtered, there exist an object $\mathfrak{m} \in \mathcal{I}$ and, for every $\langle \mathfrak{i}, \mathfrak{k} \rangle \in X$, a morphism $f_{\mathfrak{i}} : \mathfrak{i} \to \mathfrak{m}$ in $\mathcal{I}$. Let $\theta^{\mathfrak{i}}$ be the index map of $t(f_{\mathfrak{i}})$. Since $\mathcal{K}(\mathfrak{m})$ is κ-filtered, it contains an object $\mathfrak{n} \in \mathcal{K}(\mathfrak{m})$ and morphisms $g_{\mathfrak{i},\mathfrak{k}} : \theta^{\mathfrak{i}}(\mathfrak{k}) \to \mathfrak{n}$, for every $\langle \mathfrak{i}, \mathfrak{k} \rangle \in X$. The desired family of morphisms of $\mathcal{J}$ is given by

$$h_{\mathfrak{i},\mathfrak{k}} := D(\mathfrak{m})(g_{\mathfrak{i},\mathfrak{k}}) \circ t(f_{\mathfrak{i}})_{\mathfrak{k}}\,, \quad \text{for } \langle \mathfrak{i}, \mathfrak{k} \rangle \in X\,.$$

(F2) Let $X \subseteq \mathcal{J}(\langle \mathfrak{i}, \mathfrak{k} \rangle, \langle \mathfrak{j}, \mathfrak{l} \rangle)$ be a set of size $|X| < \kappa$. For each morphism $f \in X$, we choose a factorisation

$$f = h^f_0 \circ \cdots \circ h^f_{n_f}\,,$$

where each factor h^f_i is of the form $D(\mathfrak{m})(g)$, for some $\mathfrak{m} \in \mathcal{I}^{\mathrm{obj}}$ and $g \in \mathcal{K}(\mathfrak{i})^{\mathrm{mor}}$, or of the form $t(f)_{\mathfrak{m}}$, for some $f \in \mathcal{I}^{\mathrm{mor}}$. Let $\mathcal{J}_0 \subseteq \mathcal{J}$ be the minimal subcategory of $\mathcal{J}$ that contains all these morphisms h^f_i, for $f \in X$ and $i \leq n_f$, and such that the restriction $U_0 := U \upharpoonright \mathcal{J}_0$ is a union of some restriction $D \upharpoonright \mathcal{I}_0$, for some $\mathcal{I}_0 \subseteq \mathcal{I}$. Let $F : \mathcal{I}_0 \to \mathcal{I}$ be the inclusion functor. Note that $|X| < \kappa$ implies

$$|\mathcal{I}^{\mathrm{mor}}_0| < \kappa \quad \text{and} \quad |\mathcal{J}^{\mathrm{mor}}_0| < \kappa\,.$$

As $\mathcal{I}$ is κ-filtered, we can use Lemma 1.2 to find a cocone μ_0 from F to some object $\mathfrak{m} \in \mathcal{I}$. Set $\mu := D[\mu_0]$ and let $[u]^{\approx}_{D(\mathfrak{m})}$ be the union of μ. By Lemma 1.12 (b), u is a link from U_0 to $D(\mathfrak{m})$. Hence,

$$u_{\langle \mathfrak{j},\mathfrak{l} \rangle} \circ f \approx_{D(\mathfrak{m})} u_{\langle \mathfrak{i},\mathfrak{k} \rangle}\,, \quad \text{for every } f \in X\,.$$

Let ρ be the index map of u. As $D(\mathfrak{m})$ is κ-filtered, we can use Corollary 1.3 to find morphisms

$$h_f : \rho(\langle \mathfrak{j}, \mathfrak{l} \rangle) \to \mathfrak{n}_f \quad \text{and} \quad h'_f : \rho(\langle \mathfrak{i}, \mathfrak{k} \rangle) \to \mathfrak{n}_f$$

such that

$$D(\mathfrak{m})(h_f) \circ u_{\langle \mathfrak{j},\mathfrak{l} \rangle} \circ f = D(\mathfrak{m})(h'_f) \circ u_{\langle \mathfrak{i},\mathfrak{k} \rangle}\,.$$

According to Lemma 1.2, we can find an object $\mathfrak{n} \in \mathcal{K}(\mathfrak{m})$ and morphisms $g_f : \mathfrak{n}_f \to \mathfrak{n}$, for $f \in X$, such that

$$g_f \circ h_f = g_{f'} \circ h_{f'} \quad \text{and} \quad g_f \circ h'_f = g_{f'} \circ h'_{f'} ,$$

for all $f, f' \in X$. Hence, $\varphi := D(\mathfrak{m})(g_f \circ h_f) \circ u_{\langle \mathfrak{j}, \mathfrak{l} \rangle}$ (which does not depend on f) is a morphism such that

$$\begin{aligned} \varphi \circ f &= D(\mathfrak{m})(g_f \circ h_f) \circ u_{\langle \mathfrak{j}, \mathfrak{l} \rangle} \circ f \\ &= D(\mathfrak{m})(g_f \circ h'_f) \circ u_{\langle \mathfrak{i}, \mathfrak{k} \rangle} \\ &= D(\mathfrak{m})(g_{f'} \circ h'_{f'}) \circ u_{\langle \mathfrak{i}, \mathfrak{k} \rangle} \\ &= D(\mathfrak{m})(g_{f'} \circ h_{f'}) \circ u_{\langle \mathfrak{j}, \mathfrak{l} \rangle} \circ f' = \varphi \circ f' , \end{aligned}$$

for all $f, f' \in X$.

(b) To see that μ is the desired limiting cocone, we have to check several properties. We have already seen in Lemma 1.12 (d) that each component $\mu_{\mathfrak{i}}$ is a morphism $D(\mathfrak{i}) \to U$.

Next, we prove that μ is a cocone of D. Let $f : \mathfrak{i} \to \mathfrak{j}$ be a morphism of $\mathcal{I}$ and let θ be the index map of $t(f)$. Then

$$U(t(f)_{\mathfrak{k}}) \circ \mathrm{id}_{D(\mathfrak{i})(\mathfrak{k})} = t(f)_{\mathfrak{k}} = U(\mathrm{id}_{\langle \mathfrak{j}, \theta(\mathfrak{k}) \rangle}) \circ \mathrm{id}_{D(\mathfrak{j})(\theta(\mathfrak{k}))} \circ t(f)_{\mathfrak{k}} .$$

Hence, $t(f)_{\mathfrak{k}}$ and $\mathrm{id}_{\langle \mathfrak{j}, \theta(\mathfrak{k}) \rangle}$ form an alternating path from $\mathrm{id}_{D(\mathfrak{i})(\mathfrak{k})}$ to $\mathrm{id}_{D(\mathfrak{j})(\theta(\mathfrak{k}))} \circ t(f)_{\mathfrak{k}}$ in $(D(\mathfrak{i})(\mathfrak{k}) \downarrow U)$. This implies that

$$\begin{aligned} \mu_{\mathfrak{j}} \circ D(f) &= [\mathrm{in}_{D(\mathfrak{j})}]^{\curlywedge}_{U} \circ [t(f)]^{\curlywedge}_{D(\mathfrak{j})} \\ &= [\mathrm{in}_{D(\mathfrak{j})} * t(f)]^{\curlywedge}_{U} = [\mathrm{in}_{D(\mathfrak{i})}]^{\curlywedge}_{U} = \mu_{\mathfrak{i}} . \end{aligned}$$

It remains to show that μ is limiting. Let $\mu' \in \mathrm{Cone}(D, E)$ be a cocone where $\mu'_{\mathfrak{i}} = [w^{\mathfrak{i}}]^{\curlywedge}_{E}$, and let $[w]^{\curlywedge}_{E}$ be the union of μ'. We have seen in Lemma 1.12 (b) that $[w]^{\curlywedge}_{E}$ is a morphism $U \to E$. Furthermore,

$$[w]^{\curlywedge}_{E} * \mu = \big([w^{\mathfrak{i}}]^{\curlywedge}_{E} \circ [\mathrm{in}_{D(\mathfrak{i})}]^{\curlywedge}_{U}\big)_{\mathfrak{i} \in \mathcal{I}} = \big([w^{\mathfrak{i}}]^{\curlywedge}_{E}\big)_{\mathfrak{i} \in \mathcal{I}} = (\mu'_{\mathfrak{i}})_{\mathfrak{i} \in \mathcal{I}} = \mu' .$$

Hence, the function $[w]^{\curlywedge}_{E} \mapsto [w]^{\curlywedge}_{E} * \mu$ is an inverse to the bijective function of Lemma 1.12 (c). By Lemma B3.4.2 it follows that μ is limiting.

□

It turns out that $\mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$ is the closure of $\mathcal{C}$ under κ-filtered colimits of size less than λ, i.e., it is the smallest category containing $\mathcal{C}$ that is closed under such colimits. We begin the proof with a technical lemma summarising properties of the inclusion functor $\mathcal{C} \to \mathrm{Ind}_{\mathcal{P}}(\mathcal{C})$.

Lemma 1.14. *Let $\mathcal{C}$ be a category, $\mathcal{P}$ a class of small categories containing the singleton category $[1]$, and be $I : \mathcal{C} \to \mathrm{Ind}_{\mathcal{P}}(\mathcal{C})$ be the inclusion functor.*

(a) *I is well-defined.*

(b) *For links s and t from $D \in \mathrm{Ind}_{\mathcal{P}}(\mathcal{C})$ to $I(\mathfrak{a})$,*

$$[s]^{\wedge\!\!\!\wedge}_{I(\mathfrak{a})} = [t]^{\wedge\!\!\!\wedge}_{I(\mathfrak{a})} : D \to I(\mathfrak{a}) \quad \textit{implies} \quad s = t\,.$$

(c) *I is full and faithful.*

(d) *For every $D \in \mathrm{Ind}_{\mathcal{P}}(\mathcal{C})$, the inclusion $[\mathrm{in}_D]^{\wedge\!\!\!\wedge}_U : D \to U$ is an isomorphism, where U is the union of $I \circ D$.*

(e) *For every $D \in \mathrm{Ind}_{\mathcal{P}}(\mathcal{C})$ and every object $\mathfrak{a} \in \mathcal{C}$, I induces an isomorphism*

$$\mathrm{Cone}(D, \mathfrak{a}) \to \mathrm{Ind}_{\mathcal{P}}(\mathcal{C})(D, I(\mathfrak{a})) : \mu \mapsto I[\mu]\,.$$

(f) *A family t is a link from a diagram $D : \mathcal{I} \to \mathcal{C}$ to $I(\mathfrak{a})$ if, and only if, t is a cocone from D to $\mathfrak{a}$.*

Proof. To keep notation simple, we will not distinguish below between a morphism $f : \mathfrak{a} \to \mathfrak{b}$ of $\mathcal{C}$ and the link $t = (t_i)_{i \in [1]}$ from $I(\mathfrak{a})$ to $I(\mathfrak{b})$ whose only component is $t_0 = f$.

(a) Clearly, $I(\mathfrak{a}) \in \mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$, for every object $\mathfrak{a} \in \mathcal{C}$. Furthermore, if $f : \mathfrak{a} \to \mathfrak{b}$ is a morphism of $\mathcal{C}$, then the family $I(f)$ consisting just of f is a link from $I(\mathfrak{a})$ to $I(\mathfrak{b})$ since it only has to satisfy the trivial requirement that $f \circ I(\mathrm{id}_{\mathfrak{a}}) \wedge\!\!\!\wedge_{I(\mathfrak{b})} f$.

(b) Let $\mathfrak{i} \in \mathcal{I}$. Since $[s]^{\wedge\!\!\!\wedge}_{I(\mathfrak{a})} = [t]^{\wedge\!\!\!\wedge}_{I(\mathfrak{a})}$, the comma category $(D(\mathfrak{i}) \downarrow I(\mathfrak{a}))$ contains an alternating path from $s_{\mathfrak{i}}$ to $t_{\mathfrak{i}}$. As $\mathrm{id}_{\mathfrak{a}}$ is the only morphism of $I(\mathfrak{a})$, this alternating path consists only of identity morphisms. Consequently, $s_{\mathfrak{i}} = t_{\mathfrak{i}}$.

(c) To show that I is full, let $[f]^{\wedge\!\!\!\wedge}_{I(\mathfrak{b})} : I(\mathfrak{a}) \to I(\mathfrak{b})$ be a morphism of $\mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$. Then $f = (f_i)_{i\in[0]}$ consists just of one morphism $f_0 : \mathfrak{a} \to \mathfrak{b}$ and $I(f_0) = [f]^{\wedge\!\!\!\wedge}_{I(\mathfrak{b})}$.

To prove that I is faithful, suppose that $I(f) = I(g)$ for morphisms $f, g : \mathfrak{a} \to \mathfrak{b}$. Then $[f]^{\wedge\!\!\!\wedge}_{I(\mathfrak{b})} = [g]^{\wedge\!\!\!\wedge}_{I(\mathfrak{b})}$ and (b) implies that $f = g$.

(d) Let $D : \mathcal{I} \to \mathcal{C}$ be an object of $\mathrm{Ind}_{\mathcal{P}}(\mathcal{C})$ and let $U : \mathcal{J} \to \mathcal{C}$ be the union of $I \circ D$. Note that $\mathcal{J}^{\mathrm{obj}} = \mathcal{I}^{\mathrm{obj}} \times [1]$. Since $[\mathrm{in}_D]^{\wedge\!\!\!\wedge}_U : D \to U$ only consists of identity morphisms $\mathrm{id}_{D(\mathfrak{i})} : D(\mathfrak{i}) \to U(\langle \mathfrak{i}, 0\rangle)$, it has an inverse $[t]^{\wedge\!\!\!\wedge}_D : U \to D$ where

$$t_{\langle \mathfrak{i},0\rangle} := \mathrm{id}_{D(\mathfrak{i})} : U(\langle \mathfrak{i}, 0\rangle) \to D(\mathfrak{i}), \quad \text{for } \langle \mathfrak{i}, 0\rangle \in \mathcal{J}.$$

Furthermore, as both families only consist of identity morphisms, it is straightforward to check that they are links.

(e) By (d), D is the union of $I \circ D$. Hence, the morphism

$$\mathrm{Cone}(D, \mathfrak{a}) \to \mathrm{Ind}_{\mathcal{P}}(\mathcal{C})(D, I(\mathfrak{a})) : \mu \mapsto I[\mu]$$

can be written as composition of the natural isomorphisms

$$\begin{aligned} &\tau_{I,D} : \mathrm{Cone}(D, \mathfrak{a}) \to \mathrm{Cone}(I \circ D, I(\mathfrak{a})) : \mu \mapsto I[\mu] \\ \text{and}\quad &\eta_{I(\mathfrak{a})} : \mathrm{Cone}(I \circ D, I(\mathfrak{a})) \to \mathrm{Ind}_{\mathrm{all}}(\mathcal{C})(D, I(\mathfrak{a})), \end{aligned}$$

where $\eta_{I(\mathfrak{a})}$ is the morphism from Lemma 1.12 (c).

(f) ($\Leftarrow$) Let t be a cocone from D to $\mathfrak{a}$. For every morphism $f : \mathfrak{i} \to \mathfrak{j}$ of $\mathcal{I}$, we have $t_{\mathfrak{j}} \circ D(f) = t_{\mathfrak{i}}$, which implies that $t_{\mathfrak{j}} \circ D(f) \wedge\!\!\!\wedge_{I(\mathfrak{a})} t_{\mathfrak{i}}$.

($\Rightarrow$) Let t be a link from D to $I(\mathfrak{a})$. By (e), there is a unique cocone $\mu \in \mathrm{Cone}(D, \mathfrak{a})$ such that $I[\mu] = [t]^{\wedge\!\!\!\wedge}_{I(\mathfrak{a})}$. Hence, (b) implies that $\mu = t$. In particular, $t \in \mathrm{Cone}(D, \mathfrak{a})$. □

Theorem 1.15. *Let $\mathcal{C}$ be a category, κ, λ regular cardinals (or $\lambda = \infty$), and $I : \mathcal{C} \to \mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$ the inclusion functor.*

(a) *Every κ-filtered diagram $D : \mathcal{I} \to \mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$ of size less than λ has a colimit in $\mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$.*

(b) *For every object* $\mathfrak{a} \in \mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$, *there exists a* κ*-filtered diagram* $D : \mathcal{I} \to \mathcal{C}$ *of size less than* λ *such that* $\mathfrak{a} = \varinjlim (I \circ D)$.

Proof. (a) follows immediately from Proposition 1.13.

(b) Let $D \in \mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$. By Lemma 1.14 (e), D is isomorphic to the union of $I \circ D$. Consequently, it follows by Proposition 1.13 that $D \cong \varinjlim (I \circ D)$. □

Exercise 1.1. Prove the following universal property of $\mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$: for every functor $F : \mathcal{C} \to \mathcal{D}$ into a category $\mathcal{D}$ that has κ-directed colimits of size less than λ, there exists a unique functor $G : \mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C}) \to \mathcal{D}$ such that G preserves κ-filtered colimits of size less than λ and F factorises as $F = G \circ I$, where $I : \mathcal{C} \to \mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$ is the inclusion functor.

Remark. For every κ-filtered diagram $D : \mathcal{I} \to \mathcal{C}$ of size less than λ, the inductive completion $\mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$ has a colimit: the diagram D itself. But note that, if D already has a colimit $\mathfrak{a}$ in $\mathcal{C}$, the corresponding object $I(\mathfrak{a})$ of $\mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$ will in general not be a colimit. In fact, a limiting cocone λ from D to $\mathfrak{a}$ induces a morphism $[\lambda]^{\wedge\!\!\wedge}_{I(\mathfrak{a})} : D \to I(\mathfrak{a})$ in $\mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$, but there is no reason why this morphism should be an isomorphism.

2. *Extensions of diagrams*

In this section we consider ways to extend a diagram $D : \mathcal{I} \to \mathcal{C}$ to a diagram $D^+ : \mathcal{I}^+ \to \mathcal{C}$ with a larger index category. For instance, given a κ-directed diagram and a cardinal $\lambda \geq \kappa$, we would like to construct a λ-directed diagram with the same colimit.

Completions of directed orders

We start by transforming κ-directed partial orders into λ-directed ones.

Definition 2.1. Let $\mathfrak{I}$ be a partial order and κ, λ infinite cardinals or $\lambda = \infty$. The (κ, λ)-*completion* of $\mathfrak{I}$ is the partial order $\mathfrak{I}^+ := \langle I^+, \subseteq \rangle$

where

$$I^+ := \{\, \Downarrow S \mid S \subseteq I \text{ is } \kappa\text{-directed and } |S| < \lambda \,\} .$$

Our hope is that, using a generalisation of Lemma B3.3.5, we can prove that the (κ, λ)-completion of a κ-directed partial order is λ-directed. Unfortunately, this is not true in general. In only holds for certain cardinals κ and λ.

Before characterising such cardinals, we compare the (κ, λ)-completion of a κ-directed partial order $\mathfrak{I}$ to its inductive completion. It turns out that these two categories are equivalent. Before presenting the proof, let us note that the inductive completion of a preorder is again a preorder.

Lemma 2.2. *Let κ and λ be infinite cardinals or $\lambda = \infty$. If $\mathfrak{I}$ is a preorder, then so is* $\mathrm{Ind}_\kappa^\lambda(\mathfrak{I})$.

Proof. We have to prove that between any two objects $D : \mathcal{J} \to \mathfrak{I}$ and $E : \mathcal{K} \to \mathfrak{I}$ of $\mathrm{Ind}_\kappa^\lambda(\mathfrak{I})$, there is at most one morphism. Consider two links s and t from D to E. We claim that $s \approx_E t$. Let ρ and θ be the index maps of, respectively, s and t and let $\mathfrak{j} \in \mathcal{J}$. As E is κ-filtered, there exist an index $\mathfrak{k} \in \mathcal{K}$ and morphisms $g : \rho(\mathfrak{j}) \to \mathfrak{k}$ and $h : \theta(\mathfrak{j}) \to \mathfrak{k}$. It follows that $E(g) \circ s_\mathfrak{j}$ and $E(h) \circ t_\mathfrak{j}$ are both morphisms from $D(\mathfrak{j})$ to $E(\mathfrak{k})$. Since $\mathfrak{I}$ is a preorder, this implies that $E(g) \circ s_\mathfrak{j} = E(h) \circ t_\mathfrak{j}$. Consequently, g and h form an alternating path from $s_\mathfrak{j}$ to $t_\mathfrak{j}$ in $(D(\mathfrak{j}) \downarrow E)$. This implies that $s_\mathfrak{j} \approx_E t_\mathfrak{j}$. □

Proposition 2.3. *Let $\mathfrak{I}$ be a partial order and let κ, λ be infinite cardinals or $\lambda = \infty$. The (κ, λ)-completion $\mathfrak{I}^+$ of $\mathfrak{I}$ is equivalent to* $\mathrm{Ind}_\kappa^\lambda(\mathfrak{I})$.

Proof. It is sufficient to prove that the function

$$h : \mathrm{Ind}_\kappa^\lambda(\mathfrak{I}) \to \mathfrak{I}^+ : D \mapsto \Downarrow \mathrm{rng}\, D^{\mathrm{obj}}$$

is a surjective strict homomorphism. Then h induces a full and faithful functor $\mathrm{Ind}_\kappa^\lambda(\mathfrak{I}) \to \mathfrak{I}^+$. Since, trivially, every object of $\mathfrak{I}^+$ is isomorphic to some object in the image of this functor, it follows by Theorem B1.3.14 that the functor is an equivalence.

Let $D : \mathcal{J} \to \mathfrak{J}$ and $E : \mathcal{K} \to \mathfrak{J}$ be diagrams in $\mathrm{Ind}^{\lambda}_{\kappa}(\mathfrak{J})$. To see that h is a homomorphism, suppose that there exists a morphism $[t]^{\Lsh}_{E} : D \to E$. Let θ be the index map of t. Then the morphisms $t_{\mathfrak{j}} : D(\mathfrak{j}) \to E(\theta(\mathfrak{j}))$ witness that $D(\mathfrak{j}) \leq E(\theta(\mathfrak{j}))$, for all $\mathfrak{j} \in \mathcal{J}$. This implies that

$$\operatorname{rng} D^{\mathrm{obj}} \subseteq {\Downarrow} \operatorname{rng} E^{\mathrm{obj}}.$$

Hence, $h(D) \subseteq h(E)$.

For strictness, suppose that $h(D) \subseteq h(E)$. Then $\operatorname{rng} D^{\mathrm{obj}} \subseteq {\Downarrow} \operatorname{rng} E^{\mathrm{obj}}$ implies that, for every index $\mathfrak{j} \in \mathcal{J}$, we can find some index $\theta(\mathfrak{j}) \in \mathcal{K}$ such that $D(\mathfrak{j}) \leq E(\theta(\mathfrak{j}))$. Setting

$$t_{\mathfrak{j}} := \langle D(\mathfrak{j}), E(\theta(\mathfrak{j}))\rangle\,, \quad \text{for } \mathfrak{j} \in \mathcal{J}\,,$$

we obtain a link from D to E with index map θ.

It remains to prove that h is surjective. Let $S \in I^{+}$. Then $S = {\Downarrow} S_{0}$, for a κ-directed set $S_{0} \subseteq I$ of size $|S_{0}| < \lambda$. Let $D : \mathfrak{J} \restriction S_{0} \to \mathfrak{J}$ be the inclusion functor. Then $D \in \mathrm{Ind}^{\lambda}_{\kappa}(\mathfrak{J})$ and $h(D) = {\Downarrow} S_{0} = S$. □

If the (κ, λ)-completion is equivalent to the inductive completion, why did we introduce it? The reason is that we would like to extend a κ-directed diagram $D : \mathfrak{J} \to \mathcal{C}$ to a λ-directed one $D^{+} : \mathfrak{J}^{+} \to \mathcal{C}$. We cannot take the category $\mathrm{Ind}^{\lambda}_{\kappa}(\mathfrak{J})$ as index category $\mathfrak{J}^{+}$ since it is not small. Instead, we can use the skeleton of $\mathrm{Ind}^{\lambda}_{\kappa}(\mathfrak{J})$, which is small and isomorphic to the (κ, λ)-completion of $\mathfrak{J}$.

Before doing so, we sill have to characterise the cardinals κ, λ such that the (κ, λ)-completion is λ-directed. This is achieved by the following relation.

Definition 2.4. For infinite cardinals κ, λ, we write $\kappa \trianglelefteq \lambda$ if $\kappa \leq \lambda$ and, for every set X of size $|X| < \lambda$, there exists a set $D \subseteq \wp_{\kappa}(X)$ of size $|D| < \lambda$ that is dense in the partial order $\langle \wp_{\kappa}(X), \subseteq \rangle$, where

$$\wp_{\kappa}(X) := \{\, S \subseteq X \mid |S| < \kappa \,\}\,.$$

Exercise 2.1. Let κ be a regular cardinal. Prove that a set $D \subseteq \wp_\kappa(X)$ is dense if, and only if, $\langle D, \subseteq \rangle$ is κ-directed and $\bigcup D = X$.

The next lemma summarises the basic properties of the relation $\trianglelefteq$.

Lemma 2.5. *Let* $\mathrm{Cn}_{\aleph_0}$ *be the class of all infinite cardinals.*

(a) $\trianglelefteq$ *is a partial order on* $\mathrm{Cn}_{\aleph_0}$.

(b) $\kappa \triangleleft \kappa^+$, *for every regular cardinal* κ.

(c) *If* $\kappa < \lambda$ *are cardinals such that* $\mu^{<\kappa} < \lambda$, *for all* $\mu < \lambda$, *then* $\kappa \triangleleft \lambda$.

(d) $\kappa \triangleleft (2^{<\lambda})^+$ *for all cardinals* $\kappa \leq \lambda$.

(e) *The partial order* $\langle \mathrm{Cn}_{\aleph_0}, \trianglelefteq \rangle$ *is* κ*-directed for every cardinal* κ.

Proof. (a) The relation $\trianglelefteq$ is antisymmetric since, by definition, $\kappa \trianglelefteq \lambda$ implies $\kappa \leq \lambda$. For reflexivity, let X be a set of size $|X| < \kappa$. Then $X \in \wp_\kappa(X)$ and the set $D := \{X\}$ is dense. It remains to prove transitivity. Suppose that $\kappa \trianglelefteq \lambda \trianglelefteq \mu$. If $\lambda = \mu$, we are done. Hence, suppose that $\lambda \triangleleft \mu$. To show that $\kappa \trianglelefteq \mu$, let X be a set of size $|X| < \mu$. Since $\lambda \triangleleft \mu$, there exists a dense set $D \subseteq \wp_\lambda(X)$ of size $|D| < \mu$. Since $\kappa \trianglelefteq \lambda$, we can choose, for every $Y \in D$, a dense set $E_Y \subseteq \wp_\kappa(Y)$ of size $|E_Y| < \lambda$. Set

$$F := \bigcup_{Y \in D} E_Y \,.$$

Then $|F| \leq \sum_{Y \in D} |E_Y| \leq \lambda \otimes |D| < \mu$. Hence, it remains to prove that F is dense. Let $U \in \wp_\kappa(X)$. Then $U \in \wp_\lambda(X)$ and there is some $Y \in D$ with $U \subseteq Y$. Therefore, we can find a set $Z \in E_Y \subseteq F$ with $U \subseteq Z$.

(b) Let X be a set of size $|X| < \kappa^+$. Choose an injective map $f : X \to \kappa$. We claim that the set

$$D := \{\, f^{-1}[\downarrow\alpha] \mid \alpha < \kappa \,\}$$

is dense in $\wp_\kappa(X)$. First, note that $|f^{-1}[\downarrow\alpha]| \leq |\alpha| < \kappa$, for each $\alpha < \kappa$. Hence, $D \subseteq \wp_\kappa(X)$.

Given $Y \in \wp_\kappa(X)$, set $\gamma := \sup f[Y]$. Since $|f[Y]| < \kappa$ and κ is regular, it follows that $\gamma < \kappa$. Hence, $Y \subseteq f^{-1}[\downarrow(\gamma + 1)] \in D$.

(c) Let X be a set of size $\mu := |X| < \lambda$. Then $|\wp_\kappa(X)| = \mu^{<\kappa} < \lambda$. Hence, $D := \wp_\kappa(X)$ is a dense set of size less than λ.

(d) Let $\kappa \leq \lambda$ and set $\mu := (2^{<\lambda})^+$. Then

$$\begin{aligned}(<\mu)^{<\kappa} = (2^{<\lambda})^{<\kappa} &= \sup\{ (2^{\lambda_0})^{\kappa_0} \mid \kappa_0 < \kappa,\ \lambda_0 < \lambda \} \\ &= \sup\{ 2^{\lambda_0 \otimes \kappa_0} \mid \kappa_0 < \kappa,\ \lambda_0 < \lambda \} \leq 2^{<\lambda} < \mu\,.\end{aligned}$$

Hence, (c) implies that $\kappa \lhd \mu$.

(e) Let X be a set of cardinals. We set $\mu := \sup X$ and $\lambda := (2^{<\mu})^+$. By (d), it follows that $\kappa \lhd \lambda$, for every $\kappa \leq \mu$. Hence, λ is an upper bound of X. □

Exercise 2.2. Prove that $\aleph_0 \trianglelefteq \lambda$, for all infinite cardinals λ.

Example. To show that the relation $\trianglelefteq$ is non-trivial, we prove that $\aleph_1 \ntrianglelefteq \aleph_{\omega+1}$ by showing that there is no dense set $D \subseteq \wp_{\aleph_1}(\aleph_\omega)$ of size $|D| \leq \aleph_\omega$. For a contradiction, suppose that D is such a dense set. Fix a surjective function $f : \aleph_\omega \to D$. Since

$$\bigcup f[\downarrow\aleph_n] \leq \aleph_n \otimes \aleph_0 = \aleph_n < \aleph_{n+1}\,,$$

we can pick, for every $n < \omega$, an element $z_n \in \aleph_{n+1} \smallsetminus \bigcup f[\downarrow\aleph_n]$. Set $Z := \{ z_n \mid n < \omega \}$. Then $Z \in \wp_{\aleph_1}(\aleph_\omega)$ and, as D is dense, there exists a set $Y \in D$ with $Z \subseteq Y$. Since f is surjective, there is some $y \in \aleph_\omega$ with $f(y) = Y$. Fix an index $n < \omega$ with $y \in \aleph_n$. Then

$$z_n \in \aleph_{n+1} \smallsetminus \bigcup f[\downarrow\aleph_n] \supseteq \aleph_{n+1} \smallsetminus Y$$

implies that $Z \nsubseteq Y$. A contradiction.

For regular cardinals we can characterise the relation $\trianglelefteq$ in several different equivalent ways. One of them solves our question regarding the (κ, λ)-completion. Further characterisations will be given in Theorem 4.9 below.

Theorem 2.6. *Let $\kappa \leq \lambda$ be regular cardinals. The following statements are equivalent:*

(1) $\kappa \trianglelefteq \lambda$

(2) *For each κ-directed set $\mathfrak{I}$, every subset $X \subseteq I$ of size $|X| < \lambda$ is contained in a κ-directed subset $H \subseteq I$ of size $|H| < \lambda$.*

(3) *The (κ, λ)-completion of a κ-directed partial order is λ-directed.*

(4) $\mathrm{Ind}^{\lambda}_{\kappa}(\mathfrak{I})$ *is λ-directed, for every κ-directed partial order $\mathfrak{I}$.*

Proof. (1) ⇒ (2) Let I be a κ-directed partial order and let $X \subseteq I$ be a set of size $|X| < \lambda$. If $\lambda = \kappa$, the set X has an upper bound $c \in I$ and $X \cup \{c\}$ is the desired κ-directed set containing X. Therefore, we may assume that $\lambda > \kappa$. For the construction of H, we consider the following operation $B : \wp_\lambda(I) \to \wp_\lambda(I)$. Given $U \in \wp_\lambda(I)$, we define $B(U) \in \wp_\lambda(I)$ as follows. Choose a dense set $D \subseteq \wp_\kappa(U)$ of size $|U| < \lambda$ and, for every $Z \in D$, fix an upper bound $k_Z \in I$ of $Z \subseteq I$. We set

$$B(U) := U \cup \{ k_Z \mid Z \in D \} .$$

Then $U \subseteq B(U)$ and $|B(U)| \leq |U| \oplus |D| < \lambda$.

Using this operation, we define an increasing sequence $(H^\alpha(U))_{\alpha \leq \kappa}$ of sets by

$$\begin{aligned} H^0(U) &:= U , \\ H^{\alpha+1}(U) &:= B(H^\alpha(U)) , \\ H^\delta(U) &:= \bigcup_{\alpha<\delta} H^\alpha(U) , \quad \text{for limit ordinals } \delta . \end{aligned}$$

By induction on α, it follows that $|H^\alpha(U)| < \lambda$, for $\alpha \leq \kappa$ and $|U| < \lambda$. We claim that $H^\kappa(S)$ is the desired κ-directed set containing S. Let $U \subseteq H^\kappa(S)$ be a set of size $|U| < \kappa$. Since κ is regular, there is some ordinal α such that $U \subseteq H^\alpha(S)$. Consequently, $H^{\alpha+1}(S) \subseteq H^\kappa(S)$ contains an upper bound of U.

(2) ⇒ (3) Let $\mathfrak{I}^+$ be the (κ, λ)-completion of a κ-directed partial order $\mathfrak{I}$ and let $X \subseteq I^+$ be a set of size $|X| < \lambda$. By definition of I^+, there exists a family X_0 of κ-directed subsets $s \subseteq I$ of size $|s| < \lambda$ such that $X = \{ \Downarrow s \mid s \in X_0 \}$. Set $S := \bigcup X_0$. Since λ is regular, we have $|S| < \lambda$.

By (2), we can find a κ-directed set $H \subseteq I$ such that $S \subseteq H$ and $|H| < \lambda$. For each $s \in X_0$, $s \subseteq H$ implies that $\Downarrow s \subseteq \Downarrow H$. Hence, $\Downarrow H \in I^+$ is an upper bound of X.

(3) ⇔ (4) Let $\mathfrak{J}$ be a κ-directed partial order and let $\mathfrak{J}^+$ be its (κ, λ)-completion. We have seen in Proposition 2.3 that the categories $\mathrm{Ind}^\lambda_\kappa(\mathfrak{J})$ and $\mathfrak{J}^+$ are equivalent. Hence, the former is λ-directed if, and only if, the latter is λ-directed.

(4) ⇒ (1) Let X be a set of size $|X| < \lambda$. Note that, since κ is regular, we have $\bigcup Z \in \wp_\kappa(X)$, for every subset $Z \subseteq \wp_\kappa(X)$ of size $|Z| < \kappa$. Consequently, $\langle \wp_\kappa(X), \subseteq \rangle$ is κ-directed. By (4), it follows that $\mathrm{Ind}^\lambda_\kappa(\wp_\kappa(X))$ is λ-directed. Therefore, the preorder $\mathrm{Ind}^\lambda_\kappa(\wp_\kappa(X))$ contains an upper bound $D : \mathcal{I} \to \wp_\kappa(X)$ of the set $\{ I(\{x\}) \mid x \in X \}$, where $I : \wp_\kappa(X) \to \mathrm{Ind}^\lambda_\kappa(\wp_\kappa(X))$ is the inclusion functor. For $x \in X$, let θ_x be the index map of the link from $I(\{x\})$ to D. Then $\{x\} \subseteq D(\theta_x(0))$, for all $x \in X$.

We claim that $\mathrm{rng}\, D^{\mathrm{obj}}$ is a dense subset of $\wp_\kappa(X)$. Let $Y \in \wp_\kappa(X)$. Since D is κ-filtered, there exist an index $\mathfrak{k} \in \mathcal{I}$ and morphisms $f_y : \theta_y(0) \to \mathfrak{k}$, for $y \in Y$. Consequently,

$$\{y\} \subseteq D(\theta_y(0)) \subseteq D(\mathfrak{k}) \quad \text{implies} \quad Y \subseteq D(\mathfrak{k}) \in \mathrm{rng}\, D^{\mathrm{obj}}. \qquad \square$$

Extensions of directed diagrams

Having found a λ-directed completion $\mathfrak{J}^+$ of a given κ-directed partial order $\mathfrak{J}$, we can use it to extend κ-directed diagrams $D : \mathfrak{J} \to \mathcal{C}$ to a λ-directed diagram $D^+ : \mathfrak{J}^+ \to \mathcal{C}$. This construction is defined via a detour through the inductive completion $\mathrm{Ind}^\lambda_\kappa(\mathcal{C})$. We construct two diagrams $\mathfrak{J}^+ \to \mathrm{Ind}^\lambda_\kappa(\mathcal{C})$ and $\mathrm{Ind}^\lambda_\kappa(\mathcal{C}) \to \mathcal{C}$ whose composition is the extension $\mathfrak{J}^+ \to \mathcal{C}$ we are looking for. Let us start with the first diagram.

Definition 2.7. (a) Let $D : \mathcal{I} \to \mathcal{C}$ be a diagram and $F \subseteq \wp(\mathcal{I}^{\mathrm{obj}})$. The *F-completion* of D is the diagram

$$D^+ : \langle F, \subseteq \rangle \to \mathrm{Ind}_{\mathrm{all}}(\mathcal{C})$$

defined by

$$D^+(S) := D \restriction S\,, \qquad \text{for objects } S \in F\,,$$
$$D^+(S,T) := [\mathrm{in}_{D\restriction S}]^{\approx}_{D\restriction T}\,, \qquad \text{for pairs } S \subseteq T\,.$$

(b) Let $\mathfrak{I}$ be a partial order, $D : \mathfrak{I} \to \mathcal{C}$ a diagram, and κ, λ cardinals or $\lambda = \infty$. The (κ,λ)-*completion* of D is the I^+-completion $D^+ : \mathfrak{I}^+ \to \mathrm{Ind}_{\mathrm{all}}(\mathcal{C})$ of D, where $\mathfrak{I}^+$ is the (κ,λ)-completion of $\mathfrak{I}$.

For well-behaved sets F, the F-completion preserves the colimit.

Lemma 2.8. *Let $F \subseteq \wp(\mathcal{I}^{\mathrm{obj}})$ be a directed set with $\bigcup F = \mathcal{I}^{\mathrm{obj}}$ and let D^+ be the F-completion of $D : \mathcal{I} \to \mathcal{C}$. Then $\varinjlim D^+ \cong D$.*

Proof. Let $U : \mathcal{J} \to \mathcal{C}$ be the union of D^+ where, for each pair $S \subseteq T$, we have chosen the representative $u^{S,T} := \mathrm{in}_{D\restriction S}$ of the equivalence class $D^+(S,T) = [u^{S,T}]^{\approx}_{D\restriction T}$. By Proposition 1.13 it is sufficient to show that $U \cong D$. For $\langle S, \mathfrak{i}\rangle \in \mathcal{J} = \dot\bigcup_{S\in F} S$, set

$$s_{\langle S,\mathfrak{i}\rangle} := \mathrm{id}_{D(\mathfrak{i})} : U(\langle S,\mathfrak{i}\rangle) \to D(\mathfrak{i})\,.$$

For every $\mathfrak{i} \in \mathcal{I}$, choose a set $\theta(\mathfrak{i}) \in F$ with $\mathfrak{i} \in \theta(\mathfrak{i})$ and set

$$t_{\mathfrak{i}} := \mathrm{id}_{D(\mathfrak{i})} : D(\mathfrak{i}) \to U(\langle \theta(\mathfrak{i}),\mathfrak{i}\rangle)\,.$$

We claim that $s := (s_{\langle S,\mathfrak{i}\rangle})_{\langle S,\mathfrak{i}\rangle\in\mathcal{J}}$ and $t := (t_{\mathfrak{i}})_{\mathfrak{i}\in\mathcal{I}}$ are links from, respectively, U to D and D to U such that $[s]^{\approx}_D : U \to D$ is an inverse of $[t]^{\approx}_U : D \to U$.

We start by showing that s and t are a links. For t, let $f : \mathfrak{i} \to \mathfrak{j}$ be a morphism of $\mathcal{I}$ and choose a set $S \in F$ with $\mathfrak{i}, \mathfrak{j} \in S$. Then

$$\begin{aligned} u_{\mathfrak{j}}^{\theta(\mathfrak{j}),S} \circ t_{\mathfrak{j}} \circ D(f) &= \mathrm{id}_{D(\mathfrak{j})} \circ \mathrm{id}_{D(\mathfrak{j})} \circ D(f) \\ &= D(f) \circ \mathrm{id}_{D(\mathfrak{i})} \circ \mathrm{id}_{D(\mathfrak{i})} \\ &= U(D(f)) \circ u_{\mathfrak{i}}^{\theta(\mathfrak{i}),S} \circ t_{\mathfrak{i}}\,. \end{aligned}$$

Hence, $u_{\mathfrak{j}}^{\theta(\mathfrak{j}),S}$ and $D(f) \circ u_{\mathfrak{i}}^{\theta(\mathfrak{i}),S}$ form an alternating path from $t_{\mathfrak{j}} \circ D(f)$ to $t_{\mathfrak{i}}$ in $(D(\mathfrak{i}) \downarrow U)$.

For s, note that $\mathcal{J}$ is generated by morphisms of the form $D(f)$ and $u_{\mathfrak{i}}^{S,T}$, for $f \in \mathcal{I}^{\mathrm{mor}}$, $S \subseteq T$, and $\mathfrak{i} \in \mathcal{I}^{\mathrm{obj}}$. Hence, it is sufficient to check that

$$s_{\langle T,\mathfrak{j}\rangle} \circ U(h) \curlywedge_D s_{\langle S,\mathfrak{i}\rangle} \quad \text{for such morphisms } h\,.$$

For $h = u_{\mathfrak{i}}^{S,T}$, we have

$$s_{\langle T,\mathfrak{i}\rangle} \circ U(u_{\mathfrak{i}}^{S,T}) = \mathrm{id}_{D(\mathfrak{i})} \circ \mathrm{id}_{D(\mathfrak{i})} = \mathrm{id}_{D(\mathfrak{i})} = s_{\langle S,\mathfrak{i}\rangle}\,.$$

For $h = D(f)$ with $f : \mathfrak{i} \to \mathfrak{j}$ in $\mathcal{I}$,

$$\begin{aligned} D(\mathrm{id}_{\mathfrak{j}}) \circ s_{\langle S,\mathfrak{j}\rangle} \circ U(D(f)) &= D(\mathrm{id}_{\mathfrak{j}}) \circ \mathrm{id}_{D(\mathfrak{j})} \circ D(f) \\ &= D(f) \circ \mathrm{id}_{D(\mathfrak{i})} \\ &= D(f) \circ s_{\langle S,\mathfrak{i}\rangle} \end{aligned}$$

implies that $s_{\langle S,\mathfrak{j}\rangle} \circ U(D(f)) \curlywedge_D s_{\langle S,\mathfrak{i}\rangle}$.

It remains to prove that $[s]_D^{\curlywedge}$ is an inverse of $[t]_U^{\curlywedge}$. Since

$$s * t = (s_{\langle\theta(\mathfrak{i}),\mathfrak{i}\rangle} \circ t_{\mathfrak{i}})_{\mathfrak{i}\in\mathcal{I}} = (\mathrm{id}_{D(\mathfrak{i})})_{\mathfrak{i}\in\mathcal{I}}\,,$$

s is a left inverse of t. To show that it is also a right inverse, let $\langle S, \mathfrak{i}\rangle \in \mathcal{J}$ and fix a set $T \in F$ with $\theta(\mathfrak{i}) \cup S \subseteq T$. Then

$$\begin{aligned} U(u_{\mathfrak{i}}^{\theta(\mathfrak{i}),T}) \circ (t * s)_{\langle S,\mathfrak{i}\rangle} &= \mathrm{id}_{D(\mathfrak{i})} \circ t_{\mathfrak{i}} \circ s_{\langle S,\mathfrak{i}\rangle} \\ &= \mathrm{id}_{D(\mathfrak{i})} \circ \mathrm{id}_{D(\mathfrak{i})} \circ \mathrm{id}_{D(\mathfrak{i})} \\ &= U(\mathrm{id}_{D(\mathfrak{i})}) \circ \mathrm{id}_{U(\langle S,\mathfrak{i}\rangle)} \end{aligned}$$

implies that $(t * s)_{\langle S,\mathfrak{i}\rangle} \curlywedge_U \mathrm{id}_{U(\langle S,\mathfrak{i}\rangle)}$. □

The second step of the construction uses the following functor to go back to the category $\mathcal{C}$.

Definition 2.9. Let $\mathcal{C}$ be a category with $\mathcal{P}$-colimits. Fixing, for every diagram $D \in \mathrm{Ind}_{\mathcal{P}}(\mathcal{C})$, a limiting cocone $\lambda^D \in \mathrm{Cone}(D, \mathfrak{a}_D)$ of D, we define the *canonical projection functor*

$$Q : \mathrm{Ind}_{\mathcal{P}}(\mathcal{C}) \to \mathcal{C}$$

as follows. Q^{obj} maps diagrams $D \in \mathrm{Ind}_{\mathcal{P}}(\mathcal{C})$ to their colimit $\mathfrak{a}_D$. For morphisms $[t]^{\Join}_E : D \to E$, we choose for $Q^{\mathrm{mor}}([t]^{\Join}_E)$ the unique morphism $\varphi : \mathfrak{a}_D \to \mathfrak{a}_E$ such that

$$\lambda^E * t = \varphi \circ \lambda^D .$$

Lemma 2.10. *Let $\mathcal{P}$ be a class of small categories containing the singleton category $[1]$, $\mathcal{C}$ a category with $\mathcal{P}$-colimits, and let $Q : \mathrm{Ind}_{\mathcal{P}}(\mathcal{C}) \to \mathcal{C}$ be the canonical projection functor.*

(a) *Q is well-defined.*

(b) *Q preserves colimits.*

Proof. Let $(\lambda^D)_D$ be the family of limiting cocones used to define Q and let $(\mathfrak{a}_D)_D$ be the corresponding colimits.

(a) Clearly, the object part Q^{obj} is well-defined. Hence, it remains to check the morphism part Q^{mor}. First note that, for a link t from D to E, we have shown in Lemma B3.5.8 that $\lambda^E * t$ is a cocone of D. As λ^D is limiting, there therefore exists a unique morphism φ such that

$$\lambda^E * t = \varphi \circ \lambda^D .$$

It remains to show that this morphism φ does not depend on the choice of the representative t. Suppose that $s \Join_E t$. Then

$$\lambda^E * s \Join_{I(\mathfrak{a})} \lambda^E * t$$

and it follows by Lemma 1.14 (b) that $\lambda^E * s = \lambda^E * t$.

(b) Let λ^* be a limiting cocone from $D : \mathcal{I} \to \mathrm{Ind}_{\mathcal{P}}(\mathcal{C})$ to E. By Lemma B3.4.5, $Q[\lambda^*]$ is a cocone from $Q \circ D$ to $Q(E) = \mathfrak{a}_E$. Hence, it remains to show that $Q[\lambda^*]$ is limiting.

Let $\mu \in \text{Cone}(Q \circ D, \mathfrak{b})$ be a cocone. We have to find a unique morphism $\varphi : \mathfrak{a}_E \to \mathfrak{b}$ such that $\mu = \varphi * Q[\lambda^*]$. For $\mathfrak{i} \in \mathcal{I}$, set

$$\nu_{\mathfrak{i}} := [\mu_{\mathfrak{i}} * \lambda^{D(\mathfrak{i})}]^{\mathbb{M}}_{I(\mathfrak{b})}\,.$$

We claim that $\nu := (\nu_{\mathfrak{i}})_{\mathfrak{i}\in\mathcal{I}}$ is a cocone from D to $I(\mathfrak{b})$.

Let $f : \mathfrak{i} \to \mathfrak{j}$ be a morphism of $\mathcal{I}$ and suppose that $D(f) = [t]^{\mathbb{M}}_{D(\mathfrak{j})}$. Note that, by definition of Q,

$$\lambda^{D(\mathfrak{j})} * t = Q(D(f)) * \lambda^{D(\mathfrak{i})}.$$

Since μ is a cocone of $Q \circ D$, it follows that

$$\begin{aligned} \nu_{\mathfrak{j}} \circ D(f) &= \big[\mu_{\mathfrak{j}} * \lambda^{D(\mathfrak{j})}\big]^{\mathbb{M}}_{I(\mathfrak{b})} \circ D(f) \\ &= \big[\mu_{\mathfrak{j}} * \lambda^{D(\mathfrak{j})} * t\big]^{\mathbb{M}}_{I(\mathfrak{b})} \\ &= \big[\mu_{\mathfrak{j}} * Q(D(f)) * \lambda^{D(\mathfrak{i})}\big]^{\mathbb{M}}_{I(\mathfrak{b})} \\ &= \big[\mu_{\mathfrak{i}} * \lambda^{D(\mathfrak{i})}\big]^{\mathbb{M}}_{I(\mathfrak{b})} = \nu_{\mathfrak{i}}\,, \end{aligned}$$

as desired.

As ν is a cocone of D and λ^* is limiting, there exists a unique morphism $[t]^{\mathbb{M}}_{I(\mathfrak{b})} : E \to I(\mathfrak{b})$ such that

$$\nu = [t]^{\mathbb{M}}_{I(\mathfrak{b})} * \lambda^*.$$

By Lemma 1.14 (f) it follows that t is a cocone from E to $\mathfrak{b}$. As λ^E is limiting, there exists a unique morphism $\varphi : \mathfrak{a}_E \to \mathfrak{b}$ such that $t = \varphi * \lambda^E$. Suppose that $\lambda^*_{\mathfrak{i}} = [s^{\mathfrak{i}}]^{\mathbb{M}}_E$. Then

$$Q(\lambda^*_{\mathfrak{i}}) * \lambda^{D(\mathfrak{i})} = \lambda^E * s^{\mathfrak{i}}$$

implies that

$$\big[Q(\lambda^*_{\mathfrak{i}}) * \lambda^{D(\mathfrak{i})}\big]^{\mathbb{M}}_{I(\mathfrak{a}_E)} = \big[\lambda^E\big]^{\mathbb{M}}_{I(\mathfrak{a}_E)} * \lambda^*_{\mathfrak{i}}\,.$$

For every $i \in \mathcal{I}$, it follows that

$$\begin{aligned}\left[\varphi * Q(\lambda_{\mathfrak{i}}^*) * \lambda^{D(\mathfrak{i})}\right]^{\wedge\!\!\wedge}_{I(\mathfrak{b})} &= \left[\varphi * \lambda^E\right]^{\wedge\!\!\wedge}_{I(\mathfrak{b})} * \lambda_{\mathfrak{i}}^* \\ &= [t]^{\wedge\!\!\wedge}_{I(\mathfrak{b})} * \lambda_{\mathfrak{i}}^* = \nu_{\mathfrak{i}} = [\mu_{\mathfrak{i}} * \lambda^{D(\mathfrak{i})}]^{\wedge\!\!\wedge}_{I(\mathfrak{b})} .\end{aligned}$$

Using Lemma 1.14 (b), it follows that

$$\varphi * Q(\lambda_{\mathfrak{i}}^*) * \lambda^{D(\mathfrak{i})} = \mu_{\mathfrak{i}} * \lambda^{D(\mathfrak{i})},$$

which, by Lemma B3.4.2, implies that $\varphi \circ Q(\lambda_{\mathfrak{i}}^*) = \mu_{\mathfrak{i}}$. Hence,

$$\mu = \varphi * Q[\lambda^*] .$$

It remains to prove that the morphism φ is unique. Suppose that $\psi : \mathfrak{a}_E \to \mathfrak{b}$ is a morphism such that $\mu = \psi * Q[\lambda^*]$. Then

$$\begin{aligned}\left[\psi * \lambda^E\right]^{\wedge\!\!\wedge}_{I(\mathfrak{b})} * \lambda_{\mathfrak{i}}^* &= \left[\psi * Q(\lambda_{\mathfrak{i}}^*) * \lambda^{D(\mathfrak{i})}\right]^{\wedge\!\!\wedge}_{I(\mathfrak{b})} \\ &= [\mu_{\mathfrak{i}} * \lambda^{D(\mathfrak{i})}]^{\wedge\!\!\wedge}_{I(\mathfrak{b})} = \nu_{\mathfrak{i}} = [t]^{\wedge\!\!\wedge}_{I(\mathfrak{b})} * \lambda_{\mathfrak{i}}^* ,\end{aligned}$$

and it follows by Lemma B3.4.2 that

$$[\psi * \lambda^E]^{\wedge\!\!\wedge}_{I(\mathfrak{b})} = [t]^{\wedge\!\!\wedge}_{I(\mathfrak{b})} .$$

Hence, Lemma 1.14 (b) implies that $t = \psi * \lambda^E$. By choice of φ, it follows that $\psi = \varphi$. □

Combining these two functors we obtain the desired λ-directed extension.

Proposition 2.11. *Let $\kappa \trianglelefteq \lambda$ and let $\mathcal{C}$ be a category with κ-directed colimits of size less than λ. For every κ-directed diagram $D : \mathfrak{I} \to \mathcal{C}$, there exists a λ-directed diagram $D^+ : \mathfrak{I}^+ \to \mathcal{C}$ such that*

$$\varinjlim D^+ \cong \varinjlim D$$

and, for every $i \in I^+$, there is some κ-directed set $S \subseteq I$ of size $|S| < \lambda$ such that

$$D^+(i) \cong \varinjlim (D \restriction S) .$$

Proof. Let $D^+ : \mathfrak{I}^+ \to \mathrm{Ind}^\lambda_\kappa(\mathcal{C})$ be the (κ, λ)-completion of D. By Theorem 2.6 (3), the diagram D^+ is λ-directed. Furthermore, we have seen in Lemma 2.8 that $\varinjlim D^+ \cong D$. According to Lemma 2.10, the canonical projection functor $Q : \mathrm{Ind}^\lambda_\kappa(\mathcal{C}) \to \mathcal{C}$ preserves colimits. Hence, it follows that

$$\varinjlim (Q \circ D^+) = Q(\varinjlim D^+) \cong Q(D) \cong \varinjlim D\,.$$

Furthermore, each index $i \in I^+$ is of the form $i = \Downarrow S$ for some κ-directed set $S \subseteq I$ of size $|S| < \lambda$. Since S is dense in $\Downarrow S$, it follows that

$$Q(D^+(i)) \cong \varinjlim D^+(i) \cong \varinjlim (D \restriction \Downarrow S) \cong \varinjlim (D \restriction S)\,.$$

Hence, $Q \circ D^+ : \mathfrak{I}^+ \to \mathcal{C}$ is the desired diagram. □

Example. We can also use the previous results to give a short alternative proof of Proposition B3.4.16. Let $\mathcal{C}$ be a category with directed colimits and let $\mathcal{D}$ be the class of all directed partial orders. For $D \in \mathrm{Ind}_{\mathcal{D}}(\mathcal{C})$ of size κ, we find the desired chain C as follows.

By Proposition B3.3.6, there exists a chain $(H_\alpha)_{\alpha<\kappa}$ of directed subsets $H_\alpha \subseteq I$ of size $|H_\alpha| < \kappa$ such that $I = \bigcup_{\alpha<\kappa} H_\alpha$. Set $F := \{\, H_\alpha \mid \alpha < \kappa \,\}$, let D^+ be the F-completion of D, and let $Q : \mathrm{Ind}_{\mathcal{D}}(\mathcal{C}) \to \mathcal{C}$ be the canonical projection. As above,

$$\varinjlim (Q \circ D^+) = Q(\varinjlim D^+) \cong Q(D) \cong \varinjlim D\,.$$

Since $\langle F, \subseteq \rangle \cong \langle \kappa, \leq \rangle$ it follows that $C := Q \circ D^+$ is the desired chain.

Shifted diagrams

We conclude this section by presenting a second construction of diagrams. It provides a way to modify the colimit of a κ-filtered diagram $D : \mathcal{I} \to \mathcal{C}$ by adding morphisms to the index category $\mathcal{I}$ but no new objects. We will see below that this results in a retraction of the colimit.

Definition 2.12. Let $D : \mathcal{I} \to \mathcal{C}$ be a diagram.

(a) A morphism $f : \mathfrak{a} \to \mathfrak{a}$ is *idempotent* if $f \circ f = f$. Similarly, we call a link t from D to D *idempotent* if $t \circ t \mathrel{\curlywedge\!\!\!\curlywedge}_D t$.

(b) By $\circlearrowleft$ we denote the category with a single object $*$ and two morphisms id, $e : * \to *$ where $e \circ e = e$ and id is the identity morphism.

(c) Let t be an idempotent link from D to D, let $F : \circlearrowleft \to \mathrm{Ind}_{\mathrm{all}}(\mathcal{C})$ be the diagram mapping $*$ to D and e to $[t]^{\curlywedge\!\!\!\curlywedge}_*$, and let $D^+ : \mathcal{I}^+ \to \mathcal{C}$ be the union of F where we choose t as representative of $[t]^{\curlywedge\!\!\!\curlywedge}_D$. We say that D^+ is the diagram obtained by *shifting* the diagram D by t.

Our aim is to show that the colimit of a shifted diagram is a retract of the colimit of the original one. We also characterise which retracts we can obtain in this way. The key argument is a proof that, in certain categories, every idempotent morphism factorises as a retraction followed by a section.

Lemma 2.13. *Let $D : \circlearrowleft \to \mathcal{C}$ be a diagram. A cocone $\mu \in \mathrm{Cone}(D, \mathfrak{a})$ is limiting if, and only if, the morphism $\mu_* : D(*) \to \mathfrak{a}$ has a right inverse $s : \mathfrak{a} \to D(*)$ such that*

$$D(e) = s \circ \mu_* \,.$$

Proof. $(\Rightarrow)$ Since $D(e) \circ D(e) = D(e \circ e) = D(e)$, the family consisting just of the morphism $D(e)$ is a cocone from D to $D(*)$. If μ is limiting, we can therefore find a morphism $s : \mathfrak{a} \to D(*)$ such that $D(e) = s * \mu_*$.

We claim that s is the right inverse of μ_*. Since μ is a cocone, we have

$$\mu_* \circ s \circ \mu_* = \mu_* \circ D(e) = \mu_* \,,$$

which implies by Lemma B3.4.2 that $\mu_* \circ s = \mathrm{id}_{\mathfrak{a}}$.

$(\Leftarrow)$ Let s be a right inverse of μ_* such that $D(e) = s \circ \mu_*$. Given another cocone $\mu' \in \mathrm{Cone}(D, \mathfrak{b})$, we set $\varphi := \mu'_* \circ s$. Then

$$\mu'_* = \mu'_* \circ D(e) = \mu'_* \circ s \circ \mu_* = \varphi \circ \mu_*$$

implies that $\mu' = \varphi * \mu$. To show that φ is unique, suppose that $\mu' = \psi * \mu$. Then

$$\psi = \psi \circ (\mu_* \circ s) = \mu'_* \circ s = \varphi \circ \mu_* \circ s = \varphi \,. \qquad \square$$

Corollary 2.14. *Let $\mathcal{C}$ be a category with finite κ-filtered colimits, for some cardinal κ. A morphism $p : \mathfrak{a} \to \mathfrak{a}$ is idempotent if, and only if, $p = s \circ r$ for some retraction $r : \mathfrak{a} \to \mathfrak{b}$ with right inverse $s : \mathfrak{b} \to \mathfrak{a}$.*

Proof. (⇒) Let $p : \mathfrak{a} \to \mathfrak{a}$ be idempotent and let $D : \circlearrowleft \to \mathcal{C}$ be the diagram mapping the object $*$ to $\mathfrak{a}$ and the morphism e to p. By assumption, D has a limiting cocone λ to some object $\mathfrak{b}$. Consequently, it follows by Lemma 2.13 that the morphism $r := \lambda_*$ has a right inverse s with $s \circ r = D(e) = p$.

(⇐) Let r be a retraction with right inverse s. Since $(s \circ r) \circ (s \circ r) = s \circ \mathrm{id} \circ r = s \circ r$, every morphism of the form $s \circ r$ is idempotent. □

One consequence of Lemma 2.13 is that every diagram D^+ obtained by shifting a diagram D is a retract of D in $\mathrm{Ind}_{\mathrm{all}}(\mathcal{C})$. For the proof that the same holds for their colimits, we start with a technical lemma.

Lemma 2.15. *Let $D^+ : \mathcal{I}^+ \to \mathcal{C}$ be the diagram obtained by shifting a filtered diagram $D : \mathcal{I} \to \mathcal{C}$ by an idempotent link t.*

(a) *t is a link from D^+ to D.*

(b) *Let $\mu \in \mathrm{Cone}(D, \mathfrak{a})$. Then*

$$\mu \in \mathrm{Cone}(D^+, \mathfrak{a}) \quad \text{iff} \quad \mu * t = \mu\,.$$

Proof. (a) Note that the morphism $[t]^{⩕}_D : D \to D$ forms a cocone from $F : \circlearrowleft \to \mathrm{Ind}_{\mathrm{all}}(\mathcal{C})$ to D whose union is just $[t]^{⩕}_D$. Therefore, Lemma 1.12 (b) implies that t is a link from D^+ to D.

(b) (⇒) Let θ be the index map of t. If μ is a cocone of D^+, then $\mu_{\theta(\mathrm{i})} \circ t_{\mathrm{i}} = \mu_{\mathrm{i}}$, which implies that

$$\mu * t = (\mu_{\theta(\mathrm{i})} \circ t_{\mathrm{i}})_{\mathrm{i} \in \mathcal{I}} = (\mu_{\mathrm{i}})_{\mathrm{i} \in \mathcal{I}}\,.$$

(⇐) If $\mu * t = \mu$, then it follows by (a) and Lemma B3.5.8 that

$$\mu = \mu * t = \pi_t(\mu) \in \mathrm{Cone}(D^+, \mathfrak{a})\,. \qquad \square$$

Proposition 2.16. *Let $D^+ : \mathcal{I}^+ \to \mathcal{C}$ be the diagram obtained by shifting a filtered diagram $D : \mathcal{I} \to \mathcal{C}$ by an idempotent link t and let λ be a limiting cocone from D to some object $\mathfrak{a}$. For an object $\mathfrak{b} \in \mathcal{C}$, the following two statements are equivalent.*

(1) $\varinjlim D^+ \cong \mathfrak{b}$

(2) *There exists a retraction $r : \mathfrak{a} \to \mathfrak{b}$ with right inverse $e : \mathfrak{b} \to \mathfrak{a}$ satisfying*

$$\lambda * t = (e \circ r) * \lambda .$$

Proof. (1) ⇒ (2) Let λ^+ be a limiting cocone form D^+ to $\mathfrak{b}$. Since $\lambda * t \in \mathrm{Cone}(D^+, \mathfrak{a})$ and $\lambda^+ \in \mathrm{Cone}(D, \mathfrak{b})$, there exist unique morphisms $r : \mathfrak{a} \to \mathfrak{b}$ and $e : \mathfrak{b} \to \mathfrak{a}$ such that

$$\lambda * t = e * \lambda^+ \quad \text{and} \quad \lambda^+ = r * \lambda .$$

By Lemma 2.15 (b), it follows that

$$\begin{aligned}(r \circ e) * \lambda^+ &= r * (e * \lambda^+) \\ &= r * (\lambda * t) \\ &= (r * \lambda) * t = \lambda^+ * t = \lambda^+ = \mathrm{id} * \lambda^+ .\end{aligned}$$

Therefore, Lemma B3.4.2 implies that $r \circ e = \mathrm{id}$. Consequently, $r : \mathfrak{a} \to \mathfrak{b}$ is a retraction with section $e : \mathfrak{b} \to \mathfrak{a}$. Furthermore,

$$\lambda * t = e * \lambda^+ = e * (r * \lambda) = (e \circ r) * \lambda .$$

(2) ⇒ (1) We claim that $\lambda^+ := r * \lambda$ is a limiting cocone from D^+ to $\mathfrak{b}$. Since

$$\begin{aligned}\lambda^+ * t = (r * \lambda) * t &= r * (\lambda * t) \\ &= r * ((e \circ r) * \lambda) \\ &= (r \circ e \circ r) * \lambda = r * \lambda = \lambda^+ ,\end{aligned}$$

Lemma 2.15 (b) implies that $\lambda^+ \in \mathrm{Cone}(D^+, \mathfrak{b})$. To see that λ^+ is limiting, we prove that the natural transformation

$$\eta : \mathcal{C}(\mathfrak{b}, -) \to \mathrm{Cone}(D^+, -) : f \mapsto f * \lambda^+$$

from Lemma B3.4.2 is a natural isomorphism.

We start by showing that each component $\eta_\mathfrak{c}$ of η is surjective. Let $\mu \in \mathrm{Cone}(D^+, \mathfrak{c})$. Since $\mu \in \mathrm{Cone}(D, \mathfrak{c})$ and λ is limiting, there exists a unique morphism $\varphi : \mathfrak{a} \to \mathfrak{c}$ such that $\mu = \varphi * \lambda$. Consequently,

$$\begin{aligned} \mu = \mu * t = \varphi * \lambda * t &\\ &= \varphi * (e \circ r) * \lambda \\ &= (\varphi \circ e) * (r * \lambda) \\ &= (\varphi \circ e) * \lambda^+ = \eta_\mathfrak{c}(\varphi \circ e) \in \operatorname{rng} \eta_\mathfrak{c} . \end{aligned}$$

For injectivity, suppose that $f, f' : \mathfrak{b} \to \mathfrak{c}$ are two morphisms such that $\eta_\mathfrak{c}(f) = \eta_\mathfrak{c}(f')$. Since

$$(f \circ r) * \lambda = f * (r * \lambda) = f * \lambda^+ = \eta_\mathfrak{c}(f)$$

and, analogously, $(f' \circ r) * \lambda = \eta_\mathfrak{c}(f')$, it follows that

$$(f \circ r) * \lambda = (f' \circ r) * \lambda .$$

By Lemma B3.4.2, this implies that $f \circ r = f' \circ r$. Since r is an epimorphism, we obtain $f = f'$, as desired. □

3. *Presentable objects*

When trying to find a category-theoretical generalisation of statements involving the cardinality of structures, one needs a notion of cardinality for the objects of a category. Of course, one could simply add a function $\mathcal{C}^{\mathrm{obj}} \to \mathrm{Cn}$ to a category $\mathcal{C}$ and axiomatise its properties. But it is not obvious what such axioms should look like. It turns out that, for certain categories, there is a simpler way. Without explicitly adding a notion of cardinality, we can recover it from the category. To do so we introduce the concept of a *κ-presentable* object, which generalises the concept of a κ-generated structure in $\mathfrak{Emb}(\Sigma)$.

Definition 3.1. Let $\mathcal{C}$ be a category and κ a cardinal.

(a) Let $D : \mathcal{I} \to \mathcal{C}$ be a diagram and $\mu \in \operatorname{Cone}(D, \mathfrak{b})$ a cocone. A morphism $f : \mathfrak{a} \to \mathfrak{b}$ *factorises through* μ if there exists an object $\mathfrak{i} \in \mathcal{I}$ and a morphism $f_0 : \mathfrak{a} \to D(\mathfrak{i})$ such that

$$f = \mu_{\mathfrak{i}} \circ f_0 .$$

We say that this factorisation is *essentially unique* if, for every other factorisation $f = \mu_{\mathfrak{k}} \circ f'_0$ with $\mathfrak{k} \in \mathcal{I}$ and $f'_0 : \mathfrak{a} \to D(\mathfrak{k})$, we have

$$f_0 \approx_D f'_0 .$$

(b) An object $\mathfrak{a}$ of $\mathcal{C}$ is *κ-presentable* if, for each κ-directed diagram $D : \mathfrak{J} \to \mathcal{C}$ with colimit $\mathfrak{b}$, every morphism $f : \mathfrak{a} \to \mathfrak{b}$ factorises essentially uniquely through the limiting cocone. For $\kappa = \aleph_0$, we call $\mathfrak{a}$ *finitely presentable.*

Remark. (a) Let $\kappa \leq \lambda$. Since each λ-directed diagram is also κ-directed, it follows that κ-presentable objects are λ-presentable.

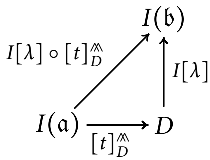
(b) For a singular cardinal κ, it follows by Lemma 1.4 that an object is κ-presentable if, and only if, it is κ^+-presentable.

Example. In $\mathfrak{Set}$ every set X is $|X|^+$-presentable.

Exercise 3.1. Prove that an object $\mathfrak{a}$ is κ-presentable if, and only if, for every κ-filtered diagram D with limiting cocone $\lambda \in \operatorname{Cone}(D, \mathfrak{b})$, the function

$$\begin{aligned} &\operatorname{Ind}_{\mathrm{all}}(\mathcal{C})(I(\mathfrak{a}), I[\lambda]) \\ &\quad : \operatorname{Ind}_{\mathrm{all}}(\mathcal{C})(I(\mathfrak{a}), D) \to \operatorname{Ind}_{\mathrm{all}}(\mathcal{C})(I(\mathfrak{a}), I(\mathfrak{b})) \\ &\quad : [t]_D^{\approx} \mapsto I[\lambda] \circ [t]_D^{\approx} \end{aligned}$$

is bijective. (I denotes the inclusion functor $\mathcal{C} \to \operatorname{Ind}_{\mathrm{all}}(\mathcal{C})$.)



Exercise 3.2. Let $D : \mathcal{I} \to \mathcal{C}$ be a κ-filtered diagram with a κ-presentable colimit $\mathfrak{a}$, and let λ be a limiting cocone from D to $\mathfrak{a}$. Prove that, in $\mathrm{Ind}^{\infty}_{\kappa}(\mathcal{C})$, the morphism $I[\lambda] : D \cong I(\mathfrak{a})$ induced by λ is an isomorphism.

First, let us show that this notion indeed generalises the concept of being κ-generated.

Proposition 3.2. *Let κ be a regular cardinal. A Σ-structure $\mathfrak{A}$ is κ-presentable in the category $\mathfrak{Emb}(\Sigma)$ if, and only if, it is κ-generated.*

Proof. $(\Rightarrow)$ Let $\mathfrak{A}$ be κ-presentable. To show that $\mathfrak{A}$ is κ-generated, let $\mathfrak{J}$ be the family of all κ-generated substructures of $\mathfrak{A}$ ordered by inclusion and let $D : \mathfrak{J} \to \mathcal{C}$ be the canonical diagram. By Proposition B3.3.16, this diagram is κ-directed and its colimit is $\mathfrak{A}$. Let λ be the limiting cocone. Since $\mathfrak{A}$ is κ-presentable, the identity $\mathrm{id}_{\mathfrak{A}} : \mathfrak{A} \to \mathfrak{A}$ factorises through λ. Therefore, we can find an index $k \in I$ and an embedding $f : \mathfrak{A} \to D(k)$ such that $\lambda_k \circ f = \mathrm{id}_{\mathfrak{A}}$. As $\lambda_k \circ f = \mathrm{id}_{\mathfrak{A}}$ is surjective, so is the embedding λ_k. Consequently, λ_k is an isomorphism and $\mathfrak{A} \cong D(k)$ is κ-generated.

$(\Leftarrow)$ Suppose that $\mathfrak{A}$ is generated by a set $X \subseteq A$ of size $|X| < \kappa$. To show that $\mathfrak{A}$ is κ-presentable, let $D : \mathfrak{J} \to \mathfrak{Emb}(\Sigma)$ be a κ-directed diagram with colimit $\mathfrak{B}$ and $f : \mathfrak{A} \to \mathfrak{B}$ an embedding. Let $\lambda \in \mathrm{Cone}(D, \mathfrak{B})$ be a limiting cocone. For every element $a \in X$, fix an index $i_a \in I$ with $f(a) \in \mathrm{rng}\,\lambda_{i_a}$ and let k be an upper bound of $\{ i_a \mid a \in X \}$. Then

$$f[X] \subseteq \bigcup_{a \in X} \mathrm{rng}\,\lambda_{i_a} \subseteq \mathrm{rng}\,\lambda_k\,,$$

which implies that $\mathrm{rng}\,f \subseteq \mathrm{rng}\,\lambda_k$. By Lemma A2.1.10, there exists a right inverse $g : \mathrm{rng}\,\lambda_k \to D(k)$ of λ_k. We set $f_0 := g \circ f$. Then

$$\lambda_k \circ f_0 = \lambda_k \circ g \circ f = f\,.$$

It remains to show that the factorisation is essentially unique. Hence, suppose that there is an index $i \in I$ and an embedding $f'_0 : \mathfrak{A} \to D(i)$ such that $\lambda_i \circ f'_0 = f$. For every element $a \in X$,

$$\lambda_i(f'_0(a)) = f(a) = \lambda_k(f_0(a))$$

implies, by the definition of a κ-directed limit of Σ-structures, that there is some index $l_a \geq i, k$ such that

$$D(i,l)(f'_0(a)) = D(k,l)(f_0(a)) .$$

Choosing an upper bound m of $\{ l_a \mid a \in X \}$, we obtain

$$D(i,m) \circ f'_0 = D(k,m) \circ f_0 .$$

This implies that $f'_0 \curlywedge_D f_0$. □

Let us present several alternative characterisations of being κ-presentable. The first one rests on the fact that, since every κ-filtered colimit can be written as a κ-directed one, we can replace in the definition κ-directed diagrams by κ-filtered ones. The second characterisation is based on hom-functors.

Theorem 3.3. *Let $\mathcal{C}$ be a category and $\mathfrak{a}$ an object. The following statements are equivalent:*

(1) *$\mathfrak{a}$ is κ-presentable.*

(2) *For each κ-filtered diagram $D : \mathcal{I} \to \mathcal{C}$ with colimit $\mathfrak{b}$, every morphism $f : \mathfrak{a} \to \mathfrak{b}$ factorises essentially uniquely through the limiting cocone.*

(3) *The covariant hom-functor $\mathcal{C}(\mathfrak{a}, -)$ preserves κ-directed colimits.*

(4) *The covariant hom-functor $\mathcal{C}(\mathfrak{a}, -)$ preserves κ-filtered colimits.*

Proof. (4) ⇒ (3) is trivial.

(3) ⇒ (1) Let $D : \mathfrak{I} \to \mathcal{C}$ be a κ-directed diagram with limiting cocone $\lambda \in \mathrm{Cone}(D, \mathfrak{b})$, and let $f : \mathfrak{a} \to \mathfrak{b}$ be a morphism. By assumption $(\mathcal{C}(\mathfrak{a}, \lambda_i))_{i \in I}$ is a limiting cocone of $\mathcal{C}(\mathfrak{a}, -) \circ D$. Consequently,

$$\mathcal{C}(\mathfrak{a}, \mathfrak{b}) = \bigcup_{i \in I} \mathcal{C}(\mathfrak{a}, \lambda_i)[\mathcal{C}(\mathfrak{a}, D(i))] \, .$$

In particular, there are an index $i \in I$ and a morphism $f_\circ \in \mathcal{C}(\mathfrak{a}, D(i))$ with

$$f = \mathcal{C}(\mathfrak{a}, \lambda_i)(f_\circ) = \lambda_i \circ f_\circ \, .$$

Hence, f factorises through λ. For essential uniqueness, suppose that there is a second index $j \in I$ and a morphism $f_\circ' : \mathfrak{a} \to D(j)$ such that $f = \lambda_j \circ f_\circ'$. Then

$$\mathcal{C}(\mathfrak{a}, \lambda_j)(f_\circ') = \lambda_j \circ f_\circ' = \lambda_i \circ f_\circ = \mathcal{C}(\mathfrak{a}, \lambda_i)(f_\circ) \, .$$

Hence, $f_\circ \in \mathcal{C}(\mathfrak{a}, D(i))$ and $f_\circ' \in \mathcal{C}(\mathfrak{a}, D(j))$ correspond to the same element of the colimit $\mathcal{C}(\mathfrak{a}, \mathfrak{b})$. This implies that there exists an index $k \geq i, j$ such that

$$\mathcal{C}(\mathfrak{a}, D(i,k))(f_\circ) = \mathcal{C}(\mathfrak{a}, D(j,k))(f_\circ') \, .$$

Consequently,

$$D(i,k) \circ f_\circ = D(j,k) \circ f_\circ' \, ,$$

which implies that $f_\circ \wedge\!\!\!\wedge_D f_\circ'$.

(1) ⇒ (2) Let λ be a limiting cocone from D to $\mathfrak{b}$. By Theorem 1.7, there exists a dense κ-directed diagram $F : \mathfrak{K} \to \mathcal{I}$. Furthermore, according to Proposition B3.5.15, the projection $\pi_{D,F}$ along F is a natural isomorphism. Consequently, it follows by Lemma B3.4.3 that the projection $\mu := \pi_{D,F}(\lambda)$ is a limiting cocone from $D \circ F$ to $\mathfrak{b}$. Therefore, every morphism $f : \mathfrak{a} \to \mathfrak{b}$ factorises essentially uniquely through μ as $f = \mu_k \circ f_\circ$, for some $k \in K$ and $f_\circ : \mathfrak{a} \to D(F(k))$.

We claim that $\lambda_{F(k)} \circ f_0$ is an essentially unique factorisation of f through λ. Note that $\lambda_{F(k)} \circ f_0 = \mu_k \circ f_0 = f$ implies that it is a factorisation of f. Hence, it remains to prove essential uniqueness.

Suppose that $f = \lambda_{\mathfrak{i}} \circ f_0'$ is a second factorisation. As F is dense, there exists an index $l \in K$ and a morphism $g : \mathfrak{i} \to F(l)$. Hence,

$$\mu_k \circ f_0 \quad \text{and} \quad \mu_l \circ D(g) \circ f_0'$$

are two factorisations of f through μ and, by essential uniqueness, we obtain

$$f_0 \approx_{D \circ F} D(g) \circ f_0' .$$

By Lemma B3.5.3 (d), this implies that $f_0 \approx_D f_0'$.

(2) ⇒ (4) Let $D : \mathcal{I} \to \mathcal{C}$ be a κ-filtered diagram with limiting cocone $\lambda \in \mathrm{Cone}(D, \mathfrak{b})$. We have to show that $\lambda' := (\mathcal{C}(\mathfrak{a}, \lambda_{\mathfrak{i}}))_{\mathfrak{i} \in \mathcal{I}}$ is a limiting cocone from $\mathcal{C}(\mathfrak{a}, -) \circ D$ to $\mathcal{C}(\mathfrak{a}, \mathfrak{b})$. By Lemma B3.4.2, it is sufficient to prove that the natural transformation

$$\eta : \mathfrak{Set}(\mathcal{C}(\mathfrak{a}, \mathfrak{b}), -) \to \mathrm{Cone}(\mathcal{C}(\mathfrak{a}, -) \circ D, -) : \varphi \mapsto \varphi * \lambda'$$

is a natural isomorphism. We define an inverse ζ of η as follows.

For each morphism $f : \mathfrak{a} \to \mathfrak{b}$, we choose an essentially unique factorisation

$$f = \lambda_{\mathfrak{i}(f)} \circ g(f) , \quad \text{with } \mathfrak{i}(f) \in \mathcal{I} \text{ and } g(f) : \mathfrak{a} \to D(\mathfrak{i}(f)) ,$$

and, for a cocone μ of $\mathcal{C}(\mathfrak{a}, -) \circ D$ and a morphism $f : \mathfrak{a} \to \mathfrak{b}$, we set

$$\zeta(\mu)(f) := \mu_{\mathfrak{i}(f)}(g(f)) .$$

It remains to show that ζ is an inverse of η. First, note that $\zeta(\lambda') = \mathrm{id}$ since

$$\begin{aligned} \zeta(\lambda')(f) &= \lambda'_{\mathfrak{i}(f)}(g(f)) \\ &= \mathcal{C}(\mathfrak{a}, \lambda_{\mathfrak{i}(f)})(g(f)) = \lambda_{\mathfrak{i}(f)} \circ g(f) = f . \end{aligned}$$

Furthermore,

$$\begin{aligned}\zeta(\varphi * \mu)(f) &= (\varphi * \mu)_{\mathfrak{i}(f)}(g(f)) \\ &= \varphi\big(\mu_{\mathfrak{i}(f)}(g(f))\big) = \varphi(\zeta(\mu)(f))\end{aligned}$$

implies that $\zeta(\varphi * \mu) = \varphi \circ \zeta(\mu)$. Consequently,

$$\zeta(\eta(\varphi)) = \zeta(\varphi * \lambda') = \varphi \circ \zeta(\lambda') = \varphi \circ \mathrm{id} = \varphi\,.$$

To show that ζ is also a right inverse of η, note that, if $f = \lambda_{\mathfrak{j}} \circ f_o$ is an arbitrary factorisation of $f : \mathfrak{a} \to \mathfrak{b}$ through λ, it follows by essential uniqueness and Corollary 1.3, that there are morphisms $h : \mathfrak{i}(f) \to \mathfrak{k}$ and $h' : \mathfrak{j} \to \mathfrak{k}$ such that

$$D(h) \circ g(f) = D(h') \circ f_o\,.$$

For a cocone μ of $\mathcal{C}(\mathfrak{a}, -) \circ D$, it therefore follows that

$$\begin{aligned}\mu_{\mathfrak{i}(f)}(g(f)) &= \big(\mu_k \circ \mathcal{C}(\mathfrak{a}, D(h))\big)(g(f)) \\ &= \mu_k\big(D(h) \circ g(f)\big) \\ &= \mu_k(D(h') \circ f_o) \\ &= \big(\mu_k \circ \mathcal{C}(\mathfrak{a}, D(h'))\big)(f_o) = \mu_{\mathfrak{j}}(f_o)\,.\end{aligned}$$

Consequently,

$$\begin{aligned}\eta(\zeta(\mu)) = \zeta(\mu) * \lambda' &= \big(\zeta(\mu) \circ \mathcal{C}(\mathfrak{a}, \lambda_{\mathfrak{j}})\big)_{\mathfrak{j}\in\mathcal{I}} \\ &= \big(f_o \mapsto \mu_{\mathfrak{i}(\lambda_{\mathfrak{j}} \circ f_o)}(g(\lambda_{\mathfrak{j}} \circ f_o))\big)_{\mathfrak{j}\in\mathcal{I}} \\ &= (f_o \mapsto \mu_{\mathfrak{j}}(f_o))_{\mathfrak{j}\in\mathcal{I}} \\ &= (\mu_{\mathfrak{j}})_{\mathfrak{j}\in\mathcal{I}}\,.\end{aligned}$$

□

Exercise 3.3. Prove that a hom-functor $\mathcal{C}(\mathfrak{a}, -)$ always preserves limits.

Corollary 3.4. *Let $\mathfrak{a}$ be κ-representable and let $D : \mathcal{I} \to \mathcal{C}$ be a κ-filtered diagram with limiting cocone λ. If $f_i : \mathfrak{a} \to D(\mathfrak{k}_i)$, $i < \gamma$, is a family of $\gamma < \kappa$ morphisms with*

$$\lambda_{\mathfrak{k}_i} \circ f_i = \lambda_{\mathfrak{k}_j} \circ f_j\,, \quad \textit{for all } i, j < \gamma\,,$$

then there exist an object $\mathfrak{l} \in \mathcal{I}$ and morphisms $g_i : \mathfrak{k}_i \to \mathfrak{l}$, $i < \gamma$, such that

$$D(g_i) \circ f_i = D(g_j) \circ f_j\,, \quad \textit{for all } i, j < \gamma\,.$$

Proof. For every pair $i, j < \gamma$, we apply Theorem 3.3 (b) to the morphism $\lambda_{\mathfrak{k}_i} \circ f_i = \lambda_{\mathfrak{k}_j} \circ f_j$. By essential uniqueness and Corollary 1.3, there are morphisms $h_{ij} : \mathfrak{k}_i \to \mathfrak{l}_{ij}$ and $h'_{ij} : \mathfrak{k}_j \to \mathfrak{l}_{ij}$ such that

$$D(h_{ij}) \circ f_i = D(h'_{ij}) \circ f_j\,.$$

By Lemma 1.2, there exist an object $\mathfrak{m} \in \mathcal{I}$ and morphisms

$$g_i : \mathfrak{k}_i \to \mathfrak{m} \quad \text{and} \quad g_{ij} : \mathfrak{l}_{ij} \to \mathfrak{m}\,, \quad \text{for } i, j < \gamma\,,$$

such that

$$g_i = g_{ij} \circ h_{ij} \quad \text{and} \quad g_j = g_{ij} \circ h'_{ij}\,, \quad \text{for all } i, j < \gamma\,.$$

Consequently,

$$\begin{aligned} D(g_i) \circ f_i &= D(g_{ij}) \circ D(h_{ij}) \circ f_i \\ &= D(g_{ij}) \circ D(h'_{ij}) \circ f_j = D(g_j) \circ f_j\,. \end{aligned}$$ □

To prove that an object of a full subcategory is κ-presentable, the next lemma is sometimes useful.

Lemma 3.5. *Let $F : \mathcal{C} \to \mathcal{D}$ be a full and faithful functor that preserves κ-directed colimits. Then F reflects κ-presentable objects.*

Proof. Let $\mathfrak{a} \in \mathcal{C}$ be an object such that $F(\mathfrak{a})$ is κ-presentable. To show that $\mathfrak{a}$ is also κ-presentable, let $D : \mathfrak{J} \to \mathcal{C}$ be a κ-directed diagram with colimit $\mathfrak{b}$, let λ be a corresponding limiting cocone, and let $f : \mathfrak{a} \to \mathfrak{b}$ be a morphism. Then $F[\lambda]$ is a limiting cocone of the κ-directed diagram $F \circ D : \mathfrak{J} \to \mathcal{D}$. Hence, $F(f)$ factorises essentially uniquely as $F(f) = F(\lambda_i) \circ g$, for some $g : F(\mathfrak{a}) \to F(D(i))$. As F is full, we can find a morphism $f_0 : \mathfrak{a} \to D(i)$ with $F(f_0) = g$. Consequently, $F(f) = F(\lambda_i \circ f_0)$ which, by faithfulness of F, implies that $f = \lambda_i \circ f_0$.

We claim that this factorisation is essentially unique. Suppose that $f = \lambda_k \circ f_0'$ is a second factorisation. Then $F(f) = F(\lambda_k) \circ F(f_0')$ is a factorisation of $F(f)$ and it follows by essential uniqueness that

$$F(f_0) \curlywedge_{F \circ D} F(f_0') .$$

By Corollary 1.3, there exist an index $l \geq i, k$ such that

$$F(D(i,l)) \circ F(f_0) = F(D(k,l)) \circ F(f_0') .$$

Since F is faithful, this implies that

$$D(i,l) \circ f_0 = D(k,l) \circ f_0' .$$

Consequently, $f_0 \curlywedge_D f_0'$. ☐

Cardinality

In the next section we will define a notion of cardinality such that κ-presentable objects have size less than κ. The aim of the following results is to show that κ-presentability does indeed behave as we would expect for a notion of cardinality: an object consisting of λ parts of size less than κ has size less than $\kappa \oplus \lambda^+$. Before giving the proof, we start with a technical result about diagrams of κ-presentable objects.

Lemma 3.6. *Let $E : \mathcal{J} \to \mathcal{C}$ be a κ-filtered diagram with limiting cocone $\mu \in \mathrm{Cone}(E, \mathfrak{b})$, and let $D : \mathcal{I} \to \mathcal{C}$ a diagram where each object $D(\mathfrak{i})$ is κ-presentable.*

(a) *For all links s and t from D to E,*

$$s \approx_E t \quad \text{iff} \quad \mu * s = \mu * t\,.$$

(b) *Given a limiting cocone* $\lambda \in \mathrm{Cone}(D, \mathfrak{a})$ *and a morphism* $f : \mathfrak{a} \to \mathfrak{b}$*, there exists a link t from D to E such that*

$$\mu * t = f * \lambda\,.$$

Furthermore, this link t is unique up to a.p.-equivalence.

Proof. (a) Let ρ and θ be the index maps of, respectively, s and t. For every $\mathfrak{i} \in \mathcal{I}$, we have

$$s_{\mathfrak{i}} \approx_E t_{\mathfrak{i}} \quad \text{iff} \quad \mu_{\rho(\mathfrak{i})} \circ s_{\mathfrak{i}} = \mu_{\theta(\mathfrak{i})} \circ t_{\mathfrak{i}}\,,$$

where one direction follows by Lemma B3.5.4 and the other one by Theorem 3.3 (b), which implies that the morphism $\mu_{\rho(\mathfrak{i})} \circ s_{\mathfrak{i}} = \mu_{\theta(\mathfrak{i})} \circ t_{\mathfrak{i}}$ factorises essentially uniquely through μ.

(b) Since $D(i)$ is κ-presentable, it follows by Theorem 3.3 (b) that $f \circ \lambda_{\mathfrak{i}}$ has an essentially unique factorisation

$$f \circ \lambda_{\mathfrak{i}} = \mu_{\theta(\mathfrak{i})} \circ t_{\mathfrak{i}}\,,$$

where $\theta(\mathfrak{i}) \in \mathcal{I}$ and $t_{\mathfrak{i}} : D(\mathfrak{i}) \to E(\theta(\mathfrak{i}))$. Setting $t := (t_{\mathfrak{i}})_{\mathfrak{i} \in \mathcal{I}}$ it follows that

$$f * \lambda = \mu * t\,.$$

Hence, it remains to show that t is a link and that it is unique. For uniqueness, note that, according to (a)

$$\mu * t' = f * \lambda = \mu * t \quad \text{implies} \quad t' \approx_E t\,.$$

To show that t is a link, let $g : \mathfrak{i} \to \mathfrak{j}$ be a morphism of $\mathcal{I}$. Then

$$\mu_{\theta(\mathfrak{i})} \circ t_{\mathfrak{i}} = f \circ \lambda_{\mathfrak{i}} = f \circ \lambda_{\mathfrak{j}} \circ D(g) = \mu_{\theta(\mathfrak{j})} \circ t_{\mathfrak{j}} \circ D(g)$$

are two factorisations of the same morphism through μ. By essential uniqueness, it therefore follows that $t_{\mathfrak{i}} \approx_D t_{\mathfrak{j}} \circ D(g)$. □

Proposition 3.7. *Let $D : \mathcal{I} \to \mathcal{C}$ be a diagram where each $D(\mathfrak{i})$ is κ-presentable. If it exists, the colimit of D is $(\kappa \oplus |\mathcal{I}^{\mathrm{mor}}|^+)$-presentable.*

Proof. Let λ be a limiting cocone from D to $\mathfrak{a} \in \mathcal{C}$ and set $\mu := \kappa \oplus |\mathcal{I}^{\mathrm{mor}}|^+$. To show that $\mathfrak{a}$ is μ-presentable, consider a morphism $f : \mathfrak{a} \to \mathfrak{b}$ where $\mathfrak{b}$ is the colimit of a μ-directed diagram $E : \mathfrak{K} \to \mathcal{C}$. Let $\lambda' \in \mathrm{Cone}(E, \mathfrak{b})$ be the corresponding limiting cocone. By Lemma 3.6 (b), there exists a link t from D to E such that

$$\lambda' * t = f * \lambda\,.$$

Let $\theta : \mathcal{I}^{\mathrm{obj}} \to K$ be the index map of t. For $h : \mathfrak{i} \to \mathfrak{j}$ in $\mathcal{I}$, we have

$$\lambda'_{\theta(\mathfrak{i})} \circ t_\mathfrak{i} = f \circ \lambda_\mathfrak{i} = f \circ \lambda_\mathfrak{j} \circ D(h) = \lambda'_{\theta(\mathfrak{j})} \circ t_\mathfrak{j} \circ D(h)\,.$$

As $D(\mathfrak{i})$ is μ-presentable, it follows by essential uniqueness and Corollary 1.3 that we can find an index $k_h \in K$ such that

$$E(\theta(\mathfrak{i}), k_h) \circ t_\mathfrak{i} = E(\theta(\mathfrak{j}), k_h) \circ t_\mathfrak{j} \circ D(h)\,.$$

Let $l \in K$ be an upper bound of $\{\, k_h \mid h \in \mathcal{I}^{\mathrm{mor}} \,\}$ and set

$$\nu_\mathfrak{i} := E(\theta(\mathfrak{i}), l) \circ t_\mathfrak{i}\,, \quad \text{for } \mathfrak{i} \in \mathcal{I}\,.$$

Then $\nu = (\nu_\mathfrak{i})_{\mathfrak{i} \in \mathcal{I}}$ is a cocone from D to $E(l)$.

Since λ is limiting, there exists a morphism $\varphi : \mathfrak{a} \to E(l)$ such that $\nu = \varphi * \lambda$. It follows that

$$f \circ \lambda_\mathfrak{i} = \lambda'_{\theta(\mathfrak{i})} \circ t_\mathfrak{i} = \lambda'_l \circ E(\theta(\mathfrak{i}), l) \circ t_\mathfrak{i} = \lambda'_l \circ \nu_\mathfrak{i} = \lambda'_l \circ \varphi \circ \lambda_\mathfrak{i}\,,$$

for every $\mathfrak{i} \in \mathcal{I}$. By Lemma B3.4.2, this implies that $f = \lambda'_l \circ \varphi$.

It remains to check that φ is essentially unique. Suppose that there is a second morphism $\psi : \mathfrak{a} \to E(m)$, for some $m \in K$, such that $f = \lambda'_m \circ \psi$. For $\mathfrak{i} \in \mathcal{I}$, it follows that

$$\lambda'_m \circ \psi \circ \lambda_\mathfrak{i} = f \circ \lambda_\mathfrak{i} = \lambda'_l \circ \varphi \circ \lambda_\mathfrak{i}\,.$$

As $D(\mathfrak{i})$ is μ-presentable, it follows by essential uniqueness and Corollary 1.3 that there is an index $n_\mathfrak{i} \geq l, m$ such that

$$E(m, n_\mathfrak{i}) \circ \psi \circ \lambda_\mathfrak{i} = E(l, n_\mathfrak{i}) \circ \varphi \circ \lambda_\mathfrak{i} \,.$$

Let $n_* \in K$ be an upper bound of $\{ n_\mathfrak{i} \mid \mathfrak{i} \in \mathcal{I} \}$. Then

$$E(m, n_*) \circ \psi \circ \lambda_\mathfrak{i} = E(l, n_*) \circ \varphi \circ \lambda_\mathfrak{i} \,, \quad \text{for all } \mathfrak{i} \in \mathcal{I} \,.$$

Consequently, it follows by Lemma B3.4.2 that

$$E(m, n_*) \circ \psi = E(l, n_*) \circ \varphi \,.$$

This implies that $\psi \approx_E \varphi$. □

For the converse of this statement we need additional requirements on the category $\mathcal{C}$.

Theorem 3.8. *Let $\kappa \trianglelefteq \lambda$ be regular cardinals and $\mathcal{C}$ a category with κ-directed colimits of size less than λ. Suppose that there exists a class $\mathcal{K} \subseteq \mathcal{C}^{\mathrm{obj}}$ of κ-presentable objects such that every object of $\mathcal{C}$ can be written a κ-filtered colimit of objects in $\mathcal{K}$.*

An object $\mathfrak{a} \in \mathcal{C}$ is λ-presentable if, and only if, it is the colimit of a κ-filtered diagram $D : \mathcal{I} \to \mathcal{C}$ of size less than λ where each $D(\mathfrak{i}) \in \mathcal{K}$.

Proof. ($\Leftarrow$) was already shown in Proposition 3.7.

($\Rightarrow$) Let $\mathfrak{a}$ be λ-presentable and let $D : \mathfrak{I} \to \mathcal{C}$ be a κ-directed diagram with colimit $\mathfrak{a}$ such that each $D(\mathfrak{i})$ belongs to $\mathcal{K}$. Since $\kappa \trianglelefteq \lambda$, we can use Proposition 2.11 to find a λ-directed diagram $D^+ : \mathfrak{I}^+ \to \mathcal{C}$ with colimit $\mathfrak{a}$ such that, for every $i \in I^+$, there exists a κ-directed subset $S \subseteq I$ of size less than λ such that

$$D^+(i) \cong \varinjlim (D \restriction S) \,.$$

Let μ^+ be a limiting cocone from D^+ to $\mathfrak{a}$. Since $\mathfrak{a}$ is λ-presentable, there exists an essentially unique factorisation $\mathrm{id}_\mathfrak{a} = \mu^+_S \circ e$, for some index $i \in I^+$ and morphism $e : \mathfrak{a} \to D^+(i)$. Set

$$\mathfrak{b} := D^+(i) \quad \text{and} \quad r := \mu^+_i \,.$$

By construction of D^+, there exists a κ-directed subset $S \subseteq I$ of size $|S| < \lambda$ such that $D^+(i) \cong \varinjlim (D \upharpoonright S)$. Let μ be a limiting cocone form $D \upharpoonright S$ to $\mathfrak{b}$.

It follows that $r : \mathfrak{b} \to \mathfrak{a}$ is a retraction with right inverse $e : \mathfrak{a} \to \mathfrak{b}$. By Lemma 3.6 (b), there exists a link t from $D \upharpoonright S$ to $D \upharpoonright S$ such that

$$\mu * t = (e \circ r) * \mu .$$

Furthermore, according to Lemma 3.6 (a),

$$\begin{aligned} \mu * t * t &= (e \circ r) * \mu * t \\ &= (e \circ r) * (e \circ r) * \mu \\ &= (e \circ r \circ e \circ r) * \mu = (e \circ r) * \mu = \mu * t \end{aligned}$$

implies that $t \circ t \approx_D t$. Hence, the link t is idempotent and we can shift $D \upharpoonright S$ by t to obtain a diagram $E : \mathcal{J} \to \mathcal{C}$. By Proposition 1.13 and Proposition 2.16, it follows that E is a κ-filtered diagram of size less than λ and that $\varinjlim E \cong \mathfrak{a}$. Finally, note that, for every $\mathfrak{j} \in \mathcal{J}$, there is some $\mathfrak{i} \in \mathcal{I}$ with $E(\mathfrak{j}) = D(\mathfrak{i}) \in \mathcal{K}$. □

As a further indication that our notion of cardinality is well-behaved, let us conclude this section with the remark that retracts do not increase the size.

Proposition 3.9. *Every retract of a κ-presentable object is κ-presentable.*

Proof. Let $\mathfrak{a}$ be κ-presentable and let $r : \mathfrak{a} \to \mathfrak{b}$ be a retraction with right inverse $e : \mathfrak{b} \to \mathfrak{a}$. To show that $\mathfrak{b}$ is also κ-presentable, let $D : \mathfrak{I} \to \mathcal{C}$ be a κ-directed diagram with limiting cocone $\lambda \in \mathrm{Cone}(D, \mathfrak{c})$, and let $f : \mathfrak{b} \to \mathfrak{c}$ be a morphism. Since $\mathfrak{a}$ is κ-presentable, $f \circ r$ factorises essentially uniquely through λ as

$$f \circ r = \lambda_i \circ g , \quad \text{for some } g : \mathfrak{a} \to D(i) .$$

We obtain a factorisation

$$f = f \circ r \circ e = \lambda_i \circ g \circ e$$

of f. We claim that this factorisation is essentially unique.

Suppose that $f = \lambda_k \circ h$ is a second factorisation. Then $\lambda_k \circ (h \circ r)$ is a factorisation of $f \circ r$ and essential uniqueness implies that $g \approx_D h \circ r$. By Lemma B3.5.3 (b), it follows that

$$g \circ e \approx_D h \circ r \circ e = h\,,$$

as desired. □

4. Accessible categories

Using the notion of κ-presentability, we can define a class of categories where one can associate a cardinality with each object.

Definition 4.1. Let κ be a cardinal. A category $\mathcal{C}$ is *κ-accessible* if

- it has κ-directed colimits,
- every object $\mathfrak{a} \in \mathcal{C}$ is a κ-directed colimit of κ-presentable objects,
- up to isomorphism, there exists only a set of κ-presentable objects.

It follows by Proposition 3.7 that every object of a κ-accessible category is λ-presentable, for some cardinal λ. We can use this fact to define a notion of cardinality for the objects of such a category.

Definition 4.2. Let $\mathcal{C}$ be a κ-accessible category. The *cardinality* $\|\mathfrak{a}\|$ of an object $\mathfrak{a} \in \mathcal{C}$ is the least cardinal λ such that $\mathfrak{a}$ is λ^+-presentable.

Example. The categories $\mathfrak{Emb}(\Sigma)$ and $\mathfrak{Set}$ are κ-accessible, for all regular cardinals κ. We have $\|X\| = |X|$, for every infinite set $X \in \mathfrak{Set}$. Similarly, if $\mathfrak{A}$ is a Σ-structure in $\mathfrak{Emb}(\Sigma)$ with $|A_s| \geq |\Sigma|^+$, for every sort s, then $\|\mathfrak{A}\| = |A|$.

The following theorem immediately follows from Theorem 3.8.

Theorem 4.3. *Let $\kappa \trianglelefteq \lambda$ be regular cardinals and $\mathcal{C}$ a κ-accessible category. An object $\mathfrak{a} \in \mathcal{C}$ is λ-presentable if, and only if, it is the colimit of a κ-filtered diagram $D : \mathcal{I} \to \mathcal{C}$ of size less than λ where each $D(\mathfrak{i})$ is κ-presentable.*

Let us give some non-trivial examples of κ-accessible categories. The first one is the category of all κ-directed partial orders.

Definition 4.4. Let κ be a cardinal. We denote by $\mathfrak{Dir}(\kappa)$ the full subcategory of $\mathfrak{Emb}(\leq)$ induced by all κ-directed partial orders.

Proposition 4.5. *Let κ be a cardinal and let $J : \mathfrak{Dir}(\kappa) \to \mathfrak{Emb}(\leq)$ be the inclusion functor.*

(a) *For every κ-directed diagram $D : \mathfrak{I} \to \mathfrak{Dir}(\kappa)$, the colimit of $J \circ D$ in $\mathfrak{Emb}(\leq)$ is a κ-directed partial order.*

(b) *J preserves κ-directed colimits.*

(c) *Let $\lambda \geq \kappa$ be a regular cardinal. An object $\mathfrak{I} \in \mathfrak{Dir}(\kappa)$ is λ-presentable if, and only if, $|I| < \lambda$.*

(d) *$\mathfrak{Dir}(\kappa)$ is κ-accessible.*

Proof. (a) Let $D : \mathfrak{I} \to \mathfrak{Dir}(\kappa)$ be a κ-directed diagram. Since $\mathfrak{Emb}(\leq)$ has colimits, the diagram $J \circ D$ has a colimit $\mathfrak{A} = \langle A, \leq \rangle \in \mathfrak{Emb}(\leq)$. Let λ be a limiting cocone from $J \circ D$ to $\mathfrak{A}$.

To show that $\mathfrak{A}$ is a partial order, consider elements $a, b, c \in A$. Since D is κ-directed, there exists an index $i \in I$ such that $a, b, c \in \operatorname{rng} \lambda_i$.

For reflexivity, note that λ_i is an embedding and that $D(i)$ is a partial order. Hence, $\lambda_i^{-1}(a) \leq \lambda_i^{-1}(a)$ implies that $a \leq a$.

For antisymmetry, suppose that $a \leq b$ and $b \leq a$. Then we have $\lambda_i^{-1}(a) \leq \lambda_i^{-1}(b)$ and $\lambda_i^{-1}(b) \leq \lambda_i^{-1}(a)$, which implies that $\lambda_i^{-1}(a) = \lambda_i^{-1}(b)$. Hence, $a = b$.

For transitivity, suppose that $a \leq b \leq c$. Then $\lambda_i^{-1}(a) \leq \lambda_i^{-1}(b) \leq \lambda_i^{-1}(c)$, which implies that $\lambda_i^{-1}(a) \leq \lambda_i^{-1}(c)$. Hence, $a \leq c$.

It remains to prove that $\mathfrak{A}$ is κ-directed. Let $X \subseteq A$ be a set of size $|X| < \kappa$. Since D is κ-directed, we can find an index $i \in I$ such that $X \subseteq \operatorname{rng} \lambda_i$. As $D(i)$ is κ-directed, $\lambda_i^{-1}[X]$ has an upper bound $c \in D(i)$. Hence, $\lambda_i(c)$ is an upper bound of X.

(b) Consider a κ-directed diagram $D : \mathfrak{I} \to \mathfrak{Dir}(\kappa)$. Since $\mathfrak{Emb}(\leq)$ has colimits, the diagram $J \circ D$ has a limiting cocone λ to some structure $\mathfrak{A} = \langle A, \leq \rangle$. We have seen in (a) that $\mathfrak{A} \in \mathfrak{Dir}(\kappa)$. Since the inclusion functor is full and faithful, it follows that λ is a cocone from D to $\mathfrak{A}$ in $\mathfrak{Dir}(\kappa)$. Furthermore, note that J reflects colimits by Lemma B3.4.7. Hence, λ is also limiting in $\mathfrak{Dir}(\kappa)$.

To show that J preserves κ-directed colimits, let $\mu \in \mathrm{Cone}(D, \mathfrak{B})$ be a limiting cocone. As both λ and μ are limiting, there exists a (unique) isomorphism $\pi : \mathfrak{B} \to \mathfrak{A}$ such that $\lambda = \pi * \mu$. Since $\lambda = J[\lambda] = J(\pi) * J[\mu]$ is limiting in $\mathfrak{Emb}(\leq)$ and since $J(\pi)$ is an isomorphism, it follows that $J[\mu]$ is also limiting.

(c) ($\Leftarrow$) Let $\mathfrak{I}$ be a κ-directed partial order of size $|I| < \lambda$. According to Proposition 3.2, $\mathfrak{I}$ is λ-presentable in $\mathfrak{Emb}(\leq)$. By (b) and Lemma 3.5, the inclusion functor $\mathfrak{Dir}(\kappa) \to \mathfrak{Emb}(\leq)$ reflects λ-presentability. Hence, $\mathfrak{I}$ is also λ-presentable in $\mathfrak{Dir}(\kappa)$.

($\Rightarrow$) For a partial order $\mathfrak{I}$, we denote by $\mathfrak{I}^\top$ the extension of $\mathfrak{I}$ by a new greatest element $\top$.

Suppose that $\mathfrak{I}$ is λ-presentable. To show that $|I| < \lambda$, let $\mathcal{S}$ be the family of all substructures of $\mathfrak{I}^\top$ of size less than λ, and let $D : \mathcal{S} \to \mathfrak{Emb}(\leq)$ be the canonical diagram. By Proposition B3.3.16, we have $\mathfrak{I}^\top = \varinjlim D$. Let $\mathcal{S}_0 \subseteq \mathcal{S}$ be the subfamily of all substructures of $\mathfrak{I}^\top$ that contain the element $\top$. Note that every such substructure is κ-directed and that $\mathcal{S}_0$ is dense in $\mathcal{S}$. Consequently, the restriction $D \restriction \mathcal{S}_0$ also has the colimit $\mathfrak{I}^\top$ and it factorises as $D \restriction \mathcal{S}_0 = J \circ D_0$ for some $D_0 : \mathcal{S}_0 \to \mathfrak{Dir}(\kappa)$. By Lemma B3.4.7, J reflects colimits. Therefore, $J(\mathfrak{I}^\top) = \mathfrak{I}^\top = \varinjlim (J \circ D_0)$ implies that $\mathfrak{I}^\top = \varinjlim D_0$.

Let μ be a corresponding limiting cocone. As $\mathfrak{I}$ is λ-presentable, the inclusion $h : \mathfrak{I} \to \mathfrak{I}^\top$ factorises as $h = \mu_{\mathfrak{A}} \circ g$, for some $\mathfrak{A} \in \mathcal{S}_0$ and some embedding $g : \mathfrak{I} \to \mathfrak{A}$. Since g is injective, it follows that $|I| = |\mathrm{rng}\, g| \leq |A| < \lambda$.

(d) To show that $\mathfrak{Dir}(\kappa)$ has κ-directed colimits, let $D : \mathfrak{I} \to \mathfrak{Dir}(\kappa)$ be a κ-directed diagram. By (a), the colimit $\mathfrak{A}$ of $J \circ D$ in $\mathfrak{Emb}(\leq)$ belongs to $\mathfrak{Dir}(\kappa)$. By Lemma B3.4.7, the inclusion functor J reflects colimits. Consequently, $\mathfrak{A}$ is also the colimit of D in $\mathfrak{Dir}(\kappa)$.

Furthermore, note that (c) implies that, up to isomorphism, there exist only a set of κ-presentable objects in $\mathfrak{Dir}(\kappa)$.

Hence, it remains to show that every object of $\mathfrak{Dir}(\kappa)$ can be written as a κ-directed diagram of κ-presentable objects. Given $\mathfrak{I} \in \mathfrak{Dir}(\kappa)$, let $\mathcal{S}$ be the family of all substructures of $\mathfrak{I}$ of size less than κ and let $D : \mathcal{S} \to \mathfrak{Emb}(\leq)$ be the canonical diagram. By Proposition B3.3.16, we have $\mathfrak{I} = \varinjlim D$. Let $\mathcal{S}_0 \subseteq \mathcal{S}$ be the subfamily of all substructures of $\mathfrak{I}$ that have a greatest element. We claim that $\mathcal{S}_0$ is dense in $\mathcal{S}$. Let $\mathfrak{A} \in \mathcal{S}$. Then $|A| < \kappa$ and, since $\mathfrak{I}$ is κ-directed, the set $A \subseteq I$ has an upper bound $b \in I$. Consequently, $\mathfrak{I}|_{A\cup\{b\}}$ is an element of $\mathcal{S}_0$ containing $\mathfrak{A}$.

Note that every substructure in $\mathcal{S}_0$ is κ-directed and that $\mathcal{S}_0$ is dense in $\mathcal{S}$. It follows that the restriction $D \upharpoonright \mathcal{S}_0$ also has the colimit $\mathfrak{I}$ and that $D \upharpoonright \mathcal{S}_0$ factorises as $D \upharpoonright \mathcal{S}_0 = J \circ D_0$ for some $D_0 : \mathcal{S}_0 \to \mathfrak{Dir}(\kappa)$. By Lemma B3.4.7, J reflects colimits. Therefore, $J(\mathfrak{I}) = \mathfrak{I} = \varinjlim (J \circ D_0)$ implies that $\mathfrak{I} = \varinjlim D_0$, as desired. □

A further important example of a κ-accessible category is the inductive completion of a category.

Lemma 4.6. *Let $\mathcal{C}$ be a category, κ a regular cardinal, and let $I : \mathcal{C} \to \mathrm{Ind}^\infty_\kappa(\mathcal{C})$ be the inclusion functor. In $\mathrm{Ind}^\infty_\kappa(\mathcal{C})$ every object of the form $I(\mathfrak{a})$ is κ-presentable.*

Proof. To keep notation simple, we will not distinguish below between a morphism $f : \mathfrak{a} \to \mathfrak{b}$ of $\mathcal{C}$ and the link $t = (t_i)_{i\in[1]}$ whose only component is $t_0 = f$.

Let $D : \mathcal{I} \to \mathrm{Ind}^\infty_\kappa(\mathcal{C})$ be a κ-directed diagram with union $U : \mathcal{J} \to \mathcal{C}$. By Proposition 1.13, the family $\mu = (\mu_\mathfrak{i})_{\mathfrak{i}\in\mathcal{I}}$ with $\mu_\mathfrak{i} = [\mathrm{in}_{D(\mathfrak{i})}]^{\Lsh}_U$ is a limiting cocone from D to U.

To show that $I(\mathfrak{a})$ is κ-presentable, let $[f]^{\Lsh}_U : I(\mathfrak{a}) \to U$ be a morphism. We have to show that $[f]^{\Lsh}_U$ factorises essentially uniquely through the cocone μ. Suppose that $f : \mathfrak{a} \to U(\langle\mathfrak{i},\mathfrak{k}\rangle)$. Then we can regard f as a link from $I(\mathfrak{a})$ to $D(\mathfrak{i})$. Let $[f]^{\Lsh}_{D(\mathfrak{i})} : I[\mathfrak{a}] \to D(\mathfrak{i})$ be the corresponding morphism of $\mathrm{Ind}^\infty_\kappa(\mathcal{C})$. Then

$$\mu_\mathfrak{i} \circ [f]^{\Lsh}_{D(\mathfrak{i})} = [\mathrm{in}_{D(\mathfrak{i})}]^{\Lsh}_U \circ [f]^{\Lsh}_{D(\mathfrak{i})} = [\mathrm{id}_{D(\mathfrak{i})(\mathfrak{k})} \circ f]^{\Lsh}_U = [f]^{\Lsh}_U .$$

We claim that this factorisation of $[f]^{⩕}_U$ is essentially unique.

Let $[f]^{⩕}_U = \mu_{\mathfrak{j}} \circ [g]^{⩕}_{D(\mathfrak{j})}$ be a second factorisation where $[g]^{⩕}_{D(\mathfrak{j})} : I(\mathfrak{a}) \to D(\mathfrak{j})$. Then $g : \mathfrak{a} \to D(\mathfrak{j})(\mathfrak{l})$, for some index $\mathfrak{l}$, and, as above, it follows that

$$[f]^{⩕}_U = \mu_{\mathfrak{j}} \circ [g]^{⩕}_{D(\mathfrak{j})} = [\mathrm{id}_{D(\mathfrak{j})(\mathfrak{l})} \circ g]^{⩕}_U = [g]^{⩕}_U .$$

Hence, $f \mathrel{⩕}_U g$ and there are morphisms

$$h : \langle \mathfrak{i}, \mathfrak{k} \rangle \to \langle \mathfrak{m}, \mathfrak{n} \rangle \quad \text{and} \quad h' : \langle \mathfrak{j}, \mathfrak{l} \rangle \to \langle \mathfrak{m}, \mathfrak{n} \rangle$$

of $\mathcal{J}$ such that

$$U(h) \circ f = U(h') \circ g .$$

By definition of the union, we can express h and h' as finite compositions

$$h = h_{u-1} \circ \cdots \circ h_0 \quad \text{and} \quad h' = h'_{v-1} \circ \cdots \circ h'_0$$

of morphisms of the form $D(\mathfrak{x})(\varphi)$ and $t(\mathfrak{x}, \mathfrak{y})_{\mathfrak{r}}$, for indices $\mathfrak{x} \in \mathcal{I}$, morphisms φ in the index category of $D(\mathfrak{x})$, and links $t(\mathfrak{x}, \mathfrak{y})$ such that $D(\mathfrak{x}, \mathfrak{y}) = [t(\mathfrak{x}, \mathfrak{y})]^{⩕}_{D(\mathfrak{y})}$. By induction on u and v it follows that

$$[h_{u-1} \circ \cdots \circ h_0 \circ f]^{⩕}_{D(\mathfrak{m})} \mathrel{⩕}_D [f]^{⩕}_{D(\mathfrak{i})}$$
$$\text{and} \quad [h'_{v-1} \circ \cdots \circ h'_0 \circ g]^{⩕}_{D(\mathfrak{m})} \mathrel{⩕}_D [g]^{⩕}_{D(\mathfrak{j})} .$$

Hence, $h \circ f = h' \circ g$ implies that

$$[f]^{⩕}_{D(\mathfrak{i})} \mathrel{⩕}_D [h_{u-1} \circ \cdots \circ h_0 \circ f]^{⩕}_{D(\mathfrak{m})} = [h'_{v-1} \circ \cdots \circ h'_0 \circ g]^{⩕}_{D(\mathfrak{m})} \mathrel{⩕}_D [g]^{⩕}_{D(\mathfrak{j})} .$$

□

Proposition 4.7. $\mathrm{Ind}^{\infty}_{\kappa}(\mathcal{C})$ *is κ-accessible, for every small category $\mathcal{C}$.*

Proof. Let $I : \mathcal{C} \to \mathrm{Ind}^{\infty}_{\kappa}(\mathcal{C})$ be the inclusion functor. We have seen in Theorem 1.15 that the category $\mathrm{Ind}^{\infty}_{\kappa}(\mathcal{C})$ has κ-directed colimits and that every object of $\mathrm{Ind}^{\infty}_{\kappa}(\mathcal{C})$ can be written as a κ-filtered diagram of objects in rng I. Hence, it follows from Lemma 4.6 that every object of $\mathrm{Ind}^{\infty}_{\kappa}(\mathcal{C})$ is a κ-filtered colimit of κ-presentable objects.

Consequently, it remains to prove that, up to isomorphism, the κ-presentable objects of $\mathrm{Ind}^{\infty}_{\kappa}(\mathcal{C})$ form a set. By Theorem 3.8, every κ-presentable object can be written as a κ-filtered colimit of size less than κ where all objects are in rng $I \cong \mathcal{C}$. Consequently, an object is κ-presentable if, and only if, it belongs to $\mathrm{Ind}^{\kappa}_{\kappa}(\mathcal{C})$. Since $\mathcal{C}$ is small, there exist, up to isomorphism, only a set of diagrams $D : \mathcal{I} \to \mathcal{C}$ of size less than κ. Therefore, $\mathrm{Ind}^{\kappa}_{\kappa}(\mathcal{C})$ is small (up to isomorphism). □

In fact, all κ-accessible categories are of this form.

Theorem 4.8. *A category $\mathcal{C}$ is κ-accessible if, and only if, it is equivalent to a category of the form* $\mathrm{Ind}^{\infty}_{\kappa}(\mathcal{C}_{\mathrm{o}})$, *for some small category $\mathcal{C}_{\mathrm{o}}$.*

Proof. ($\Leftarrow$) We have seen in Proposition 4.7 that $\mathrm{Ind}^{\infty}_{\kappa}(\mathcal{C}_{\mathrm{o}})$ is κ-accessible. Hence, all categories $\mathcal{C}$ equivalent to $\mathrm{Ind}^{\infty}_{\kappa}(\mathcal{C}_{\mathrm{o}})$ are κ-accessible.

($\Rightarrow$) Suppose that $\mathcal{C}$ is κ-accessible, let $\mathcal{C}_1$ be the full subcategory of all κ-presentable objects of $\mathcal{C}$, and let $\mathcal{C}_{\mathrm{o}}$ be a skeleton of $\mathcal{C}_1$. We claim that $\mathcal{C}$ is equivalent to $\mathrm{Ind}^{\infty}_{\kappa}(\mathcal{C}_{\mathrm{o}})$.

Let $Q_{\mathrm{o}} : \mathrm{Ind}^{\infty}_{\kappa}(\mathcal{C}_{\mathrm{o}}) \to \mathcal{C}$ be the restriction of the canonical projection $Q : \mathrm{Ind}^{\infty}_{\kappa}(\mathcal{C}) \to \mathcal{C}$ to $\mathrm{Ind}^{\infty}_{\kappa}(\mathcal{C}_{\mathrm{o}})$. We claim that Q_{o} is the desired equivalence. By Theorem B1.3.14, it is sufficient to prove that Q_{o} is full and faithful and that every object of $\mathcal{C}$ is isomorphic to some object in rng $Q_{\mathrm{o}}^{\mathrm{obj}}$.

Let $D : \mathcal{I} \to \mathcal{C}_{\mathrm{o}}$ and $E : \mathcal{J} \to \mathcal{C}_{\mathrm{o}}$ be objects of $\mathrm{Ind}^{\infty}_{\kappa}(\mathcal{C}_{\mathrm{o}})$ and let λ^D and λ^E be the limiting cocones used to define $Q_{\mathrm{o}}(D)$ and $Q_{\mathrm{o}}(E)$.

To show that Q_{o} is faithful, let $[f]^{\mathbin{\wedge\!\!\!\wedge}}_{E}, [g]^{\mathbin{\wedge\!\!\!\wedge}}_{E} : D \to E$ be morphisms of $\mathrm{Ind}^{\infty}_{\kappa}(\mathcal{C}_{\mathrm{o}})$ with $Q_{\mathrm{o}}([f]^{\mathbin{\wedge\!\!\!\wedge}}_{E}) = Q_{\mathrm{o}}([g]^{\mathbin{\wedge\!\!\!\wedge}}_{E})$. Then

$$\lambda^E * f = Q_{\mathrm{o}}([f]^{\mathbin{\wedge\!\!\!\wedge}}_{E}) * \lambda^D = Q_{\mathrm{o}}([g]^{\mathbin{\wedge\!\!\!\wedge}}_{E}) * \lambda^D = \lambda^E * g\,.$$

By Lemma 3.6, this implies that $f \mathbin{\wedge\!\!\!\wedge}_E g$. Hence, $[f]^{\mathbin{\wedge\!\!\!\wedge}}_{E} = [g]^{\mathbin{\wedge\!\!\!\wedge}}_{E}$.

To prove that Q_{o} is full, let $f : Q_{\mathrm{o}}(D) \to Q_{\mathrm{o}}(E)$ be a morphism of $\mathcal{C}$. By Lemma 3.6 (b), there exists a link t from D to E such that

$$\lambda^E * t = f * \lambda^D\,.$$

By definition of Q_0^{mor}, this implies that $Q_0([t]_E^{\approx}) = f$.

Hence, it remains to prove that every object $\mathfrak{a} \in \mathcal{C}$ is isomorphic to some object in rng Q_0^{obj}. Let $D : \mathfrak{I} \to \mathcal{C}$ be a κ-directed diagram with colimit $\mathfrak{a}$ where every object $D(i)$ belongs to $\mathcal{C}_1$. For every index $i \in I$, let $E(i)$ be the unique object of $\mathcal{C}_0$ isomorphic to $D(i)$. This defines the object part of a functor $E : \mathfrak{I} \to \mathcal{C}_0$. To define the morphism part, we fix isomorphisms $\eta_i : D(i) \cong E(i)$ and we set

$$E(i,j) := \eta_j \circ D(i,j) \circ \eta_i^{-1}, \quad \text{for } i \leq j.$$

Then E is a κ-directed diagram in $\mathrm{Ind}_\kappa^\infty(\mathcal{C}_0)$ and $\eta := (\eta_i)_{i \in I}$ is a natural isomorphism $\eta : D \cong E$. Consequently, it follows by Lemma B3.4.3 that

$$Q_0(E) = \varinjlim E \cong \varinjlim D = \mathfrak{a},$$

as desired. □

Finally, let us show that in general it is not true that a κ-accessible category is also λ-accessible for larger cardinals λ. Studying this question, we again meet the relation $\trianglelefteq$.

Theorem 4.9. *Let $\kappa \leq \lambda$ be regular cardinals. The following statements are equivalent:*

(1) $\kappa \trianglelefteq \lambda$

(2) *Every κ-accessible category is λ-accessible.*

(3) *Let $\mathcal{C}$ be a category with κ-directed colimits. For each κ-directed diagram $D : \mathfrak{I} \to \mathcal{C}$ of κ-presentable objects, there exists a λ-directed diagram $D^+ : \mathfrak{I}^+ \to \mathcal{C}$ of λ-presentable objects with the same colimit.*

(4) *For every set X of size $|X| < \lambda$, we can write the partial order $\langle \wp_\kappa(X), \subseteq \rangle$ as the colimit of a λ-directed diagram $D : \mathfrak{I} \to \mathfrak{Dir}(\kappa)$ of partial orders of size $|D(i)| < \lambda$.*

Proof. (1) ⇒ (3) Let $D : \mathfrak{I} \to \mathcal{C}$ be a κ-directed diagram of κ-presentable objects. By (1) and Proposition 2.11, there exists a λ-directed diagram

$D^+ : \mathfrak{I}^+ \to \mathcal{C}$ with the same colimit as D where every object $D^+(i)$ is of the form $\varinjlim (D \upharpoonright S)$, for some κ-directed subset $S \subseteq I$ of size $|S| < \lambda$. By Proposition 3.7, it follows that each $D^+(i)$ is λ-presentable.

(3) ⇒ (2) Let $\mathcal{C}$ be a κ-accessible category. Since every λ-directed diagram is also κ-directed, it follows that $\mathcal{C}$ has λ-directed colimits.

We claim that every $\mathfrak{a} \in \mathcal{C}$ is a λ-directed colimit of λ-presentable objects. As $\mathcal{C}$ is κ-accessible, there exists a κ-directed diagram $D : \mathfrak{I} \to \mathcal{C}$ of κ-presentable objects with colimit $\mathfrak{a}$. By (3), it follows that $\mathfrak{a}$ is the colimit of a λ-directed diagram D^+ of λ-presentable objects.

It remains to prove that the λ-presentable objects form a set. By Theorem 4.3, we can write every λ-presentable object as a κ-directed diagram D of size less than λ such that each $D(\mathfrak{i})$ is κ-presentable. Since, up to isomorphism, there exists only a set of κ-presentable objects, it follows that, up to isomorphism, there also exists only a set of such diagrams.

(2) ⇒ (4) Let X be a set of size less than λ. Since κ is regular, the partial order $\langle \wp_\kappa(X), \subseteq \rangle$ is κ-directed. Hence, it is an object of the category $\mathfrak{Dir}(\kappa)$. We have shown in Proposition 4.5 that $\mathfrak{Dir}(\kappa)$ is κ-accessible. By (2), it is also λ-accessible. Consequently, we can write $\wp_\kappa(X)$ as the colimit of a λ-directed diagram $D : \mathfrak{I} \to \mathfrak{Dir}(\kappa)$ of λ-presentable objects. By Proposition 4.5 (c), it follows that every $D(i)$ has size less than λ.

(4) ⇒ (1) Let X be a set of size less than λ. We have to find a dense set $H \subseteq \wp_\kappa(X)$ of size $|H| < \lambda$. By (4), there exists a λ-directed diagram $D : \mathfrak{I} \to \mathfrak{Dir}(\kappa)$ of partial orders of size less than λ with $\varinjlim D = \wp_\kappa(X)$. Let μ be the corresponding limiting cocone. For each element $x \in X$, we select an index $i(x) \in I$ such that $\{x\} \in \operatorname{rng} \mu_{i(x)}$. Since $\mathfrak{I}$ is λ-directed, there exists an index $k \in I$ with $k \geq i(x)$, for all $x \in X$. This implies that $\{ \{x\} \mid x \in X \} \subseteq \operatorname{rng} \mu_k$.

We claim that the range $H := \operatorname{rng} \mu_k$ is the desired dense set. Since $|H| = |D(k)| < \lambda$, it remains to show that H is dense. Let $Y \in \wp_\kappa(X)$. As $D(k)$ is κ-directed, it contains an upper bound c of the set $\{ \mu_k^{-1}(\{y\}) \mid y \in Y \}$. Consequently, $\mu_k(c) \in H$ is an upper bound of $\{ \{y\} \mid y \in Y \}$. This implies that $Y \subseteq \mu_k(c)$. □

Substructures

We have shown in Proposition B3.3.16, that every Σ-structure can be written as a κ-directed colimit of its κ-generated substructures. This statement can be generalised to arbitrary κ-accessible categories. We start by introducing a notion of substructure for accessible categories.

Definition 4.10. Let $\mathcal{C}$ be a category, $\mathcal{K} \subseteq \mathcal{C}^{\mathrm{obj}}$ a class of objects, and $\mathfrak{a} \in \mathcal{C}$.

(a) We define the arrow category

$$\mathfrak{Sub}_{\mathcal{K}}(\mathfrak{a}) := (\mathcal{K} \downarrow \mathfrak{a}) ,$$

where we have written $\mathcal{K}$ for the inclusion functor $\mathcal{K} \to \mathcal{C}$.

For the class $\mathcal{K}$ of all κ-presentable objects, we also write $\mathfrak{Sub}_{\kappa}(\mathfrak{a})$ instead of $\mathfrak{Sub}_{\mathcal{K}}(\mathfrak{a})$.

(b) The *canonical diagram* $D : \mathfrak{Sub}_{\mathcal{K}}(\mathfrak{a}) \to \mathcal{C}$ of $\mathfrak{a}$ over $\mathcal{K}$ is defined by

$$D(f) := \mathfrak{c} , \quad \text{for objects } f : \mathfrak{c} \to \mathfrak{a} ,$$
$$\text{and} \quad D(\varphi) := \varphi , \quad \text{for morphisms } \varphi : f \to f' .$$

Before generalising Proposition B3.3.16 we prove a technical lemma.

Lemma 4.11. *Let $\mathcal{C}$ be a category, $D : \mathfrak{Sub}_{\kappa}(\mathfrak{a}) \to \mathcal{C}$ the canonical diagram of $\mathfrak{a} \in \mathcal{C}$, and $E : \mathcal{I} \to \mathcal{C}$ a diagram with colimit $\mathfrak{a}$ such that every $E(\mathfrak{i})$ is κ-presentable.*

(a) *E factorises as $E = D \circ F$, for a suitable functor $F : \mathcal{I} \to \mathfrak{Sub}_{\kappa}(\mathfrak{a})$.*

(b) *If $\mathcal{I}$ is κ-filtered, we can choose F to be dense.*

Proof. Let λ be a limiting cocone from E to $\mathfrak{a}$. We define

$$F(\mathfrak{i}) := \lambda_{\mathfrak{i}} , \quad \text{for } \mathfrak{i} \in \mathcal{I}^{\mathrm{obj}},$$
$$F(f) := E(f) , \quad \text{for } f \in \mathcal{I}^{\mathrm{mor}}.$$

To see that F is indeed a functor $\mathcal{I} \to \mathfrak{Sub}_{\kappa}(\mathfrak{a})$, note that, for a morphism $f : \mathfrak{i} \to \mathfrak{j}$ of $\mathcal{I}$, $\lambda_{\mathfrak{i}} = \lambda_{\mathfrak{j}} \circ E(f)$ implies that $F(f) \in \mathfrak{Sub}_{\kappa}(\mathfrak{a})(\lambda_{\mathfrak{i}}, \lambda_{\mathfrak{j}})$.

(a) We have

$$\begin{aligned} (D \circ F)(\mathfrak{i}) &= D(\lambda_{\mathfrak{i}}) = E(\mathfrak{i}), && \text{for } \mathfrak{i} \in \mathcal{I}^{\mathrm{obj}}, \\ (D \circ F)(f) &= D(E(f)) = E(f), && \text{for } f \in \mathcal{I}^{\mathrm{mor}}. \end{aligned}$$

(b) (D1) Consider $g \in \mathfrak{Sub}_\kappa(\mathfrak{a})$. Since g factorises essentially uniquely through λ, there are $\mathfrak{i} \in \mathcal{I}$ and a morphism g_0 such that $g = \lambda_{\mathfrak{i}} \circ g_0$. Since $F(\mathfrak{i}) = \lambda_{\mathfrak{i}}$, it follows that $g_0 : g \to F(\mathfrak{i})$ is a morphism in $\mathfrak{Sub}_\kappa(\mathfrak{a})$.

(D2) Let $f : g \to F(\mathfrak{i})$ and $f' : g \to F(\mathfrak{i}')$ be morphisms of $\mathfrak{Sub}_\kappa(\mathfrak{a})$. Then

$$\lambda_{\mathfrak{i}} \circ f = F(\mathfrak{i}) \circ f = g = F(\mathfrak{i}') \circ f' = \lambda_{\mathfrak{i}'} \circ f'.$$

Consequently, $\lambda_{\mathfrak{i}} \circ f$ and $\lambda_{\mathfrak{i}'} \circ f$ are two factorisations of g through λ. As E is κ-filtered and the domain of g is κ-presentable, it follows by essential uniqueness and Corollary 1.3 that there are morphisms $h : \mathfrak{i} \to \mathfrak{k}$ and $h' : \mathfrak{i}' \to \mathfrak{k}$ such that

$$E(h) \circ f = E(h') \circ f'.$$

Consequently,

$$F(h) \circ f = F(h') \circ f',$$

which implies that $f \approx_F f'$. □

Proposition 4.12. *Let $\mathcal{C}$ be a κ-accessible category and $\mathfrak{a} \in \mathcal{C}$ an object. The canonical diagram $D : \mathfrak{Sub}_\kappa(\mathfrak{a}) \to \mathcal{C}$ of $\mathfrak{a}$ is κ-filtered and $\varinjlim D = \mathfrak{a}$.*

Proof. Fix a κ-directed diagram $E : \mathfrak{I} \to \mathcal{C}$ of κ-presentable objects with colimit $\mathfrak{a}$ and let λ be the corresponding limiting cocone. To show that $\mathfrak{Sub}_\kappa(\mathfrak{a})$ is κ-filtered, we have to check two conditions.

(F1) Let $X \subseteq \mathfrak{Sub}_\kappa(\mathfrak{a})^{\mathrm{obj}}$ be a set of size $|X| < \kappa$. Every $g : \mathfrak{c}_g \to \mathfrak{a}$ in X factorises essentially uniquely through λ as $g = \lambda_{k_g} \circ g_0$, for suitable $k_g \in I$ and $g_0 : \mathfrak{c}_g \to E(k_g)$. Since $\mathfrak{I}$ is κ-directed, there exists an upper bound $l \in I$ of $\{ k_g \mid g \in X \}$. Consequently, $\lambda_l : E(l) \to \mathfrak{a}$ is an object of $\mathfrak{Sub}_\kappa(\mathfrak{a})$ and

$$E(k_g, l) \circ g_0 : g \to \lambda_l\,, \quad \text{for } g \in X\,,$$

is the desired family of morphisms of $\mathfrak{Sub}_\kappa(\mathfrak{a})$.

(F2) Let $X \subseteq \mathfrak{Sub}_\kappa(\mathfrak{a})(g, g')$ be a set of size $|X| < \kappa$. There are essentially unique factorisations

$$g = \lambda_i \circ g_0 \quad \text{and} \quad g' = \lambda_j \circ g'_0\,, \quad \text{for suitable } i, j \in I\,.$$

For every $f \in X$,

$$\lambda_j \circ (g'_0 \circ f) = g' \circ f = g\,,$$

is another factorisation of g. Consequently, $g'_0 \circ f \approx_E g_0$ and, by Corollary 1.3, we can find an index $k_f \geq i, j$ such that

$$E(j, k_f) \circ g'_0 \circ f = E(i, k_f) \circ g_0\,.$$

Let l be an upper bound of $\{ k_f \mid f \in X \}$. Then

$$E(j, l) \circ g'_0 \circ f = E(i, l) \circ g_0 = E(j, l) \circ g'_0 \circ f'\,,$$

for all $f, f' \in X$. Since $\lambda_l : E(l) \to \mathfrak{a}$ is an object of $\mathfrak{Sub}_\kappa(\mathfrak{a})$ and $E(j, l) \circ g'_0 : g' \to \lambda_l$ is a morphism, the claim follows.

It remains to prove that D has the colimit $\mathfrak{a}$. Let $F : \mathcal{I} \to \mathfrak{Sub}_\kappa(\mathfrak{a})$ be the dense functor from Lemma 4.11 with $E = D \circ F$. Then

$$\varinjlim D = \varinjlim (D \circ F) = \varinjlim E = \mathfrak{a}\,. \qquad \square$$

B5. Topology

1. Open and closed sets

Definition 1.1. A *topology* on a set X is a system $\mathcal{C} \subseteq \wp(X)$ of subsets of X that satisfies the following conditions:

- $\varnothing, X \in \mathcal{C}$
- If $Z \subseteq \mathcal{C}$ then $\bigcap Z \in \mathcal{C}$.
- If $C_0, C_1 \in \mathcal{C}$ then $C_0 \cup C_1 \in \mathcal{C}$.

A *topological space* is a pair $\mathfrak{X} = \langle X, \mathcal{C} \rangle$ consisting of a set X and a topology $\mathcal{C}$ on X. The elements of $\mathcal{C}$ are called *closed sets.* A set O is *open* if its complement $X \smallsetminus O$ is closed. Sets that are both closed and open are called *clopen.* A set U is a *neighbourhood* of an element $x \in X$ if there exists an open set O with $x \in O \subseteq U$. The elements of a topological space X are usually called *points.*

Example. (a) In the usual topology $\langle \mathbb{R}, \mathcal{C} \rangle$ of the real numbers a subset $A \subseteq \mathbb{R}$ is open if and only if, for every $a \in A$, there exists an open interval $(c, d) \subseteq A$ with $a \in (c, d)$. Correspondingly, a set $A \subseteq \mathbb{R}$ is closed if it contains all elements $a \in \mathbb{R}$ such that, for every open interval (c, d) with $a \in (c, d)$, there exists an element $b \in (c, d) \cap A$. The only clopen sets are $\varnothing$ and $\mathbb{R}$.

(b) Consider the space $\mathbb{R}^n$. We denote the usual Euklidean norm of a tuple $\bar{a} \in \mathbb{R}^n$ by

$$\|\bar{a}\| := \sqrt{a_0^2 + \dots + a_{n-1}^2}\,,$$

and the *ε-ball* around $\bar{a}$ by

$$B_\varepsilon(\bar{a}) := \{\, \bar{b} \in \mathbb{R}^n \mid \|\bar{b} - \bar{a}\| < \varepsilon \,\}.$$

A set $A \subseteq \mathbb{R}^n$ is open if and only if, for every $\bar{a} \in A$, there is some $\varepsilon > 0$ such that $B_\varepsilon(\bar{a}) \subseteq A$. The set A is closed if, whenever $\bar{a} \in \mathbb{R}^n$ is a tuple such that $B_\varepsilon(\bar{a}) \cap A \neq \emptyset$, for all $\varepsilon > 0$, then we have $\bar{a} \in A$.

(c) Let X be an arbitrary set. The *trivial topology* of X is given by the set $\mathcal{C} = \{\emptyset, X\}$ where only $\emptyset$ and X are closed.

(d) The *discrete topology* of a set X is its power set $\mathcal{C} = \wp(X)$ where every set is clopen.

(e) We can define a topology on any set X by

$$\mathcal{C} := \{\, C \subseteq X \mid C \text{ is finite} \,\}.$$

(f) Let $\mathfrak{K}$ be a field and $n < \omega$. For a set $I \subseteq K[x_0, \dots, x_{n-1}]$ of polynomials over $\mathfrak{K}$, define

$$Z(I) := \{\, \bar{a} \in K^n \mid p(\bar{a}) = 0 \text{ for all } p \in I \,\}.$$

We can equip K^n with the *Zariski topology*

$$\mathcal{Z} := \{\, Z(I) \mid I \subseteq K[\bar{x}] \,\}.$$

Let us prove that $\mathcal{Z}$ is indeed a topology. Clearly,

$$\emptyset = Z(\{1\}) \in \mathcal{Z} \quad \text{and} \quad K^n = Z(\{0\}) \in \mathcal{Z}.$$

Let $X \subseteq \mathcal{Z}$ and set $\mathcal{I} := \{\, I \mid Z(I) \in X \,\}$. Then we have

$$\bigcap X = \bigcap \{\, Z(I) \mid I \in \mathcal{I} \,\} = Z(\bigcup \mathcal{I}) \in \mathcal{Z}.$$

Finally, suppose that $Z(I_0), Z(I_1) \in \mathcal{Z}$. Then

$$Z(I_0) \cup Z(I_1) = Z(J), \quad \text{where} \quad J := \{\, pq \mid p \in I_0,\ q \in I_1 \,\}.$$

Note that, for $n = 1$, $\mathcal{Z}$ consists of all finite subsets of K. If $K = \mathbb{R}$ and $\mathcal{C}$ is the usual topology on $\mathbb{R}$ then we have $\mathcal{Z} \subset \mathcal{C}$. An example of a $\mathcal{C}$-closed set that is not $\mathcal{Z}$-closed is $[0,1]^n$.

Remark. (a) Note that the system $\mathcal{O}$ of open sets satisfies:

- $\emptyset, X \in \mathcal{O}$
- If $Z \subseteq \mathcal{O}$ then $\bigcup Z \in \mathcal{O}$.
- If $O_0, O_1 \in \mathcal{O}$ then $O_0 \cap O_1 \in \mathcal{O}$.

Conversely, given any system $\mathcal{O}$ with these properties we can define a topology by

$$\mathcal{C} := \{ X \setminus O \mid O \in \mathcal{O} \} .$$

(b) The family of clopen sets of a topological space $\mathfrak{X}$ forms a boolean algebra.

Lemma 1.2. *Let $\mathfrak{X}$ be a topological space. A set $A \subseteq X$ is open if and only if it is a neighbourhood of all of its elements.*

Proof. Clearly, if A is open and $x \in A$ then we have $x \in A \subseteq A$ and A is a neighbourhood of x. Conversely, suppose that, for every $x \in A$, there is an open set O_x with $x \in O_x \subseteq A$. Then $A = \bigcup_{x \in A} O_x$ is open. □

Remark. The family of all neighbourhoods of a point $x \in X$ forms a filter in the power-set lattice $\wp(X)$.

Note that every topological space is a closure space. Hence, we can use Lemma A2.4.8 to assign to each topology a corresponding closure operator.

Definition 1.3. Let $\mathfrak{X} = \langle X, \mathcal{C} \rangle$ be a topological space.

(a) The *topological closure* of a set $A \subseteq X$ is

$$\mathrm{cl}(A) := \bigcap \{ C \in \mathcal{C} \mid A \subseteq C \} .$$

(b) The *interior* of A is the set

$$\mathrm{int}(A) := \bigcup \{ O \mid O \subseteq A \text{ is open} \} .$$

(c) The *boundary* of A is the set

$$\partial A := \mathrm{cl}(A) \setminus \mathrm{int}(A) .$$

Example. (a) Consider the space $\mathbb{R}$. We have $\mathrm{cl}(\mathbb{Q}) = \mathbb{R}$, $\mathrm{int}(\mathbb{Q}) = \varnothing$, and $\partial\mathbb{Q} = \mathbb{R}$.

(b) The interior of a closed interval $[a, b]$ is the corresponding open interval (a, b). Its boundary is $\{a, b\}$.

Exercise 1.1. Prove that

$$\mathrm{int}(A) = A \smallsetminus \mathrm{cl}(X \smallsetminus A) \quad \text{and} \quad \partial A = \mathrm{cl}(A) \cap \mathrm{cl}(X \smallsetminus A)\,.$$

Lemma 1.4. *Let X be a set.*

(a) *If $\mathcal{C}$ is a topology on X, the corresponding operation* cl *forms a topological closure operator on X.*

(b) *Conversely, if c is a topological closure operator on X, then* fix c *is a topology on X.*

As seen in the examples above, it can be quite cumbersome to describe a topology by defining when a set is closed. Instead, it is usually easier to define only some especially simple closed sets. Note that the intersection of a family of topologies is again a topology. Hence, the collection of all topologies on a set X form a complete partial order and we can assign to each family $\mathcal{B} \subseteq \wp(X)$ the least topology containing $\mathcal{B}$.

Definition 1.5. Let $\mathfrak{X} = \langle X, \mathcal{C} \rangle$ be a closure space.

(a) A *closed base* of $\mathcal{C}$ is a system $\mathcal{B} \subseteq \wp(X)$ such that

$$\mathcal{C} = \{ \textstyle\bigcap Z \mid Z \subseteq \mathcal{B} \}\,.$$

(By convention, we set $\bigcap \varnothing := X$.)

(b) An *open base* of $\mathcal{C}$ is a system $\mathcal{B} \subseteq \wp(X)$ such that

$$\mathcal{C} = \{ X \smallsetminus \textstyle\bigcup Z \mid Z \subseteq \mathcal{B} \}\,.$$

(c) A *closed subbase* of $\mathcal{C}$ is a system $\mathcal{B} \subseteq \wp(X)$ such that the set

$$\{ B_0 \cup \cdots \cup B_{n-1} \mid n < \omega,\ B_i \in \mathcal{B} \}$$

forms a closed base of $\mathcal{C}$.

(d) An *open subbase* of $\mathcal{C}$ is a system $\mathcal{B} \subseteq \wp(X)$ such that the set

$$\{ B_0 \cap \dots \cap B_{n-1} \mid n < \omega,\ B_i \in \mathcal{B} \}$$

forms an open base of $\mathcal{C}$.

(e) If $\mathcal{B}$ is a base or subbase of $\mathcal{C}$ then we say that $\mathcal{B}$ *induces* the topology $\mathcal{C}$.

Every family $\mathcal{B} \subseteq \wp(X)$ is a closed base for the closure space $\langle X, \mathcal{C} \rangle$ where

$$\mathcal{C} := \{ \bigcap Z \mid Z \subseteq \mathcal{B} \}.$$

In the following lemma we characterise those families $\mathcal{B}$ where resulting closure space is topological.

Lemma 1.6. *Let X be a set and $\mathcal{B} \subseteq \wp(X)$.*

(a) *$\mathcal{B}$ forms a closed base of some topology $\mathcal{C}$ on X if and only if it satisfies the following conditions:*
   - $\bigcap \mathcal{B} = \varnothing$.
   - *For all $C_0, C_1 \in \mathcal{B}$, there exists a set $Z \subseteq \mathcal{B}$ such that $C_0 \cup C_1 = \bigcap Z$.*

(b) *$\mathcal{B}$ forms an open base of some topology $\mathcal{C}$ on X if and only if it satisfies the following conditions:*
   - $\bigcup \mathcal{B} = X$.
   - *For all $O_0, O_1 \in \mathcal{B}$, there is a set $Z \subseteq \mathcal{B}$ such that $O_0 \cap O_1 = \bigcup Z$.*

Remark. (a) The set of all open intervals forms an open base for the topology of $\mathbb{R}$. An open subbase is given by the set of all intervals of the form $\downarrow a$ and $\uparrow a$, for $a \in \mathbb{R}$. Similarly, the set of all intervals of the form $\Downarrow a$ and $\Uparrow a$ is a closed subbase for this topology.

(b) The usual topology of $\mathbb{R}^n$ has an open base consisting of all balls $B_\varepsilon(\bar{a})$ with $\bar{a} \in \mathbb{R}^n$ and $\varepsilon > 0$.

Definition 1.7. Let $\mathfrak{X} = \langle X, \mathcal{C} \rangle$ be a closure space and $Y \subseteq X$. The *closure subspace* of $\mathfrak{X}$ induced by Y is the closure space

$$\mathfrak{X}|_Y := \langle X, \mathcal{C}|_Y \rangle \quad \text{where} \quad \mathcal{C}|_Y := \{ C \cap Y \mid C \in \mathcal{C} \} .$$

$\mathcal{C}|_Y$ is called the system of closed sets on Y *induced* by $\mathcal{C}$.

Lemma 1.8. *If $\mathfrak{X}$ is a topological space then so is $\mathfrak{X}|_Y$, for every $Y \subseteq X$.*

Example. Let $X = \mathbb{R}^2$ with the usual topology and $Y := \mathbb{R} \times \{0\} \subseteq X$. The set $A := (0,1) \times \{0\} = (0,1) \times \mathbb{R} \cap Y$ is an open subset of Y in the subspace topology. Clearly, A is not an open subset of X.

2. *Continuous functions*

As usual we employ structure preserving maps to compare topological spaces.

Definition 2.1. Let $f : \mathfrak{X} \to \mathfrak{Y}$ be a function between closure spaces.
(a) f is *continuous* if $f^{-1}[C]$ is closed, for every closed set $C \subseteq Y$.
(b) f is *closed* if $f[C]$ is closed, for every closed set $C \subseteq X$.
(c) f is a *homeomorphism* if it is bijective, closed, and continuous.

Exercise 2.1. Let $f : \mathbb{R} \to \mathbb{R}$. Show that f is continuous if and only if, for every element $x \in \mathbb{R}$ and all $\varepsilon > 0$, there exists a number $\delta > 0$ such that $|f(y) - f(x)| < \varepsilon$, for all y with $|y - x| < \delta$. Hence, for the standard topology of the real numbers the above definition coincides with the well-known definition from analysis.

Lemma 2.2. *Let $f : \mathfrak{X} \to \mathfrak{Y}$ be a function between closure spaces. The following statements are equivalent:*

(1) *f is continuous.*

(2) *$f^{-1}[O]$ is open, for every open set $O \subseteq Y$.*

(3) *$f^{-1}[O]$ is open, for every basic open set $O \subseteq Y$.*

(4) *$f^{-1}[C]$ is closed, for every basic closed set $C \subseteq Y$.*

Proof. (1) ⇒ (2) If O is open then $Y \setminus O$ is closed. Hence,

$$X \setminus f^{-1}[O] = f^{-1}[Y \setminus O]$$

is closed and $f^{-1}[O]$ is open.

(3) ⇒ (4) follows analogously. If $\mathcal{B}$ is a closed base for the topology of $\mathfrak{Y}$ then $\{ Y \setminus B \mid B \in \mathcal{B} \}$ is an open base for this topology. Hence, if $B \in \mathcal{B}$ then

$$X \setminus f^{-1}[B] = f^{-1}[Y \setminus B]$$

is open and $f^{-1}[B]$ is closed.

(2) ⇒ (3) is trivial.

(4) ⇒ (1) Let $C \subseteq Y$ be closed. Then there exists a family S of basic closed sets such that $C = \bigcap S$. Hence,

$$f^{-1}[C] = \bigcap \{ f^{-1}[B] \mid B \in S \}$$

is closed. □

Example. We claim that addition of real numbers is a continuous function $+ : \mathbb{R}^2 \to \mathbb{R}$ with regard to the usual topologies on $\mathbb{R}$ and $\mathbb{R}^2$. Since the open intervals form a base for the topology of $\mathbb{R}$ it is sufficient to check that the preimage of every open interval (a, b) is open. This preimage is the set

$$\{ \langle x, y \rangle \in \mathbb{R}^2 \mid a - x < y < b - x \}$$

which is open in the topology of $\mathbb{R}^2$.

Exercise 2.2. Prove that multiplication $\cdot : \mathbb{R}^2 \to \mathbb{R}$ is also continuous.

Lemma 2.3. *Let $f : \mathfrak{X} \to \mathfrak{Y}$ be a function between topological spaces.*

(a) *f is continuous if, and only if, there exists a closed subbase $\mathcal{B}$ of $\mathfrak{Y}$ such that $f^{-1}[B]$ is closed, for every $B \in \mathcal{B}$.*

(b) *If f is injective, then f is closed if, and only if, there exists a closed subbase $\mathcal{B}$ of $\mathfrak{X}$ such that $f[B]$ is closed, for every $B \in \mathcal{B}$.*

Proof. (a) ($\Rightarrow$) is trivial. For ($\Leftarrow$), note that

$$f^{-1}[B_0 \cup \cdots \cup B_{n-1}] = f^{-1}[B_0] \cup \cdots \cup f^{-1}[B_{n-1}]$$

is closed, for all $B_0, \dots, B_{n-1} \in \mathcal{B}$. Hence, there is a close base

$$\mathcal{B}_+ := \{ B_0 \cup \cdots \cup B_{n-1} \mid n < \omega,\ B_0, \dots, B_{n-1} \in \mathcal{B} \}$$

of $\mathfrak{Y}$ such that $f^{-1}[B]$ is closed, for all $B \in \mathcal{B}$. Consequently, we can use Lemma 2.2 to show that that f is continuous.

(b) ($\Rightarrow$) is trivial. For ($\Leftarrow$), let $C \subseteq X$ be closed. Then there is a family $(F_i)_{i \in I}$ of finite subsets $F_i \subseteq \mathcal{B}$ such that

$$C = \bigcap_{i \in I} \bigcup F_i \,.$$

Since f is injective, it follows that

$$f[C] = f\big[\textstyle\bigcap_{i \in I} \bigcup F_i\big] = \displaystyle\bigcap_{i \in I} f\big[\textstyle\bigcup F_i\big] = \displaystyle\bigcap_{i \in I} \bigcup_{B \in F_i} f[B] \,.$$

This set is closed. □

Lemma 2.4. *Let $f : \mathfrak{X} \to \mathfrak{Y}$ and $g : \mathfrak{Y} \to \mathfrak{Z}$ be functions between closure spaces.*

(a) *If f and g are continuous then so is $g \circ f$.*

(b) *If f and g are closed then so is $g \circ f$.*

The following lemma comes in handy when one wants to prove that a piecewise defined function is continuous.

Lemma 2.5 (Gluing Lemma)**.** *Let $f : \mathfrak{X} \to \mathfrak{Y}$ be a function between topological spaces and suppose that $C_0, \dots, C_{n-1} \subseteq X$ is a finite sequence of closed sets such that $X = C_0 \cup \cdots \cup C_{n-1}$. If each restriction $f \restriction C_i$ is continuous then so is f.*

Proof. Let $A \subseteq Y$ be closed. Since $f \upharpoonright C_i$ is continues it follows that the sets $f^{-1} \upharpoonright C_i[A]$ are closed. Hence,

$$f^{-1}[A] = f^{-1} \upharpoonright C_0[A] \cup \dots \cup f^{-1} \upharpoonright C_{n-1}[A]$$

being a finite union of closed sets is also closed. □

As an application we consider topologies on partial orders and continuous functions between them.

Definition 2.6. Let $\langle A, \leq \rangle$ be a partial order. The *order topology* of A is the topology induced by the open subbase consisting of all sets $\uparrow a$ and $\downarrow a$, for $a \in A$.

Example. (a) The order topology of $\langle \mathbb{Z}, \leq \rangle$ is the discrete topology.

(b) The order topology of $\langle \mathbb{R}, \leq \rangle$ is the usual topology.

(c) The order topology of $\langle \mathbb{Q}, \leq \rangle$ is the subspace topology induced by the inclusion $\mathbb{Q} \subseteq \mathbb{R}$. If $(a, b) \subseteq \mathbb{R}$ is an open interval with irrational endpoints then $(a, b) \cap \mathbb{Q}$ is a clopen subset of $\mathbb{Q}$.

Lemma 2.7. *Let $\mathfrak{X}$ be a topological space and $\mathfrak{L}$ a lattice with the order topology. If $f, g : \mathfrak{X} \to \mathfrak{L}$ are continuous then so are the functions $f \sqcup g, f \sqcap g : \mathfrak{X} \to \mathfrak{L}$ with*

$$(f \sqcup g)(x) := f(x) \sqcup g(x) \quad \textit{and} \quad (f \sqcap g)(x) := f(x) \sqcap g(x)\,.$$

Proof. The preimages

$$(f \sqcup g)^{-1}[\downarrow a] = f^{-1}[\downarrow a] \cap g^{-1}[\downarrow a]$$
$$(f \sqcup g)^{-1}[\uparrow a] = f^{-1}[\uparrow a] \cup g^{-1}[\uparrow a]$$

of the basic open sets $\downarrow a$ and $\uparrow a$ are open. The claim for $f \sqcap g$ follows analogously. □

Corollary 2.8. *Let $\mathfrak{L}$ be a lattice with the order topology and let $C(\mathfrak{X}, \mathfrak{L})$ be the set of all continuous functions $\mathfrak{X} \to \mathfrak{L}$. If we order $f, g \in C(\mathfrak{X}, \mathfrak{L})$ by*

$$f \sqsubseteq g \quad :\text{iff} \quad f(x) \sqsubseteq g(x)\,, \quad \textit{for all } x \in X\,,$$

then $\mathfrak{C}(\mathfrak{X}, \mathfrak{L}) := \langle C(\mathfrak{X}, \mathfrak{L}), \sqsubseteq \rangle$ forms a lattice.

Proof. We have shown in the preceding lemma that $f, g \in C(\mathfrak{X}, \mathfrak{L})$ implies $f \sqcup g, f \sqcap g \in C(\mathfrak{X}, \mathfrak{L})$. Clearly, $f \sqcup g = \sup\{f, g\}$ and $f \sqcap g = \inf\{f, g\}$. □

Definition 2.9. Let $\mathfrak{A} = \langle A, \leq \rangle$ be a partial order. The *chain topology* on A is the topology where a set $U \subseteq A$ is closed if, and only if, $\sup C \in U$, for every nonempty chain $C \subseteq U$ that has a supremum.

Lemma 2.10. *Let $\langle A, \leq \rangle$ be a complete partial order. If $C \subseteq A$ is closed in the chain topology then the suborder $\langle C, \leq \rangle$ is inductively ordered.*

Lemma 2.11. *An increasing function $f : \mathfrak{A} \to \mathfrak{B}$ between partial orders is continuous (in the sense of Definition A2.3.12) if and only if it is continuous with regard to the chain topology.*

Proof. ($\Rightarrow$) Suppose that $U \subseteq B$ is a closed set such that $f^{-1}[U]$ is not closed. Then there exists a chain $C \subseteq f^{-1}[U]$ such that $\sup C$ exists but $\sup C \notin f^{-1}[U]$. Since f is increasing it follows that $f[C]$ is a chain in U. If $\sup f[C]$ does not exist then f is not continuous and we are done. Otherwise, we have $\sup f[C] \in U$ since U is closed. Since $f(\sup C) \notin U$ it follows that $\sup f[C] \neq f(\sup C)$, as desired.

($\Leftarrow$) Suppose that there is a chain $C \subseteq A$ such that $\sup C$ exists but, either $\sup f[C]$ does not or $\sup f[C] \neq f(\sup C)$. Set $c := f(\sup C)$. Since c is an upper bound of $f[C]$ but not the least one, we can find an upper bound b of $f[C]$ with $b \not\geq c$. Since $C \subseteq f^{-1}[\Downarrow b]$ is a chain with supremum $\sup C \notin f^{-1}[\Downarrow b]$ it follows that $f^{-1}[\Downarrow b]$ is not closed. The set $\Downarrow b$, on the other hand, is closed. Consequently, f is not continuous with regard to the chain topology. □

3. *Hausdorff spaces and compactness*

The finer a topology on X is, that is, the more subsets of X are closed, the smaller the vicinity of a point becomes. One extreme is the trivial topology $\{\varnothing, X\}$ where all points are near to each other. The other extreme is the discrete topology $\wp(X)$ which consists of isolated points that are far away from each other. When we equip a set X with a topology we aim at imposing a spatial relationship on the points of X. To exclude trivial cases we will adopt the basic requirement that the topology is fine enough to separate each point from every other one. Such topologies are called *Hausdorff topologies.*

Definition 3.1. Let $\mathfrak{X}$ be a topological space.

(a) $\mathfrak{X}$ is a *Hausdorff* space if, for all $x, y \in X$ with $x \neq y$, there exist open sets U and V with $x \in U$, $y \in V$, and $U \cap V = \varnothing$.

(b) $\mathfrak{X}$ is *zero-dimensional,* or *totally disconnected,* if it has an open base of clopen sets.

Example. (a) $\mathbb{R}$ is a Hausdorff space. It is not zero-dimensional.

(b) $\mathbb{Q}$ is a zero-dimensional Hausdorff space.

(c) The Zariski topology is not Hausdorff.

A typical example for the kind of topological space we are mostly interested in is given by the Cantor discontinuum.

Definition 3.2. The *Cantor discontinuum* is the space $\mathfrak{C} := \langle 2^\omega, \mathcal{C} \rangle$ where the open sets are of the form

$$\langle W \rangle := \{ x \in 2^\omega \mid w \preceq x \text{ for some } w \in W \}$$

with $W \subseteq 2^{<\omega}$. ($\preceq$ denotes the prefix order.)

Remark. The Cantor discontinuum can be regarded as the set of all branches of the infinite binary tree $\langle 2^{<\omega}, \preceq \rangle$. An open set $\langle W \rangle$ consists of all branches that contain an element of W. Correspondingly, a set C is closed if there exists a set $W \subseteq 2^{<\omega}$ such that C consists of all branches that avoid every element of W. In particular, every singleton $\{x\}$ is closed. An open base of the Cantor topology consists of the sets $\langle \{w\} \rangle$ with $w \in 2^{<\omega}$.

Lemma 3.3. *The Cantor discontinuum is a zero-dimensional Hausdorff space.*

Proof. Let $w = c_0 \dots c_{n-1} \in 2^{<\omega}$ and set $d_i := 1 - c_i$. The complement of a basic open set $\langle\{w\}\rangle$ is the open set $\bigcup\{\, \langle c_0 \dots c_{i-1} d_i \rangle \mid i < n \,\}$. Hence, every basic open set $\langle\{w\}\rangle$ is clopen.

To show that the topology is Hausdorff let $x, y \in 2^\omega$ with $x \neq y$. Then there exists a least index $n < \omega$ with $x(n) \neq y(n)$. Let $w \in 2^{<\omega}$ be the common prefix of x and y of length n and set $c := x(n)$ and $d := y(n)$. Then we have $x \in \langle wc \rangle$, $y \in \langle wd \rangle$ and $\langle wc \rangle \cap \langle wd \rangle = \varnothing$. □

Many familiar properties of the real topology are shared by all Hausdorff spaces.

Lemma 3.4. *In a Hausdorff space $\mathfrak{X}$ every singleton $\{x\}$ is closed.*

Proof. Let $x \in X$. For every $y \neq x$, there are disjoint open sets U_y, V_y with $x \in U_y$ and $y \in V_y$. The set $O := \bigcup_{y \neq x} V_y$ is open. Since $O = X \setminus \{x\}$ it follows that $\{x\}$ is closed. □

An important property of topological spaces is *compactness* which can be regarded as a strong form of completeness (the precise statement is given in Lemma 3.6 (3) below).

Definition 3.5. Let $\mathfrak{X}$ be a topological space.

(a) A *cover* of $\mathfrak{X}$ is a subset $\mathcal{U} \subseteq \wp(X)$ such that $\bigcup \mathcal{U} = X$. The cover is called *open* if every $U \in \mathcal{U}$ is an open set. A *subcover* of $\mathcal{U}$ is a subset $\mathcal{U}_0 \subseteq \mathcal{U}$ that is still a cover of $\mathfrak{X}$.

(b) $\mathfrak{X}$ is *compact* if every open cover has a finite subcover. We call a set $A \subseteq X$ compact if the subspace induced by A is compact.

(c) $\mathfrak{X}$ is *locally compact* if every point $x \in X$ has a compact neighbourhood.

Exercise 3.1. (a) Prove that $\mathbb{R}$ is not compact.

(b) Prove that a subset $A \subseteq \mathbb{R}$ is compact if, and only if, it is closed and bounded.

(c) Prove that $\mathbb{R}$ is locally compact.

(d) Prove that $\mathbb{Q}$ is not locally compact.

Lemma 3.6. *Let $\mathfrak{X}$ be a topological space. The following statements are equivalent:*

(1) *$\mathfrak{X}$ is compact.*

(2) *The topology of $\mathfrak{X}$ has an open subbase $\mathcal{B}$ such that every cover $\mathcal{U}$ of $\mathfrak{X}$ with $\mathcal{U} \subseteq \mathcal{B}$ has a finite subcover.*

(3) *If $\mathcal{C} \subseteq \wp(X)$ is a family of closed sets with $\bigcap \mathcal{C} = \varnothing$ then there exists a finite subfamily $\mathcal{C}_0 \subseteq \mathcal{C}$ with $\bigcap \mathcal{C}_0 = \varnothing$.*

Proof. (1) ⇒ (2) is trivial. (2) ⇒ (1) Let $\mathcal{F}$ be the set of all open covers of $\mathfrak{X}$ that do not have a finite subcover. We have to show that $\mathcal{F} = \varnothing$. For a contradiction, suppose otherwise. Note that $\langle \mathcal{F}, \subseteq \rangle$ is inductively ordered. Hence, there exists a maximal element $\mathcal{U} \in \mathcal{F}$. Let $\mathcal{V} := \mathcal{U} \cap \mathcal{B}$. Since no finite subset of $\mathcal{V}$ is a cover of $\mathfrak{X}$ and $\mathcal{V} \subseteq \mathcal{B}$ it follows by (2) that $\mathcal{V}$ is not a cover of $\mathfrak{X}$. Let $x \in X \smallsetminus \bigcup \mathcal{V}$ and choose some open set $U \in \mathcal{U}$ with $x \in U$. By definition of a subbase there exist finitely many sets $B_0, \dots, B_n \in \mathcal{B}$ such that

$$x \in B_0 \cap \dots \cap B_n \subseteq U \,.$$

Since $x \notin \bigcup \mathcal{V}$ we have $B_i \notin \mathcal{U}$, for all $i < n$. By maximality of $\mathcal{U}$ it follows that $\mathcal{U} \cup \{B_i\}$ has a finite subcover. That is, for every $i < n$, there exists a finite subset $\mathcal{U}_i \subseteq \mathcal{U}$ such that $\mathcal{U}_i \cup \{B_i\}$ is a cover of $\mathfrak{X}$. It follows that

$$U \cup \bigcup_{i<n} \bigcup \mathcal{U}_i \supseteq \bigcap_{i<n} B_i \cup \bigcup_{i<n} \bigcup \mathcal{U}_i \supseteq \bigcap_{i<n} (B_i \cup \bigcup \mathcal{U}_i) = X \,.$$

Consequently, $\mathcal{U}$ contains the finite subcover $\{U\} \cup \mathcal{U}_0 \cup \dots \cup \mathcal{U}_{n-1}$. Contradiction.

(1) ⇒ (3) Set $\mathcal{U} := \{ X \smallsetminus C \mid C \in \mathcal{C} \}$. If $\bigcap \mathcal{C} = \varnothing$ then $\mathcal{U}$ is an open cover of X. Hence, there exists a finite subcover $\mathcal{U}_0 \subseteq \mathcal{U}$ which implies that $\bigcap \mathcal{C}_0 = \varnothing$ where $\mathcal{C}_0 := \{ X \smallsetminus U \mid U \in \mathcal{U}_0 \} \subseteq \mathcal{C}$.

(3) ⇒ (1) Let $\mathcal{U}$ be an open cover of X and set $\mathcal{C} := \{ X \smallsetminus U \mid U \in \mathcal{U} \}$. Then $\bigcap \mathcal{C} = \varnothing$. Hence, there exists a finite subset $\mathcal{C}_0 \subseteq \mathcal{C}$ such that $\bigcap \mathcal{C}_0 = \varnothing$. This implies that $\{ X \smallsetminus C \mid C \in \mathcal{C}_0 \}$ is a finite subcover of $\mathcal{U}$. □

Lemma 3.7. *The Cantor discontinuum is compact.*

Proof. Let $\mathcal{U}$ be a cover of 2^ω consisting of basic open sets $\langle W \rangle$ with $W \subseteq 2^{<\omega}$. Set $\mathcal{W} := \{ W \subseteq 2^{<\omega} \mid \langle W \rangle \in \mathcal{U} \}$ and

$$T := 2^{<\omega} \setminus \bigcup \mathcal{W} .$$

Note that if $w \in W$ then $\langle W \rangle = \langle W \cup \{wx\} \rangle$, for all $x \in 2^{<\omega}$. Consequently, $v \in T$ implies $u \in T$, for all $u \leq v$. Hence, T is a tree. We claim that it is finite.

Suppose otherwise. As the tree T is binary we can use Lemma B2.1.9 to find an infinite branch $\alpha \in 2^\omega$ through T. This implies that $\alpha \notin \langle W \rangle$, for all $W \in \mathcal{W}$. Hence, $\alpha \notin \bigcup \mathcal{U}$. Contradiction.

Since T is finite it follows that the partial order $\langle 2^{<\omega} \setminus T, \leq \rangle$ has finitely many minimal elements $w_0, \dots, w_{n-1}$. For every $i < n$, choose some $W_i \in \mathcal{W}$ with $w_i \in W_i$. Then $\{ \langle W_0 \rangle, \dots, \langle W_{n-1} \rangle \}$ is a finite subcover of $\mathcal{U}$. □

Lemma 3.8. *If A and B are compact then so is $A \cup B$.*

Proof. Let $\mathcal{U}$ be an open cover of $A \cup B$. Since A is compact there exists a finite subset $\mathcal{V} \subseteq \mathcal{U}$ that is a cover of A. Similarly, we find a finite cover $\mathcal{W} \subseteq \mathcal{U}$ of B. Hence, $\mathcal{V} \cup \mathcal{W} \subseteq \mathcal{U}$ is a finite cover of $A \cup B$. □

Lemma 3.9. *If $\mathfrak{X}$ is compact and $A \subseteq X$ closed then A is compact.*

Proof. We employ the characterisation of Lemma 3.6 (3). Let $\mathcal{C}$ be a family of subsets of A that are closed in A. It is sufficient to show that every set in $\mathcal{C}$ is also closed in X. For every $C \in \mathcal{C}$, there is a closed set $U \subseteq X$ with $C = U \cap A$. Since A is closed it follows that so is C. □

Lemma 3.10. *Let $f : \mathfrak{X} \to \mathfrak{Y}$ be continuous. If $K \subseteq X$ is compact then so is $f[K]$.*

Proof. Let $\mathcal{U}$ be an open cover of $f[K]$. Then $\mathcal{V} := \{ f^{-1}[U] \mid U \in \mathcal{U} \}$ is an open cover of K that, by assumption, contains a finite subcover $\mathcal{V}_0 \subseteq \mathcal{V}$. For every $V \in \mathcal{V}_0$, fix some set $U_V \in \mathcal{U}$ such that $f^{-1}[U_V] = V$.

We claim that $\mathcal{U}_0 := \{ U_V \mid V \in \mathcal{V}_0 \}$ is a cover of $f[K]$. If $y \in f[K]$ then $y = f(x)$, for some $x \in K$. Choose some $V \in \mathcal{V}_0$ with $x \in V$. Then $y = f(x) \in f[V] = U_V$ is covered by $\mathcal{U}_0$. □

Lemma 3.11. *Let $\mathfrak{X}$ be a Hausdorff space and $K \subseteq X$ a compact set.*

(a) *For every $x \in X \setminus K$, there exist disjoint open sets U and V with $x \in U$ and $K \subseteq V$.*

(b) *For every compact set $A \subseteq X$, disjoint from K, there exist disjoint open sets U and V with $A \subseteq U$ and $K \subseteq V$.*

(c) *K is closed.*

Proof. (a) Let $x \in X \setminus K$. Since $\mathfrak{X}$ is a Hausdorff space we can find, for every $y \in K$, disjoint open sets $U_y, V_y \subseteq X$ with $x \in U_y$ and $y \in V_y$. Since $K \subseteq \bigcup_y V_y$ is compact there exist finitely many points $y_0, \dots, y_{n-1} \in K$ such that $K \subseteq V_{y_0} \cup \dots \cup V_{y_{n-1}} =: V$. The set $U := U_{y_0} \cap \dots \cap U_{y_{n-1}}$ is open, disjoint from V, and it contains x.

(b) The proof is similar to that of (a). Applying (a) we fix, for every $x \in K$, disjoint open sets U_x and V_x with $x \in V_x$ and $A \subseteq U_x$. Since $K \subseteq \bigcup_x V_x$ there exist finitely many elements $x_0, \dots, x_{n-1} \in K$ with $K \subseteq V_{x_0} \cup \dots \cup V_{x_{n-1}} =: V$. The set $U := U_{x_0} \cap \dots \cap U_{x_{n-1}}$ is open, disjoint from V, and it contains A.

(c) For every $x \in X \setminus K$, we can use (a) to find an open set U_x with $x \in U_x$ and $K \cap U_x = \varnothing$. Since $X \setminus K = \bigcup_x U_x$ is open it follows that K is closed. □

We turn to an investigation of locally compact Hausdorff spaces. The following lemma shows that these are very similar to the real topology.

Lemma 3.12. *Let $\mathfrak{X}$ be a locally compact Hausdorff space.*

(a) *For every neighbourhood U of a point $x \in X$, there exists a compact neighbourhood $V \subseteq U$ of x.*

(b) *For all sets $K \subseteq O \subseteq X$ where K is compact and O is open, there exists an open set U such that $K \subseteq U \subseteq \mathrm{cl}(U) \subseteq O$ and $\mathrm{cl}(U)$ is compact.*

(c) *If $C \subseteq X$ is closed and $O \subseteq X$ is open then the subspace induced by $C \cap O$ is a locally compact Hausdorff space.*

Proof. (a) Replacing U by $\operatorname{int}(U)$ we may assume that U is open. Let K be a compact neighbourhood of x. If $K \subseteq U$ we are done. Otherwise, the set $A := K \setminus U = K \cap (X \setminus U)$ is closed. Since $A \subseteq K$ it is also compact. There exist disjoint open sets W_0, W_1 with $A \subseteq W_0$ and $x \in W_1$. The set $V := K \cap (X \setminus W_0) = K \setminus W_0$ is closed, compact, and it contains x. Furthermore, $K \setminus U \subseteq W_0$ implies that $V = K \setminus W_0 \subseteq U$.

(b) By (a), we can choose, for every $x \in K$, a compact neighbourhood $W_x \subseteq O$. The family

$$\mathcal{W} := \{ \operatorname{int}(W_x) \mid x \in K \}$$

is an open cover of K. By compactness, there exists a finite subcover $\mathcal{W}_0 \subseteq \mathcal{W}$. The set $U := \bigcup \mathcal{W}_0$ is open and we have

$$\begin{aligned}\operatorname{cl}(U) = \operatorname{cl}(\textstyle\bigcup \mathcal{W}_0) &= \bigcup \{ \operatorname{cl}(\operatorname{int}(W_x)) \mid \operatorname{int}(W_x) \in \mathcal{W}_0 \} \\ &\subseteq \bigcup \{ W_x \mid \operatorname{int}(W_x) \in \mathcal{W}_0 \} \subseteq O\,.\end{aligned}$$

Finally, $\operatorname{cl}(U)$ is compact because it is a finite union of compact sets.

(c) Every subspace of a Hausdorff space is Hausdorff. To prove that $C \cap O$ is locally compact, let $x \in C \cap O$. By (a), there exists a compact neighbourhood $K \subseteq O$ of x. The set $V := C \cap K \subseteq C \cap O$ is compact. Furthermore, V is a neighbourhood of x in $C \cap O$ since $x \in C \cap \operatorname{int}(K)$ and $C \cap \operatorname{int}(K)$ is open in $C \cap O$. □

Theorem 3.13. *A Hausdorff space $\mathfrak{X}$ is locally compact if and only if there exist a compact Hausdorff space $\mathfrak{Y}$ such that $X \subseteq Y$ is an open subset of Y.*

Proof. ($\Leftarrow$) If Y is compact and $X \subseteq Y$ is open then Lemma 3.12 (c) implies that $X = X \cap Y$ is locally compact.

($\Rightarrow$) We set $Y := X \cup \{\infty\}$ where $\infty \notin X$ is a new point. Let $\mathcal{C}$ be the topology of $\mathfrak{X}$. We define the topology of $\mathfrak{Y}$ by

$$\mathcal{D} := \{ C \cup \{\infty\} \mid C \in \mathcal{C} \} \cup \{ K \mid K \subseteq X \text{ is compact} \}\,.$$

Let us show that $\mathcal{D}$ is a topology. Since $\varnothing$ is compact we have

$$\varnothing \in \mathcal{D} \quad \text{and} \quad Y = X \cup \{\infty\} \in \mathcal{D}\,.$$

Furthermore, if $A, B \in \mathcal{D}$ then either $\infty \in A \cup B$ and $(A \cup B) \setminus \{\infty\}$ is closed in X, or A and B are compact in X and so is $A \cup B$. In both cases it follows that $A \cup B \in \mathcal{D}$.

Finally, suppose that $Z \subseteq \mathcal{D}$. If $\infty \in \bigcap Z$ then $\bigcap Z \setminus \{\infty\}$ being closed in X it follows that $\bigcap Z \in \mathcal{D}$. Otherwise, there is a compact set $K \in Z$ and $\bigcap Z \subseteq X$ is closed in X. Since $\bigcap Z \subseteq K$ it follows that it is also compact. Hence, $\bigcap Z \in \mathcal{D}$.

Since $\{\infty\} = \varnothing \cup \{\infty\} \in \mathcal{D}$ it follows that X is an open subset of Y. Hence, it remains to prove that $\mathfrak{Y}$ is a compact Hausdorff space.

If $x \neq y$ are points in X then X contains disjoint open neighbourhoods of x and y. These are also open in Y. Similarly, for $x \in X$ and ∞, we can select a compact neighbourhood $K \subseteq X$ of x. Then $\operatorname{int}(K)$ and $Y \setminus K$ are disjoint open sets with $x \in \operatorname{int}(K)$ and $\infty \in Y \setminus K$. Consequently, $\mathfrak{Y}$ is a Hausdorff space.

For compactness, let $Z \subseteq \mathcal{D}$ be a family with $\bigcap Z = \varnothing$. Since $\infty \notin \bigcap Z$ there is a set $K \in Z$ that is compact in X. The family,

$$Z' := \{\, C \cap K \mid C \in Z \,\}$$

is a family of closed subsets of K with $\bigcap Z' = \varnothing$. Since K is compact it follows that there is a finite subset $Z'_0 \subseteq Z'$ with $\bigcap Z'_0 = \varnothing$. Suppose that

$$Z'_0 = \{C_0 \cap K, \dots, C_{n-1} \cap K\}\,.$$

Then $Z_0 := \{K, C_0, \dots, C_{n-1}\}$ is a finite subset of Z with $\bigcap Z_0 = \varnothing$. □

4. *The Product topology*

Definition 4.1. Let $(\mathfrak{X}_i)_{i \in I}$ be a sequence of topological space. Their *product* $\prod_{i \in I} \mathfrak{X}_i$ is the space with universe $\prod_{i \in I} X_i$ whose topology has as open base all sets of the form $\prod_{i \in I} O_i$ where each $O_i \subseteq X_i$ is open and there are only finitely many i with $O_i \neq X_i$.

Example. The Cantor discontinuum is the product $\prod_{n<\omega}[2]$ where each factor $[2]$ is equipped with the discrete topology.

Lemma 4.2. *The product topology is the least topology such that every projection is continuous.*

Proof. Let $\mathfrak{X}_i$, $i \in I$, be a family of topological spaces and let $\mathcal{C}$ be the product topology. Set

$$\mathcal{B} := \left\{ \mathrm{pr}_k^{-1}[O] \;\middle|\; k \in I,\ O \subseteq X_k \text{ open} \right\}.$$

Since $\mathcal{B}$ is an open subbase of $\mathcal{C}$ it follows that $\mathrm{pr}_k^{-1}[O]$ is open, for every open $O \subseteq X_k$. Hence, $\mathrm{pr}_k : \prod_i X_i \to X_k$ is continuous.

Let $\mathcal{C}'$ be another topology on $\prod_i X_i$ such that all projections pr_k are continuous. If $O \subseteq X_k$ is open then $\mathrm{pr}_k^{-1}[O]$ is open in $\mathcal{C}'$. Hence, every set of $\mathcal{B}$ is open in $\mathcal{C}'$. Since $\mathcal{B}$ is a subbase of $\mathcal{C}$ it follows that every open set of $\mathcal{C}$ is open in $\mathcal{C}'$, that is, $\mathcal{C} \subseteq \mathcal{C}'$. □

Lemma 4.3. *Let $\mathfrak{X}_i$, for $i \in I$, be nonempty topological spaces.*

(a) *The product $\prod_{i\in I} \mathfrak{X}_i$ is a Hausdorff space if and only if each factor $\mathfrak{X}_i$ is a Hausdorff space.*

(b) *The product space $\prod_{i\in I} \mathfrak{X}_i$ is zero-dimensional if and only if each factor $\mathfrak{X}_i$ is zero-dimensional.*

Proof. (a) ($\Leftarrow$) Let $(x_i)_i, (y_i)_i \in \prod_i X_i$ be distinct. Fix some index i with $x_i \neq y_i$. Since X_i is Hausdorff there exist disjoint open sets $U, V \subseteq X_i$ with $x_i \in U$ and $y_i \in V$. Hence, $U_* := \mathrm{pr}_i^{-1}[U]$ and $V_* := \mathrm{pr}_i^{-1}[V]$ are disjoint open sets with $(x_i)_i \in U_*$ and $(y_i)_i \in V_*$.

($\Rightarrow$) Fix elements $z_i \in X_i$, for $i \in I$. For $x \in X_k$, let $x^* := (x_i)_i$ where

$$x_i := \begin{cases} x & \text{if } i = k\,, \\ z_i & \text{otherwise}\,. \end{cases}$$

To show that $\mathfrak{X}_k$ is a Hausdorff space let $x, y \in X_k$ be distinct. By assumption there are disjoint open sets $U, V \subseteq \prod_i X_i$ with $x^* \in U$ and

$y^* \in V$. W.l.o.g. we may assume that $U = \prod_i U_i$ and $V = \prod_i V_i$ are basic open with open sets $U_i, V_i \subseteq X_i$. It follows that $x \in U_k$ and $y \in V_k$. Furthermore, $U_k \cap V_k = \emptyset$ since $z \in U_k \cap V_k$ would imply that $z^* \in \prod_i U_i \cap \prod_i V_i = \emptyset$.

(b) ($\Rightarrow$) Suppose that $\prod_i \mathfrak{X}_i$ is zero-dimensional. Fix elements $z_i \in X_i$ and define the functions $f_k : X_k \to \prod_i X_i : x \mapsto (y_i)_i$ where

$$y_i := \begin{cases} x & \text{if } i = k\,, \\ z_i & \text{otherwise}\,. \end{cases}$$

Then f_k is a homeomorphism from $\mathfrak{X}_k$ to a subspace of $\prod_i \mathfrak{X}_i$. Since every subspace of a zero-dimensional space is zero-dimensional it follows that so is $\mathfrak{X}_k$.

($\Leftarrow$) Suppose that every factor $\mathfrak{X}_i$ has an open base $\mathcal{B}_i$ of clopen sets. The space $\prod_i \mathfrak{X}_i$ has an open base consisting of all sets of the form

$$\mathrm{pr}_{k_0}^{-1}[B_0] \cap \dots \cap \mathrm{pr}_{k_n}^{-1}[B_n]$$

where $B_i \in \mathcal{B}_{k_i}$. Since each element of $\mathcal{B}_{k_i}$ is clopen, the projections pr_{k_i} are continuous, and the family of clopen sets is closed under boolean operations it follows that these sets are clopen. □

Theorem 4.4 (Tychonoff). *Let $\mathfrak{X}_i$, for $i \in I$, be nonempty topological spaces. The product space $\prod_{i \in I} \mathfrak{X}_i$ is compact if and only if each factor $\mathfrak{X}_i$ is compact.*

Proof. ($\Rightarrow$) Let $\mathcal{U}$ be an open cover of X_i. Then

$$\mathcal{V} := \{ \mathrm{pr}_i^{-1}[U] \mid U \in \mathcal{U} \}$$

is an open cover of $\prod_i X_i$. Consequently, there exists a finite subcover $\mathcal{V}_0 \subseteq \mathcal{V}$ and $\{ U \in \mathcal{U} \mid \mathrm{pr}_i^{-1}[U] \in \mathcal{V}_0 \}$ is a finite subcover of $\mathcal{U}$.

($\Leftarrow$) Let $\mathcal{U}$ be a cover of $\prod_i \mathfrak{X}_i$. By Lemma 3.6, we may assume that every set in $\mathcal{U}$ is of the form $\mathrm{pr}_i^{-1}(U)$ where $i \in I$ and $U \subseteq X_i$ is open. For $i \in I$, let

$$\mathcal{U}_i := \{ U \subseteq X_i \mid \mathrm{pr}_i^{-1}[U] \in \mathcal{U} \}\,.$$

We claim that there is some index $i \in I$ such that $\bigcup \mathcal{U}_i = X_i$. Suppose otherwise. Then, for every $i \in I$, we can find a point $x_i \in X_i \setminus \bigcup \mathcal{U}_i$. Hence, $(x_i)_i \notin \bigcup \mathcal{U}$ and $\mathcal{U}$ is not a cover of $\prod_i \mathfrak{X}_i$. Contradiction.

Fix such an index i. Since $\mathfrak{X}_i$ is compact there exists a finite subcover $\mathcal{U}_0 \subseteq \mathcal{U}_i$ of $\mathfrak{X}_i$. It follows that $\{ \mathrm{pr}_i^{-1}[U] \mid U \in \mathcal{U}_0 \}$ is a finite subcover of $\mathcal{U}$. □

Lemma 4.5. *Let $f : \mathfrak{Y}_0 \times \cdots \times \mathfrak{Y}_{n-1} \to \mathfrak{Z}$ and $g_i : \mathfrak{X}_i \to \mathfrak{Y}_i$, for $i < n$, be functions and define $h : \mathfrak{X}_0 \times \cdots \times \mathfrak{X}_{n-1} \to \mathfrak{Z}$ by*

$$h(\bar{a}) = f(g_0(a_0), \dots, g_{n-1}(a_{n-1})) .$$

If f and all g_i are continuous then so is h.

Proof. Let $k : \mathfrak{X}_0 \times \cdots \times \mathfrak{X}_{n-1} \to \mathfrak{Y}_0 \times \cdots \times \mathfrak{Y}_{n-1}$ be the function such that

$$k(\bar{a}) := \langle g_0(a_0), \dots, g_{n-1}(a_{n-1}) \rangle .$$

Since $h = f \circ k$ it is sufficient to prove that k is continuous.

Let $O \subseteq X_0 \times \cdots \times X_{n-1}$ be a basic open set. Then $O = U_0 \times \cdots \times U_{n-1}$ where each U_i is open. Since g_i is continuous it follows that $g_i^{-1}[U_i]$ is also open. Consequently,

$$k^{-1}[O] = g_0^{-1}[U_0] \times \cdots \times g_{n-1}^{-1}[U_{n-1}]$$

is open. □

Example. From this lemma and the fact that addition and multiplication of real numbers are continuous functions, it follows immediately that every polynomial function $\mathbb{R}^n \to \mathbb{R}$ is continuous.

We conclude this section with two further lemmas showing that Hausdorff spaces exhibit properties familiar from real topology. The first one is similar to Lemma 3.4.

Lemma 4.6. *If $\mathfrak{X}$ is a Hausdorff space then the set*

$$\Delta := \{ \langle x, x \rangle \mid x \in X \}$$

is closed in $\mathfrak{X} \times \mathfrak{X}$.

Proof. If $\langle x, y\rangle \notin \Delta$ then there are disjoint open sets U and V with $x \in U$ and $y \in V$. Hence, $U \times V$ is an open neighbourhood of $\langle x, y\rangle$. Since U and V are disjoint we have $U \times V \cap \Delta = \emptyset$. It follows that $X \times X \smallsetminus \Delta$ is open and Δ closed. □

Lemma 4.7. *Let $f : \mathfrak{X} \to \mathfrak{Y}$ be a continuous function where $\mathfrak{Y}$ is a Hausdorff space. Then f is a closed subset of $\mathfrak{X} \times \mathfrak{Y}$.*

Proof. The function $g : X \times Y \to Y \times Y$ with $g(x, y) := \langle f(x), y\rangle$ is continuous, by Lemma 4.5. Since Δ is closed in $\mathfrak{Y} \times \mathfrak{Y}$ and

$$f = \{ \langle x, f(x)\rangle \mid x \in X \} = g^{-1}[\Delta]$$

if follows that f is closed in $\mathfrak{X} \times \mathfrak{X}$. □

5. *Dense sets and isolated points*

In this section we study two different approaches to classify subsets of a space into 'thin' and 'thick' ones. The first one is the property of Baire and the second one the Cantor-Bendixson rank.

Definition 5.1. A set $A \subseteq X$ is *dense* if $A \cap O \neq \emptyset$, for every nonempty open set O.

Example. The set $\mathbb{Q}$ is dense in $\mathbb{R}$.

Lemma 5.2. *Let $\mathfrak{X}$ be a topological space and $A \subseteq X$.*

(a) *A is dense if and only if* $\mathrm{cl}(A) = X$.

(b) $\mathrm{int}(A) = \emptyset$ *if and only if $X \smallsetminus A$ is dense.*

Proof. (a) ($\Leftarrow$) Let O be a nonempty open set. Then $C := X \smallsetminus O \neq X$. Since $\mathrm{cl}(A) = X$ it follows that $C \not\supseteq A$. This implies that $O \cap A \neq \emptyset$.

($\Rightarrow$) Let $C \supseteq A$ be closed and set $O := X \smallsetminus C$. If $O \neq \emptyset$ then we have $O \cap A \neq \emptyset$ since A is dense. It follows that $A \smallsetminus C \neq \emptyset$. Contradiction. Hence, X is the only closed set containing A, which implies that $\mathrm{cl}(A) = X$.

(b) Let $O \neq \varnothing$ be open. If $O \cap (X \smallsetminus A) = \varnothing$ then $O \subseteq A$ which implies that $\mathrm{int}(A) \neq \varnothing$. Conversely, if $O \subseteq A$ then $O \cap (X \smallsetminus A) = \varnothing$ and $X \smallsetminus A$ is not dense. □

Definition 5.3. Let $\mathfrak{X}$ be a topological space and $A \subseteq X$.

(a) A is *nowhere dense* if its closure has empty interior.

(b) A is *meagre* if A is a countable union of nowhere dense sets.

Lemma 5.4. *Let $\mathfrak{X}$ be a topological space and $A \subseteq X$.*

(a) *If A is meagre and $B \subseteq A$ then B is meagre.*

(b) *If $A = \bigcup_{n<\omega} B_n$ where each B_n is meagre then A is meagre.*

(c) *If $D \subseteq X$ is dense and $A \cap D$ is meagre in D then A is meagre in X.*

Proof. (a) Fix nowhere dense sets C_n, $n < \omega$, such that $A = \bigcup_n C_n$. Since $B = \bigcup_n (C_n \cap B)$ and every $C_n \cap B$ is nowhere dense it follows that B is also meagre.

(b) Fix nowhere dense sets C_n^k, $k, n < \omega$, such that $B_n = \bigcup_k C_n^k$. Then

$$A = \bigcup_n B_n = \bigcup_n \bigcup_k C_n^k$$

is a countable union of nowhere dense sets.

(c) Let $A = \bigcup_n B_n$ where each set $B_n \cap D$ is nowhere dense in D. It is sufficient to prove that every B_n is nowhere dense in D. Let O be the interior of the closure of B_n in X. For a contradiction, suppose that $O \neq \varnothing$. Then $O \subseteq \mathrm{cl}_{\mathfrak{X}}(B)$ implies $O \cap D \subseteq \mathrm{cl}_{\mathfrak{D}}(B \cap D)$. Since $O \cap D$ is open in D we have $O \cap D \subseteq \mathrm{int}_{\mathfrak{D}}(\mathrm{cl}_{\mathfrak{D}}(B \cap D))$. But D is dense in X and O is open. Hence, $O \cap D \neq \varnothing$ and $B \cap D$ is not nowhere dense in D. Contradiction. □

This lemma shows that the meagre subsets $A \subseteq X$ form an ideal in $\wp(X)$ that is closed under countable unions. We are interested in spaces $\mathfrak{X}$ where this ideal is proper. The next lemma gives several equivalent characterisations of such spaces.

Lemma 5.5. *Let $\mathfrak{X}$ be a topological space. The following statements are equivalent:*

1. *If, for every $n < \omega$, A_n is a closed set with empty interior then $\bigcup_{n<\omega} A_n$ has empty interior.*
2. *If A_n is open and dense, for every $n < \omega$, then $\bigcap_{n<\omega} A_n$ is dense.*
3. *If A is open and nonempty then A is not meagre.*
4. *If A is meagre then $X \smallsetminus A$ is dense.*

Proof. (1) ⇒ (2) If A_n is open and dense then $X \smallsetminus A_n$ is a closed set with empty interior. By (1), it follows that $B = \bigcup_n (X \smallsetminus A_n)$ has empty interior. Consequently, $\bigcap_{n<\omega} A_n = X \smallsetminus B$ is dense.

(2) ⇒ (3) Suppose that A is open, nonempty, and meagre. Then there are nowhere dense sets B_n such that $A = \bigcup_{n<\omega} B_n$. Since the interior of $\mathrm{cl}(B_n)$ is empty it follows that $O_n := X \smallsetminus \mathrm{cl}(B_n)$ is dense and open. (2) implies that the set $X \smallsetminus A = \bigcap_n O_n$ is dense. Consequently, A has empty interior and, since A is open it follows that $A = \varnothing$. A contradiction.

(3) ⇒ (4) Suppose that A is meagre but $X \smallsetminus A$ is not dense. Then $\mathrm{int}(A) \neq \varnothing$ and there exists a nonempty open subset $O = \mathrm{int}(A) \subseteq A$ of A. By (3), it follows that O is not meagre. This contradicts Lemma 5.4.

(4) ⇒ (1) Let $B = \bigcup_{n<\omega} A_n$ where each A_n is a closed set with empty interior. Then B is meagre and it follows by (4) that $X \smallsetminus B$ is dense. Consequently, we have $\mathrm{int}(B) = \varnothing$. □

Definition 5.6. A topological space $\mathfrak{X}$ has the *property of Baire* if there is no set $A \subseteq X$ that is nonempty, open, and meagre.

Lemma 5.7. *Let $\mathfrak{X}$ be a topological space with the property of Baire. If A is a meagre set then the subspace $X \smallsetminus A$ has the property of Baire. In particular, $X \smallsetminus A$ is not meagre.*

Proof. Let A be a meagre subset of X. By Lemma 5.5 (4), it follows that $X \smallsetminus A$ is dense. According to Lemma 5.4 (c), if B is a meagre set in $X \smallsetminus A$ then B is also meagre in $\mathfrak{X}$. By Lemma 5.4 it follows that $A \cup B$ is also meagre. Consequently, $C = (X \smallsetminus A) \smallsetminus B = X \smallsetminus (A \cup B)$ is dense in $\mathfrak{X}$ and, therefore, C is also dense in $X \setminus A$. By Lemma 5.5, it follows that $X \setminus A$ has the property of Baire. □

Theorem 5.8 (Baire). *Every locally compact Hausdorff space $\mathfrak{X}$ has the property of Baire.*

Proof. We show that $\mathfrak{X}$ has the property of Lemma 5.5 (2). Let $(A_n)_{n<\omega}$ be a family of open dense subsets of $\mathfrak{X}$. Let O_0 be an arbitrary nonempty open set in $\mathfrak{X}$. We have to prove that $O_0 \cap \bigcap_n A_n \neq \varnothing$. We construct a decreasing chain

$$O_0 \supseteq \mathrm{cl}(O_0) \supseteq O_1 \supseteq \mathrm{cl}(O_1) \supseteq \dots$$
$$\dots \supseteq O_n \supseteq \mathrm{cl}(O_n) \supseteq O_{n+1} \supseteq \mathrm{cl}(O_{n+1}) \supseteq \dots$$

where each O_n is nonempty and open, $\mathrm{cl}(O_n)$ is compact, and $\mathrm{cl}(O_n) \subseteq A_n$.

Suppose that O_n is already defined. Since A_n is dense there exists an element $a_n \in O_n \cap A_n$. Since the singleton $\{a_n\}$ is compact we can use Lemma 3.12 (b) to find an open set O_{n+1} such that

$$a_n \in O_{n+1} \subseteq \mathrm{cl}(O_{n+1}) \subseteq O_n \cap A_n$$

and $\mathrm{cl}(O_{n+1})$ is compact.

Since $C := \bigcap_n \mathrm{cl}(O_n)$ is the intersection of a decreasing sequence of nonempty compact sets it follows that $C \neq \varnothing$. Furthermore, we have $C \subseteq O_0$ and $C \subseteq A_n$, for every n. □

Definition 5.9. Let $\mathfrak{X}$ be a topological space and $A \subseteq X$. A point $x \in X$ is an *accumulation point* of A if $x \in \mathrm{cl}(A \setminus \{x\})$. A point $a \in A$ that is not an accumulation point of A is called *isolated*.

Remark. x is an isolated point of X if and only if the set $\{x\}$ is open.

Lemma 5.10. *Let $\mathfrak{X}$ be a topological space. The following statements are equivalent:*

(1) *$\mathfrak{X}$ is a finite Hausdorff space.*

(2) $\mathfrak{X}$ *is a Hausdorff space with a finite dense subset.*

(3) $\mathfrak{X}$ *is a finite space with discrete topology.*

(4) $\mathfrak{X}$ *is compact and every point is isolated.*

Proof. (1) ⇒ (2) is trivial.

(2) ⇒ (1) Suppose that $A = \{a_0, \dots, a_{n-1}\}$ is dense in $\mathfrak{X}$. Each singleton $\{a_i\}$ is closed since $\mathfrak{X}$ is a Hausdorff space. Hence, their union $A = \{a_0\} \cup \dots \cup \{a_{n-1}\}$ is also closed. Since A is dense in X it follows by Lemma 5.2 that $A = \mathrm{cl}(A) = X$. Thus, X is finite.

(1) ⇒ (3) Suppose that $X = \{x_0, \dots, x_{n-1}\}$ and let $A \subseteq X$ be an arbitrary set. We claim that A is open. Since $\mathfrak{X}$ is Hausdorff we can choose open sets U_{ik}, for $i \neq k$, such that $x_i \in U_{ik}$ and $x_k \notin U_{ik}$. Let $O_i := \bigcap_{k \neq i} U_{ik}$. Then we have $O_i = \{x_i\}$ and $A = \bigcup \{ O_i \mid x_i \in A \}$ and these sets are open.

(3) ⇒ (4) Let $X = \{x_0, \dots, x_{n-1}\}$. Since $\{x_i\}$ is open it follows that every element is isolated. For compactness, suppose that $(U_i)_{i \in I}$ is an open cover of X. For every x_k, we fix some $i_k \in I$ with $x_k \in U_{i_k}$. Then $(U_{i_k})_{k<n}$ is a finite subcover of X.

(4) ⇒ (1) For every pair $x \neq y$ of distinct points we have the disjoint open neighbourhoods $\{x\}$ and $\{y\}$. Hence, $\mathfrak{X}$ is a Hausdorff space.

To show that $\mathfrak{X}$ is finite fix, for every $x \in X$, an open neighbourhood U_x isolating x, i.e., $U_x = \{x\}$. Then $\mathcal{U} = \{ U_x \mid x \in X \}$ is an open cover of X. By compactness, we can find a finite subcover $\mathcal{U}_0 = \{ U_x \mid x \in X_0 \}$ with $X_0 \subseteq X$. It follows that

$$X = \bigcup_{x \in X} U_x = \bigcup_{x \in X_0} U_x = X_0$$

is also finite. □

Definition 5.11. Let $\mathfrak{X}$ be a topological space and $A \subseteq X$. The *Cantor-Bendixson rank* $\mathrm{rk}_{\mathrm{CB}}(x/A)$ of an element $x \in X$ with respect to A is defined as follows:

- $\mathrm{rk}_{\mathrm{CB}}(x/A) = -1$ iff $x \notin A$.

- $\mathrm{rk}_{\mathrm{CB}}(x/A) \geq 0$ iff $x \in A$.
- $\mathrm{rk}_{\mathrm{CB}}(x/A) \geq \alpha + 1$ if $\mathrm{rk}_{\mathrm{CB}}(x/A) \geq \alpha$ and x is an accumulation point of the set $\{ a \in A \mid \mathrm{rk}_{\mathrm{CB}}(a/A) \geq \alpha \}$.
- For limit ordinals δ, we set $\mathrm{rk}_{\mathrm{CB}}(x/A) \geq \delta$ if $\mathrm{rk}_{\mathrm{CB}}(x/A) \geq \alpha$, for all $\alpha < \delta$.

The Cantor-Bendixson rank of A is

$$\mathrm{rk}_{\mathrm{CB}}(A) := \sup \{ \mathrm{rk}_{\mathrm{CB}}(a/A) \mid a \in A \} .$$

Remark. A point a is an isolated point of A if and only if $\mathrm{rk}_{\mathrm{CB}}(a/A) = 0$.

Proposition 5.12. *Let $\mathfrak{X}$ be a topological space. For $\alpha \in \mathrm{On} \cup \{\infty\}$, define*

$$X^{<\alpha} := \{ x \in X \mid \mathrm{rk}_{\mathrm{CB}}(x/X) < \alpha \}$$

and set $X^{\geq\alpha} := X \setminus X^{<\alpha}$ and $X^{\alpha} := X^{\geq\alpha} \cap X^{<\alpha+1}$.

(a) $\mathrm{rk}_{\mathrm{CB}}(X) \geq |X|^+$ *implies* $\mathrm{rk}_{\mathrm{CB}}(X) = \infty$.

(b) *Each set $X^{<\alpha}$ is open, while $X^{\geq\alpha}$ is closed.*

(c) *X^{∞} is a closed set without isolated points.*

(d) *The following statements are equivalent:*

   (1) *The isolated points are dense in X.*

   (2) *X^{∞} is nowhere dense.*

   (3) $\mathrm{int}(X^{\infty}) = \varnothing$.

Proof. (a) By definition, $X^{\geq\alpha} = X^{\geq\alpha+1}$ implies $X^{\geq\alpha} = X^{\infty}$. Since the sequence $(X^{\geq\alpha})_{\alpha}$ is decreasing it follows that there is some $\alpha < \kappa^+$ with $X^{\geq\alpha} \setminus X^{\geq\alpha+1} = \varnothing$. Consequently, $X^{\geq\alpha} = X^{\infty}$. If $X^{\geq\alpha} = \varnothing$ then we have $\mathrm{rk}_{\mathrm{CB}}(X) \leq \alpha < \kappa^+$. Otherwise, $\mathrm{rk}_{\mathrm{CB}}(X) = \infty$.

(b) Suppose that there is some element $x \in \mathrm{cl}(X^{\geq\alpha}) \setminus X^{\geq\alpha}$. Let $\beta := \mathrm{rk}_{\mathrm{CB}}(x/X) < \alpha$. Then $x \in \mathrm{cl}(X^{\geq\alpha}) = \mathrm{cl}(X^{\geq\alpha} \setminus \{x\}) \subseteq \mathrm{cl}(X^{\geq\beta} \setminus \{x\})$ implies that x is an accumulation point of $X^{\geq\beta}$. This implies that $x \in X^{\geq\beta+1}$. A contradiction.

(c) We have seen in (b) that X^∞ is closed. Fix some $\alpha < |X|^+$ with $X^{\geq\alpha} = X^\infty$. If $X^{\geq\alpha}$ had an isolated point then we would have $X^\infty \subseteq X^{\geq\alpha+1} \subset X^{\geq\alpha}$. Contradiction.

(d) The equivalence (2) $\Leftrightarrow$ (3) follows from the fact that X^∞ is closed. It remains to prove (1) $\Leftrightarrow$ (3). If X° is dense in X then so is $X^{<\infty} \supseteq X^\circ$. By Lemma 5.2 (b), it follows that $\operatorname{int}(X^\infty) = \emptyset$. Conversely, let $O \subseteq X$ be a nonempty open set. Choose some $a \in O$ such that $\alpha := \operatorname{rk}_{\mathrm{CB}}(a/X) < \infty$ is minimal. Since a is an isolated point of $X^{\geq\alpha}$ it follows that there is an open set U with $U \cap X^{\geq\alpha} = \{a\}$. By choice of a we have $O \subseteq X^{\geq\alpha}$ and it follows that $U \cap O = \{a\}$. Hence, $\{a\}$ is open and a is an isolated point of X. Therefore, $a \in O \cap X^\circ \neq \emptyset$, as desired. □

Lemma 5.13. *Let $\mathfrak{X}$ be a topological space and $C \subseteq X$ a closed set. For every $c \in C$, we have*

$$\operatorname{rk}_{\mathrm{CB}}(c/C) = \operatorname{rk}_{\mathrm{CB}}(c/X) .$$

Proof. We prove by induction on α that

$$\operatorname{rk}_{\mathrm{CB}}(c/C) = \alpha \quad \text{iff} \quad \operatorname{rk}_{\mathrm{CB}}(c/X) = \alpha .$$

Set

$$X^\alpha := \{ x \in X \mid \operatorname{rk}_{\mathrm{CB}}(x/X) < \alpha \} ,$$
$$C^\alpha := \{ x \in C \mid \operatorname{rk}_{\mathrm{CB}}(x/C) < \alpha \} .$$

By inductive hypothesis, we have

$$C^\alpha = X^\alpha \cap C \quad \text{and} \quad C \setminus C^\alpha = (X \setminus X^\alpha) \cap C .$$

It follows that

$$\begin{aligned} \operatorname{rk}_{\mathrm{CB}}(c/C) = \alpha \quad &\text{iff} \quad c \text{ is isolated in } C \setminus C^\alpha \\ &\text{iff} \quad c \text{ is isolated in } X \setminus X^\alpha \\ &\text{iff} \quad \operatorname{rk}_{\mathrm{CB}}(c/X) = \alpha . \end{aligned}$$

□

Lemma 5.14. *Let $f : \mathfrak{X} \to \mathfrak{Y}$ be injective and continuous. For every $x \in X$, we have*

$$\mathrm{rk}_{\mathrm{CB}}(x/X) \leq \mathrm{rk}_{\mathrm{CB}}(f(x)/Y) .$$

Proof. We prove by induction on α that

$$\mathrm{rk}_{\mathrm{CB}}(x/X) \geq \alpha \quad \text{implies} \quad \mathrm{rk}_{\mathrm{CB}}(f(x)/Y) \geq \alpha .$$

For $\alpha = 0$, there is nothing to do and, if α is a limit ordinal then the claim follows immediately from the inductive hypothesis. For the successor step, suppose that $\mathrm{rk}_{\mathrm{CB}}(x/X) \geq \alpha + 1$. Set

$$X^{\geq\alpha} := \{ x \in X \mid \mathrm{rk}_{\mathrm{CB}}(x/X) \geq \alpha \} ,$$
$$Y^{\geq\alpha} := \{ y \in Y \mid \mathrm{rk}_{\mathrm{CB}}(y/Y) \geq \alpha \} .$$

By inductive hypothesis, we know that $f[X^{\geq\alpha}] \subseteq Y^{\geq\alpha}$. For a contradiction, suppose that $\mathrm{rk}_{\mathrm{CB}}(f(x)/Y) = \alpha$. Then $f(x)$ is an isolated point of $Y^{\geq\alpha}$ and we can find an open neighbourhood O of $f(x)$ such that $Y^{\geq\alpha} \cap O = \{f(x)\}$. Hence,

$$\{x\} = f^{-1}[\{f(x)\}] = f^{-1}[Y^{\geq\alpha} \cap O] = f^{-1}[Y^{\geq\alpha}] \cap f^{-1}[O]$$
$$\supseteq X^{\geq\alpha} \cap f^{-1}[O] \supseteq \{x\} .$$

It follows that $X^{\geq\alpha} \cap f^{-1}[O] = \{x\}$ and x is an isolated point of $X^{\geq\alpha}$. Contradiction. □

Lemma 5.15. *Let $\mathfrak{X}$ be a compact Hausdorff space and $C \subseteq X$ a closed set. If $\mathrm{rk}_{\mathrm{CB}}(C) < \infty$ then the set*

$$\{ c \in C \mid \mathrm{rk}_{\mathrm{CB}}(c/C) = \mathrm{rk}_{\mathrm{CB}}(C) \}$$

is finite and nonempty.

Proof. Let $\mathfrak{C} \subseteq \mathfrak{X}$ be the subspace induced by C. By Lemma 3.9, $\mathfrak{C}$ is also a compact Hausdorff space. Replacing $\mathfrak{X}$ by $\mathfrak{C}$, we may therefore assume w.l.o.g. that $C = X$.

Let $\alpha := \mathrm{rk}_{\mathrm{CB}}(X)$. By Proposition 5.12 (b), the set $X^\alpha = X^{\geq\alpha}$ is closed. Consequently, X^α is a compact subspace of $\mathfrak{X}$ where every point is isolated. By Lemma 5.10, it follows that X^α is finite.

It remains to prove that it is nonempty. Suppose otherwise. Then $\{ X^{<\beta} \mid \beta < \alpha \}$ is an open cover of $\mathfrak{X}$. By compactness, we can find an open subcover $\{X^{<\beta_0}, \dots, X^{<\beta_n}\}$. Set $\gamma := \max\{\beta_0, \dots, \beta_n\}$. Then $X = X^{<\gamma}$ implies that $\mathrm{rk}_{\mathrm{CB}}(X) \leq \gamma < \alpha$. Contradiction. □

Lemma 5.16. *Let $\mathfrak{X}$ be a locally compact Hausdorff space. If* $\mathrm{rk}_{\mathrm{CB}}(X) = \infty$ *then* $|X| \geq 2^{\aleph_0}$.

Proof. Let $A := \{ x \in X \mid \mathrm{rk}_{\mathrm{CB}}(x/X) = \infty \}$. We prove that $|A| \geq 2^{\aleph_0}$. We choose points $x_w \in A$, for $w \in 2^{<\omega}$, and open neighbourhoods U_w of x_w such that, for all $v, w \in 2^{<\omega}$,

- $U_v \subseteq U_w$ iff $v \leq w$,
- if $v \nleq w$ and $w \nleq v$ then $U_v \cap U_w = \varnothing$.

By assumption $A \neq \varnothing$. Choose an arbitrary element $x_{\langle\rangle} \in A$, let K be a compact neighbourhood of $x_{\langle\rangle}$, and set $U_{\langle\rangle} := \mathrm{int}(K)$. Suppose that x_w has already been chosen. Since A has no isolated points there is some element

$$y \in (A \setminus \{x_w\}) \cap U_w \,.$$

We set $x_{w0} := x_w$ and $x_{w1} := y$. As $\mathfrak{X}$ is a Hausdorff space there are disjoint open sets V_0 and V_1 with $x_{w0} \in V_0$ and $x_{w1} \in V_1$. We set $U_{w0} := U_w \cap V_0$ and $U_{w1} := U_w \cap V_1$. For every $\sigma \in 2^\omega$, let

$$C_\sigma := \bigcap_{w < \sigma} \mathrm{cl}(U_w)\,.$$

Since K is compact and $\mathrm{cl}(U_w) \subseteq K$ it follows that $C_\sigma \neq \varnothing$. Furthermore, we have $C_\sigma \cap C_\rho = \varnothing$, for $\sigma \neq \rho$. Consequently,

$$|A| \geq \sum_{\sigma \in 2^\omega} |C_\sigma| \geq 2^{\aleph_0}.$$

□

6. Spectra and Stone duality

Boolean algebras can be characterised in terms of topological spaces. With every boolean algebra we can associate a topological space in such a way that we can recover the original algebra from the topology.

Definition 6.1. Let $\mathfrak{L}$ be a lattice. The *spectrum* of $\mathfrak{L}$ is the set

$$\operatorname{spec}(\mathfrak{L}) := \{\, \mathfrak{u} \subseteq L \mid \mathfrak{u} \text{ an ultrafilter} \,\}$$

of all ultrafilters of $\mathfrak{L}$. We equip $\operatorname{spec}(\mathfrak{L})$ with the topology consisting of all sets of the form

$$\langle X \rangle := \{\, \mathfrak{u} \in \operatorname{spec}(\mathfrak{L}) \mid X \subseteq \mathfrak{u} \,\}\,, \quad \text{for } X \subseteq L\,.$$

For $X = \{x\}$, we simply write $\langle x \rangle$.

Remark. Note that the sets $\langle X \rangle$ really form a topology since,

$$\operatorname{spec}(\mathfrak{L}) = \langle \varnothing \rangle\,, \qquad \varnothing = \langle L \rangle\,,$$
$$\bigcap_{i \in I} \langle X_i \rangle = \langle \textstyle\bigcup_{i \in I} X_i \rangle\,,$$
$$\langle X \rangle \cup \langle Y \rangle = \langle \{\, x \sqcup y \mid x \in X,\, y \in Y \,\} \rangle\,.$$

Lemma 6.2. *Let $\mathfrak{L}$ be a lattice.*

(a) *The sets of the form $\langle x \rangle$, for $x \in L$, form a closed base of the topology of* $\operatorname{spec}(\mathfrak{L})$.

(b) *If $\mathfrak{L}$ is a boolean algebra then every basic closed set $\langle x \rangle$ is clopen.*

Proof. (a) Every closed set $\langle X \rangle = \bigcap \{\, \langle x \rangle \mid x \in X \,\}$ is an intersection of basic closed sets.

(b) The complement $L \setminus \langle x \rangle = \langle x^* \rangle$ of a basic closed set is closed. □

Example. Let A be an infinite set. For the lattice $\mathfrak{F} = \langle F, \subseteq \rangle$ with

$$F := \{\, X \subseteq A \mid X \text{ or } A \setminus X \text{ is finite} \,\}\,,$$

we have $\operatorname{spec}(\mathfrak{F}) = \{\mathfrak{u}_\infty\} \cup \{\, \mathfrak{u}_a \mid a \in A \,\}$ where

$$\mathfrak{u}_a := \Uparrow\{a\} \quad \text{and} \quad \mathfrak{u}_\infty := \{\, X \subseteq A \mid A \smallsetminus X \text{ is finite} \,\}\,.$$

The basic closed sets are

$$\langle X \rangle = \begin{cases} \{\, \mathfrak{u}_a \mid a \in X \,\}\,, & \text{if } X \text{ is finite,} \\ \{\, \mathfrak{u}_a \mid a \in X \,\} \cup \{\mathfrak{u}_\infty\}\,, & \text{if } X \text{ is infinite.} \end{cases}$$

Each $\mathfrak{u}_a$ is isolated while $\mathfrak{u}_\infty$ is an accumulation point. Consequently, we have $\mathrm{rk}_{\mathrm{CB}}(\operatorname{spec}(\mathfrak{F})) = 1$.

Exercise 6.1. Let $\mathfrak{B}$ be a boolean algebra. Prove that a point $\mathfrak{u} \in \operatorname{spec}(\mathfrak{B})$ is isolated if, and only if, $\mathfrak{u}$ is principal.

Exercise 6.2. Prove that $\langle x \sqcup y \rangle = \langle x \rangle \cup \langle y \rangle$, $\langle x \sqcap y \rangle = \langle x \rangle \cap \langle y \rangle$, and $\langle x^* \rangle = \operatorname{spec}(\mathfrak{B}) \smallsetminus \langle x \rangle$.

Lemma 6.3. *Let $f : \mathfrak{L} \to \mathfrak{K}$ be a homomorphism between lattices. If $\mathfrak{u}$ is an ultrafilter of $\mathfrak{K}$ such that $f^{-1}[\mathfrak{u}] \neq L$, then $f^{-1}[\mathfrak{u}]$ is an ultrafilter of $\mathfrak{L}$.*

Proof. If $a \in f^{-1}[\mathfrak{u}]$ and $a \sqsubseteq b$ then $f(a) \sqsubseteq f(b) \in \mathfrak{u}$ implies $b \in f^{-1}[\mathfrak{u}]$. Similarly, if $a, b \in f^{-1}[\mathfrak{u}]$ then $f(a \sqcap b) = f(a) \sqcap f(b) \in \mathfrak{u}$ implies $a \sqcap b \in f^{-1}[\mathfrak{u}]$. Finally, if $a \sqcup b \in f^{-1}[\mathfrak{u}]$ then $f(a \sqcup b) = f(a) \sqcup f(b) \in \mathfrak{u}$ implies $f(a) \in \mathfrak{u}$ or $f(b) \in \mathfrak{u}$. Hence, $a \in f^{-1}[\mathfrak{u}]$ or $b \in f^{-1}[\mathfrak{u}]$. It follows that either $f^{-1}[\mathfrak{u}] = L$ or it is an ultrafilter. ☐

Definition 6.4. Let $f : \mathfrak{L} \to \mathfrak{K}$ be a homomorphism between lattices. If there is no ultrafilter of $\mathfrak{K}$ containing $\operatorname{rng} f$ then we can define

$$\operatorname{spec}(f) : \operatorname{spec}(\mathfrak{K}) \to \operatorname{spec}(\mathfrak{L}) : \mathfrak{u} \mapsto f^{-1}[\mathfrak{u}]\,.$$

Remark. Note that $\operatorname{spec}(f)$ is defined if (a) f is surjective, or (b) $\mathfrak{K}$ is a boolean algebra.

Lemma 6.5. *Let $f : \mathfrak{L} \to \mathfrak{K}$ be a homomorphism between lattices such that* $\operatorname{spec}(f)$ *is defined.*

(a) *The function* spec$(f) : \operatorname{spec}(\mathfrak{K}) \to \operatorname{spec}(\mathfrak{L})$ *is continuous.*

(b) *If f is surjective, then* spec(f) *is injective.*

Proof. (a) For every basic closed set $\langle a \rangle_{\mathfrak{L}} \subseteq \operatorname{spec}(\mathfrak{L})$,

$$\operatorname{spec}(f)^{-1}[\langle a \rangle_{\mathfrak{L}}] = \{ \mathfrak{u} \in \operatorname{spec}(\mathfrak{K}) \mid a \in f^{-1}[\mathfrak{u}] \} = \langle f(a) \rangle_{\mathfrak{K}} .$$

Hence, spec(f) is continuous.

(b) Let $\mathfrak{u}, \mathfrak{v} \in \operatorname{spec}(\mathfrak{K})$. If $f^{-1}[\mathfrak{u}] = f^{-1}[\mathfrak{v}]$ then Lemma A2.1.10 implies

$$\mathfrak{u} = f[f^{-1}[\mathfrak{u}]] = f[f^{-1}[\mathfrak{v}]] = \mathfrak{v} . \qquad \square$$

Since for boolean algebras the function spec is always defined, we obtain the following corollary.

Proposition 6.6. spec *is a contravariant functor from the category* $\mathfrak{Bool}$ *of boolean algebras to the category* $\mathfrak{Top}$ *of topological spaces.*

Lemma 6.7. *Let $f : \mathfrak{A} \to \mathfrak{B}$ be a homomorphism between boolean algebras.*

(a) *If f is surjective then* spec(f) *is continuous and injective.*

(b) *If f is injective then* spec(f) *is a closed continuous surjection.*

(c) *If* spec(f) *is injective then f is surjective.*

(d) *If* spec(f) *is surjective then f is injective.*

Proof. (a) was already proved in Lemma 6.5.

(b) We have already seen in Lemma 6.5 that spec(f) is continuous. To show that spec(f) is surjective let $\mathfrak{u} \in \operatorname{spec}(\mathfrak{A})$. We have to find some $\mathfrak{v} \in \operatorname{spec}(\mathfrak{B})$ with $f^{-1}[\mathfrak{v}] = \mathfrak{u}$. Set $\mathfrak{v}_0 := f[\mathfrak{u}]$. If there is some ultrafilter $\mathfrak{v} \supseteq \mathfrak{v}_0$, then $f^{-1}[\mathfrak{v}] \supseteq f^{-1}[f[\mathfrak{u}]] = \mathfrak{u}$, by injectivity of f and Lemma A2.1.10, and we are done. Hence, suppose that such an ultrafilter does not exist. By Corollary B2.4.10, we can find elements $b_0, \dots, b_n \in \mathfrak{v}_0$ with $b_0 \sqcap \dots \sqcap b_n = \bot$. Choosing elements $a_i \in \mathfrak{u}$ with $f(a_i) = b_i$ it follows that

$$f(a_0 \sqcap \dots \sqcap a_n) = b_0 \sqcap \dots \sqcap b_n = \bot .$$

Since f is injective this implies that $a_0 \sqcap \dots \sqcap a_n = \bot$. Hence, $\bot \in \mathfrak{u}$. Contradiction.

It remains to prove that $\operatorname{spec}(f)$ is closed. For $X \subseteq B$, we have to show that $f^{-1}[\langle X \rangle]$ is closed. Since $\langle X \rangle = \langle c_\uparrow(X) \rangle$ we may assume that $X = c_\uparrow(X)$ is a filter. We claim that $f^{-1}[\langle X \rangle] = \langle f^{-1}[X] \rangle$.

($\subseteq$) If $\mathfrak{u} \in \langle X \rangle$ then $X \subseteq \mathfrak{u}$ implies that $f^{-1}[X] \subseteq f^{-1}[\mathfrak{u}]$. Hence, $f^{-1}[\mathfrak{u}] \in \langle f^{-1}[X] \rangle$.

($\supseteq$) For a contradiction suppose that there is some element

$$\mathfrak{u} \in \langle f^{-1}[X] \rangle \setminus f^{-1}[\langle X \rangle] .$$

Then there is no ultrafilter $\mathfrak{v} \in \langle X \rangle$ with $f^{-1}[\mathfrak{v}] = \mathfrak{u}$. Note that every ultrafilter $\mathfrak{v}$ containing the set $X \cup f[\mathfrak{u}]$ satisfies $\mathfrak{v} \in \langle X \rangle$ and $f^{-1}[\mathfrak{v}] \supseteq f^{-1}[f[\mathfrak{u}]] = \mathfrak{u}$, by injectivity of f and Lemma A2.1.10. Hence, there is no such ultrafilter and we can use Corollary B2.4.10 to find finite subsets $C \subseteq \mathfrak{u}$ and $D \subseteq X$ such that

$$\bigsqcap f[C] \sqcap \bigsqcap D = \bot .$$

Set $c := \bigsqcap C \in \mathfrak{u}$ and $d := \bigsqcap D \in X$. Then

$$f(c) \sqcap d = \bot \quad \text{implies} \quad d \sqsubseteq f(c)^* = f(c^*) .$$

Since X is a filter it follows that $f(c^*) \in X$. Hence, $c^* \in f^{-1}[X] \subseteq \mathfrak{u}$ which implies that $\bot = c \sqcap c^* \in \mathfrak{u}$. Contradiction.

(c) Note that $\operatorname{rng} f$ induces a subalgebra of $\mathfrak{B}$. Hence, if $\operatorname{rng} f \subset B$, we can use Proposition B2.4.14 to find distinct ultrafilters $\mathfrak{u}, \mathfrak{v} \in \operatorname{spec}(\mathfrak{B})$ with $\mathfrak{u} \cap \operatorname{rng} f = \mathfrak{v} \cap \operatorname{rng} f$. Consequently, $f^{-1}[\mathfrak{u}] = f^{-1}[\mathfrak{v}]$ and $\operatorname{spec}(f)$ is not injective.

(d) For a contradiction, suppose that $\operatorname{spec}(f)$ is surjective, but f is not injective. Then there are elements $a, b \in A$ with $a \neq b$ and $f(a) = f(b)$. We distinguish three cases.

If $a \sqcap b^* \neq \bot$, there is some ultrafilter $\mathfrak{u} \in \operatorname{spec}(\mathfrak{A})$ with $a \sqcap b^* \in \mathfrak{u}$. As $\operatorname{spec}(f)$ is surjective, we can find some $\mathfrak{v} \in \operatorname{spec}(\mathfrak{B})$ with $f^{-1}[\mathfrak{v}] = \mathfrak{u}$. It follows that

$$\begin{aligned} a \in \mathfrak{u} = f^{-1}[\mathfrak{v}] \quad &\Rightarrow \quad f(a) \in \mathfrak{v} \\ &\Rightarrow \quad f(b) \in \mathfrak{v} \quad \Rightarrow \quad b \in f^{-1}[\mathfrak{v}] = \mathfrak{u}\,. \end{aligned}$$

Since $b^* \in \mathfrak{u}$ we obtain $\bot = b \sqcap b^* \in \mathfrak{u}$. A contradiction.

If $b \sqcap a^* \neq \bot$, we analogously choose an ultrafilter $\mathfrak{u}$ with $b \sqcap a^* \in \mathfrak{u}$ and we obtain $a \sqcap a^* \in \mathfrak{u}$ as above.

Hence, it remains to consider the case that $a \sqcap b^* = \bot = b \sqcap a^*$. Then $a \sqcup b^* = (a^* \sqcap b)^* = \bot^* = \top$. Hence, b^* satisfies the defining equations for the complement of a. Since complements are unique, it follow that $b^* = a^*$. Hence, $b = a$. A contradiction. □

We will show below that the functor spec has an inverse. But first let us show that the class of topological spaces of the form $\operatorname{spec}(\mathfrak{B})$, for a boolean algebra $\mathfrak{B}$, can be characterised in purely topological terms.

Definition 6.8. (a) A *Stone space* is a nonempty Hausdorff space that is compact and zero-dimensional.

(b) If $\mathfrak{S}$ is a Stone space then we denote by $\operatorname{clop}(\mathfrak{S})$ the lattice of all clopen subsets of $\mathfrak{S}$.

Example. The Cantor discontinuum $\mathfrak{C}$ is a Stone space. $\operatorname{clop}(\mathfrak{C})$ consists of all sets

$$\langle W \rangle := \{\, x \in 2^\omega \mid w \preceq x \text{ for some } w \in W \,\}$$

where $W \subseteq 2^{<\omega}$ is finite.

It follows from Lemma 4.3 and Theorem 4.4 that the class of Stone spaces is closed under products.

Lemma 6.9. *Let $\mathfrak{X}_i$, $i \in I$, be a family of nonempty topological spaces. The product $\prod_i \mathfrak{X}_i$ is a Stone space if and only if every factor $\mathfrak{X}_i$ is a Stone space.*

The next theorem states that the functors spec and clop form an equivalence between the category of boolean algebras and the category of Stone spaces.

Theorem 6.10. *Let $\mathfrak{B}$ be a boolean algebra and $\mathfrak{S}$ a Stone space.*

(a) $\operatorname{spec}(\mathfrak{B})$ *is a Stone space.*

(b) $\operatorname{clop}(\mathfrak{S})$ *is a boolean algebra.*

(c) *The function*

$$g : \mathfrak{B} \to \operatorname{clop}(\operatorname{spec}(\mathfrak{B})) : x \mapsto \langle x \rangle$$

is an isomorphism.

(d) *The function*

$$h : \mathfrak{S} \to \operatorname{spec}(\operatorname{clop}(\mathfrak{S})) : x \mapsto \{ C \in \operatorname{clop}(\mathfrak{S}) \mid x \in C \}$$

is a homeomorphism.

Proof. (a) Every basic closed set $\langle x \rangle$ is open since $\langle x \rangle = \operatorname{spec}(\mathfrak{B}) \setminus \langle x^* \rangle$. Hence, the topology is zero-dimensional.

Next, we show that it is Hausdorff. If $\mathfrak{u} \neq \mathfrak{v}$ are distinct points of $\operatorname{spec}(\mathfrak{B})$ then we can find some element $x \in \mathfrak{u} \setminus \mathfrak{v}$. This implies that $x^* \in \mathfrak{v} \setminus \mathfrak{u}$. The sets $\langle x \rangle$ and $\langle x^* \rangle$ are disjoint, open, and we have $\mathfrak{u} \in \langle x \rangle$ and $\mathfrak{v} \in \langle x^* \rangle$, as desired.

It remains to prove that $\operatorname{spec}(\mathfrak{B})$ is compact. Let $\langle x_i \rangle_{i \in I}$ be a cover of $\operatorname{spec}(\mathfrak{B})$ consisting of basic open sets. Set $X := \{ x_i \mid i \in I \}$ and let $\mathfrak{a} := c_\downarrow(X)$ be the ideal generated by X. We claim that $\mathfrak{a}$ is non-proper.

Suppose otherwise. Then we can use Theorem B2.4.7 to find an ultrafilter $\mathfrak{u}$ with $\mathfrak{u} \cap \mathfrak{a} = \varnothing$. In particular, we have $x_i \notin \mathfrak{u}$, for all i. Hence, $\mathfrak{u} \notin \bigcup_{i \in I} \langle x_i \rangle$ and $\langle x_i \rangle_i$ is not a cover of $\operatorname{spec}(\mathfrak{B})$. A contradiction.

Consequently, we have $\top \in \mathfrak{a}$. By definition of $c_\downarrow(X)$ it follows that there is a finite subset $X_0 \subseteq X$ with $\top = \bigsqcup X_0$. If $\mathfrak{v}$ is an ultrafilter then $\bigsqcup X_0 = \top \in \mathfrak{v}$ implies, by definition of an ultrafilter, that there is some

$x \in X_0$ with $x \in \mathfrak{v}$. Hence, we have found a finite subcover

$$\operatorname{spec}(\mathfrak{B}) = \bigcup_{x \in X_0} \langle x \rangle \,.$$

(b) Clearly, the complement of a clopen set is clopen. Since the class of open sets and the class of closed sets are both closed under finite intersections and unions so is the class of clopen sets. Hence, $\operatorname{clop}(\mathfrak{S})$ forms a boolean algebra.

(c) The function g is clearly an embedding. We only need to prove that it is surjective. Let U be a clopen subset of $\operatorname{spec}(\mathfrak{B})$. By (a), we can find a finite cover $\bigcup_{i \leq n} \langle x_i \rangle$ of U consisting of basic clopen sets. Since

$$U = \langle x_0 \rangle \cup \cdots \cup \langle x_n \rangle = \langle x_0 \sqcup \cdots \sqcup x_n \rangle$$

we have $U \in \operatorname{rng} g$.

(d) The set $h(x)$ is a final segment of $\operatorname{clop}(\mathfrak{S})$ and it is closed under finite intersections. Furthermore, if $C \cup D \in h(x)$ then at least one of C and D is also in $h(x)$. Hence, $h(x)$ is an ultrafilter and h is well-defined.

Since $\mathfrak{S}$ is a zero-dimensional Hausdorff space we have $\langle x \rangle \in h(x)$. Hence, $h(x) \neq h(y)$, for $x \neq y$, and h is injective. For surjectivity, let $\mathfrak{u} \in \operatorname{spec}(\operatorname{clop}(\mathfrak{S}))$. Since $\mathfrak{S}$ is compact we have $\bigcap \mathfrak{u} \neq \varnothing$. Fix some element $x \in \bigcap \mathfrak{u}$. We claim that $h(x) = \mathfrak{u}$.

Let C be a clopen set in $\mathfrak{S}$. If $C \in \mathfrak{u}$ then we have $x \in C$. Conversely, $x \notin S \setminus C$ implies that $S \setminus C \notin \mathfrak{u}$. Therefore, it follows that

$$C \in \mathfrak{u} \quad \text{iff} \quad x \in C \quad \text{iff} \quad C \in h(x)\,.$$

It remains to prove that h is a homeomorphism. Note that, if $C \in \operatorname{clop}(\mathfrak{S})$ then

$$h(x) \in \langle C \rangle \quad \text{iff} \quad C \in h(x) \quad \text{iff} \quad x \in C\,.$$

Consequently, if $\langle C \rangle \in \operatorname{spec}(\operatorname{clop}(\mathfrak{S}))$ then $h^{-1}[\langle C \rangle] = C \in \operatorname{clop}(\mathfrak{S})$. Conversely, if $C \in \operatorname{clop}(\mathfrak{S})$ then $h[C] = \{\, h(x) \mid x \in C \,\} = \langle C \rangle$ is clopen. □

Corollary 6.11. *The functor* spec *forms an equivalence between the category* $\mathfrak{Bool}$ *of boolean algebras and the opposite* $\mathfrak{Stone}^{\mathrm{op}}$ *of the category of Stone spaces. Its inverse is the functor* clop.

An immediate consequence of Theorem 6.10 is that every boolean algebra is isomorphic to an algebra of sets.

Corollary 6.12. *For every boolean algebra* $\mathfrak{B}$*, there exists a set* X *such that* $\mathfrak{B}$ *is isomorphic to a substructure of* $\langle \wp(X), \cap, \cup, {}^*, \varnothing, X \rangle$.

Corollary 6.13. *Every boolean algebra* $\mathfrak{A}$ *is a subdirect product of two-element boolean algebras* $\mathfrak{B}_2$*. In particular,* $\mathfrak{B}_2$ *is the only subdirectly irreducible boolean algebra.*

Proof. The power-set algebra $\wp(X)$ is isomorphic to $\mathfrak{B}_2^X$. □

7. Stone spaces and Cantor-Bendixson rank

The structure of Stone spaces will play an important part in the following chapters. In particular, we will be interested in their cardinality and their Cantor-Bendixson rank. We start with an observation that immediately follows from Lemma 5.10.

Lemma 7.1. *If* $\mathfrak{S}$ *is a Stone space with* $\mathrm{rk}_{\mathrm{CB}}(\mathfrak{S}) = 0$ *then* $\mathfrak{S}$ *is finite.*

A generalisation of this result is given in the next lemma which shows that the size of a Stone space is minimal if the corresponding boolean algebra has a partition rank.

Lemma 7.2. *Let* $\mathfrak{B}$ *be a boolean algebra. If* $\mathrm{rk}_{\mathrm{P}}(a) < \infty$*, for every* $a \in B$*, then then* $|\mathrm{spec}(\mathfrak{B})| \leq |B|$.

Proof. This follows immediately from Corollary B2.5.22. □

Conversely, if the boolean algebra has infinite partition rank then its Stone space is large.

Lemma 7.3. *Let $\mathfrak{B}$ be a boolean algebra and let κ, λ be cardinals. If there exists an embedding of $\lambda^{<\kappa}$ into $\mathfrak{B}$, then $|\operatorname{spec}(\mathfrak{B})| \geq \lambda^\kappa$.*

Proof. Let $(a_w)_{w \in \lambda^{<\kappa}}$ be an embedding of $\lambda^{<\kappa}$ into $\mathfrak{B}$. For sequences $\alpha \in \lambda^\kappa$, define

$$X_\alpha := \bigcap \{ \langle a_w \rangle \mid w \prec \alpha \} .$$

($\preceq$ denotes the prefix order.) If $\alpha \neq \beta$, then there exists some prefix $w \in \lambda^{<\kappa}$ and ordinals $i, k < \lambda$ with $i \neq k$ such that $wi \prec \alpha$ and $wk \prec \beta$. Consequently, we have $X_\alpha \subseteq \langle a_{wi} \rangle$ and $X_\beta \subseteq \langle a_{wk} \rangle$. Since $a_{wi} \sqcap a_{wk} = \bot$ it follows that $X_\alpha \cap X_\beta = \emptyset$.

Hence, it is sufficient to prove that $X_\alpha \neq \emptyset$, for all $\alpha \in \lambda^\kappa$. For finitely many elements $w_0 \prec \dots \prec w_n \prec \alpha$, we have

$$\langle a_{w_0} \rangle \cap \dots \cap \langle a_{w_n} \rangle = \langle a_{w_0} \sqcap \dots \sqcap a_{w_n} \rangle = \langle a_{w_n} \rangle \neq \emptyset .$$

Thus, the family $\langle a_w \rangle_{w \prec \alpha}$ has the finite intersection property and, by compactness, it follows that $X_\alpha = \bigcap_{w \prec \alpha} \langle a_w \rangle \neq \emptyset$. □

Corollary 7.4. *Let $\mathfrak{B}$ be a boolean algebra. If there is an element $a \in B$ with $\operatorname{rk}_P(a) = \infty$ then $|\operatorname{spec}(\mathfrak{B})| \geq 2^{\aleph_0}$.*

Proof. By Lemma B2.5.15, there exists an embedding $(b_w)_{w \in 2^{<\omega}}$ of $2^{<\omega}$ into $\mathfrak{B}$. Hence, the claim follows by Lemma 7.3. □

Remark. In Theorem 7.8 below we will prove that Cantor-Bendixson rank and partition rank are the same. Hence, Corollary 7.4 is just a special case of Lemma 5.16.

Combining Corollary 7.4 with Lemma 7.2, we obtain the following result.

Corollary 7.5. *Let $\mathfrak{B}$ be a countable boolean algebra. If $|\operatorname{spec}(\mathfrak{B})| > \aleph_0$ then $|\operatorname{spec}(\mathfrak{B})| = 2^{\aleph_0}$.*

In the remainder of this section we provide tools to compute the Cantor-Bendixson rank of a Stone space. First, we show that it coincides with the partition rank of the associated Boolean algebra, which is usually easier to compute.

Lemma 7.6. *Let $\mathfrak{B}$ be a boolean algebra and $a \in B$. If* $\mathrm{rk_P}(a) < \infty$ *then there exists an ultrafilter* $\mathfrak{u} \in \langle a \rangle$ *with* $\mathrm{rk_P}(\mathfrak{u}) = \mathrm{rk_P}(a)$.

Proof. For every $\mathfrak{u} \in \langle a \rangle$, choose an element $c_\mathfrak{u} \in \mathfrak{u}$ of minimal rank and degree. Then

$$\langle a \rangle = \bigcup_{\mathfrak{u} \in \langle a \rangle} \langle a \sqcap c_\mathfrak{u} \rangle \,.$$

By compactness, there exists a finite subcover

$$\langle a \rangle = \langle a \sqcap c_{\mathfrak{u}_0} \rangle \cup \dots \cup \langle a \sqcap c_{\mathfrak{u}_n} \rangle \,.$$

Hence, $a = (a \sqcap c_{\mathfrak{u}_0}) \sqcup \dots \sqcup (a \sqcap c_{\mathfrak{u}_n})$. By Lemma B2.5.11, there is some index $i \leq n$ such that

$$\mathrm{rk_P}(a) = \mathrm{rk_P}(a \sqcap c_{\mathfrak{u}_i}) \,.$$

This implies that

$$\mathrm{rk_P}(\mathfrak{u}_i) \leq \mathrm{rk_P}(a) = \mathrm{rk_P}(a \sqcap c_{\mathfrak{u}_i}) \leq \mathrm{rk_P}(c_{\mathfrak{u}_i}) = \mathrm{rk_P}(\mathfrak{u}_i) \,. \qquad \square$$

Corollary 7.7. *Let $\mathfrak{B}$ be a boolean algebra and $a \in B$.*

$$\mathrm{rk_P}(a) = \sup \{ \mathrm{rk_P}(\mathfrak{u}) \mid \mathfrak{u} \in \langle a \rangle \} \,.$$

Proof. If $\mathfrak{u} \in \langle a \rangle$, then $a \in \mathfrak{u}$ implies that $\mathrm{rk_P}(\mathfrak{u}) \leq \mathrm{rk_P}(a)$. Conversely, we can use Lemma 7.6 to find some ultrafilter $\mathfrak{u} \in \langle a \rangle$ with $\mathrm{rk_P}(\mathfrak{u}) = \mathrm{rk_P}(a)$. $\square$

Theorem 7.8. *Let $\mathfrak{B}$ be a boolean algebra. For every* $\mathfrak{u} \in \mathrm{spec}(\mathfrak{B})$, *we have*

$$\mathrm{rk_P}(\mathfrak{u}) = \mathrm{rk_{CB}}(\mathfrak{u} / \mathrm{spec}(\mathfrak{B})) \,.$$

Proof. We prove by induction on α that

$$\mathrm{rk}_{\mathrm{P}}(\mathfrak{u}) \geq \alpha \quad \text{iff} \quad \mathrm{rk}_{\mathrm{CB}}(\mathfrak{u}/\operatorname{spec}(\mathfrak{B})) \geq \alpha .$$

For $\alpha = 0$ the claim holds trivially and, if α is a limit ordinal, it follows immediately from the inductive hypothesis. Thus, suppose that $\alpha = \beta + 1$ is a successor ordinal. Let

$$X := \{ \mathfrak{u} \in \operatorname{spec}(\mathfrak{B}) \mid \mathrm{rk}_{\mathrm{P}}(\mathfrak{u}) \geq \beta \} .$$

By inductive hypothesis, we know that

$$X = \{ \mathfrak{u} \in \operatorname{spec}(\mathfrak{B}) \mid \mathrm{rk}_{\mathrm{CB}}(\mathfrak{u}/\operatorname{spec}(\mathfrak{B})) \geq \beta \} .$$

Suppose that $\mathrm{rk}_{\mathrm{P}}(\mathfrak{u}) = \beta$. Fix an element $a \in \mathfrak{u}$ of minimal partition rank and degree. If $\mathfrak{v} \in \langle a \rangle$ is an ultrafilter with $\mathfrak{v} \neq \mathfrak{u}$ then we have $\mathrm{rk}_{\mathrm{P}}(\mathfrak{v}) < \mathrm{rk}_{\mathrm{P}}(\mathfrak{u}) = \beta$, by Proposition B2.5.21. Hence, $\langle a \rangle \cap X = \{\mathfrak{u}\}$ and $\mathfrak{u}$ is an isolated point of X. This implies that $\mathrm{rk}_{\mathrm{CB}}(\mathfrak{u}/\operatorname{spec}(\mathfrak{B})) = \beta$.

Conversely, suppose that $\mathrm{rk}_{\mathrm{CB}}(\mathfrak{u}/\operatorname{spec}(\mathfrak{B})) = \beta$. Then there is a basic open set $\langle a \rangle$ such that $\langle a \rangle \cap X = \{\mathfrak{u}\}$. By inductive hypothesis it follows that $\mathrm{rk}_{\mathrm{P}}(a) \geq \mathrm{rk}_{\mathrm{P}}(\mathfrak{u}) \geq \beta$. Let P be a partition of a with $\mathrm{rk}_{\mathrm{P}}(p) = \beta$, for all $p \in P$. By Lemma 7.6, there are ultrafilters $\mathfrak{v}_p \in \langle p \rangle$, for $p \in P$, such that $\mathrm{rk}_{\mathrm{P}}(\mathfrak{v}_p) = \mathrm{rk}_{\mathrm{P}}(p) = \beta$. Hence, $\mathfrak{v}_p \in X$. It follows that

$$\mathfrak{v}_p \in \langle p \rangle \cap X \subseteq \langle a \rangle \cap X = \{\mathfrak{u}\} .$$

Consequently, $\mathfrak{v}_p = \mathfrak{u}$ and $\mathrm{rk}_{\mathrm{P}}(\mathfrak{u}) = \mathrm{rk}_{\mathrm{P}}(\mathfrak{v}_p) = \beta$. □

Corollary 7.9. *Let $\mathfrak{B}$ be a boolean algebra and $a \in B$. Then*

$$\mathrm{rk}_{\mathrm{CB}}(\langle a \rangle) = \mathrm{rk}_{\mathrm{P}}(a) .$$

Proof. By Lemma 5.13, Theorem 7.8, and Corollary 7.7, it follows that

$$\begin{aligned}
\mathrm{rk}_{\mathrm{CB}}(\langle a \rangle) &= \sup \{ \mathrm{rk}_{\mathrm{CB}}(\mathfrak{u}/\langle a \rangle) \mid \mathfrak{u} \in \langle a \rangle \} \\
&= \sup \{ \mathrm{rk}_{\mathrm{CB}}(\mathfrak{u}/\operatorname{spec}(\mathfrak{B})) \mid \mathfrak{u} \in \langle a \rangle \} \\
&= \sup \{ \mathrm{rk}_{\mathrm{P}}(\mathfrak{u}) \mid \mathfrak{u} \in \langle a \rangle \} \\
&= \mathrm{rk}_{\mathrm{P}}(a) .
\end{aligned}$$

□

Corollary 7.10. *Let $\mathfrak{S}$ be a Stone space and $C \subseteq S$ closed.*

$$\mathrm{rk}_{\mathrm{CB}}(C) = \mathrm{rk}_{\mathrm{P}}(C/\mathrm{clop}(C))$$

Proof. Let $\mathfrak{C}$ be the subspace of $\mathfrak{S}$ induced by C. By Lemma 3.9, $\mathfrak{C}$ is compact. Since every subspace of a zero-dimensional Hausdorff space is itself a zero-dimensional Hausdorff space, it follows that $\mathfrak{C}$ is a Stone space. Let $\mathfrak{B} := \mathrm{clop}(\mathfrak{C})$. Then $\mathrm{spec}(\mathfrak{B}) \cong \mathfrak{C}$ and Corollary 7.9 implies that

$$\mathrm{rk}_{\mathrm{CB}}(C) = \mathrm{rk}_{\mathrm{CB}}(\mathrm{spec}(\mathfrak{B})) = \mathrm{rk}_{\mathrm{P}}(\top/\mathfrak{B}) = \mathrm{rk}_{\mathrm{P}}(C/\mathrm{clop}(\mathfrak{C}))\,. \qquad \square$$

When applying Corollary 7.10, we have to consider clopen sets in a closed subspace of the given Stone space. The following lemma shows that such clopen sets are just restrictions of sets that are clopen in the ambient space.

Lemma 7.11. *Let $\mathfrak{B}$ be a boolean algebra, $A \subseteq B$, and let $\mathfrak{S}_A$ be the subspace of* $\mathrm{spec}(\mathfrak{B})$ *induced by $\langle A\rangle$. A set $C \subseteq \langle A\rangle$ is clopen in $\mathfrak{S}_A$ if, and only if, it is of the form $C = \langle b\rangle \cap \langle A\rangle$, for some $b \in B$.*

Proof. ($\Leftarrow$) A set of the form $C = \langle b\rangle \cap \langle A\rangle$ is obviously closed. It is open since its complement $\langle A\rangle \smallsetminus C = \langle b^*\rangle \cap \langle A\rangle$ is also closed.

($\Rightarrow$) Suppose that $C \subseteq \langle A\rangle$ is clopen in $\mathfrak{S}_A$. Then there are sets $D, E \subseteq B$ such that

$$C = \langle D\rangle \cap \langle A\rangle \quad \text{and} \quad \langle A\rangle \smallsetminus C = \langle E\rangle \cap \langle A\rangle\,.$$

Consequently,

$$\langle A\rangle \cap \langle E\rangle \cap \bigcap_{d \in D} \langle d\rangle = \langle A\rangle \cap \langle E\rangle \cap \langle D\rangle = \varnothing\,.$$

As $\mathrm{spec}(\mathfrak{B})$ is compact, there exists a finite subset $D_0 \subseteq D$ such that

$$\langle A\rangle \cap \langle E\rangle \cap \bigcap_{d \in D_0} \langle d\rangle = \varnothing\,.$$

It follows that

$$C = \langle D \rangle \cap \langle A \rangle \subseteq \langle D_0 \rangle \cap \langle A \rangle \subseteq \langle A \rangle \setminus \langle E \rangle = C\,.$$

Hence, $C = \langle b \rangle \cap \langle A \rangle$ for $b := \prod D_0$. □

Corollary 7.12. *Let $\mathfrak{S}$ be a Stone space, $C \subseteq S$ closed, and $D \in \operatorname{clop}(C)$. Then*

$$\operatorname{clop}(D) = \{\, E \in \operatorname{clop}(C) \mid E \subseteq D \,\}\,.$$

Proof. Let $\mathfrak{B} := \operatorname{clop}(\mathfrak{S})$. By Lemma 7.11, there is some $A \in B$ such that $D = A \cap C$. By the same lemma it follows that

$$\begin{aligned} E \in \operatorname{clop}(D) \quad &\text{iff} \quad E = A' \cap D \text{ for some } A' \in B \\ &\text{iff} \quad E = A' \cap A \cap C \text{ for some } A' \in B \\ &\text{iff} \quad E = A'' \cap C \text{ for some } A'' \in B \text{ with } A'' \subseteq A \\ &\text{iff} \quad E \in \operatorname{clop}(C) \text{ and } E \subseteq D\,. \end{aligned}$$

□

Corollary 7.13. *Let $\mathfrak{S}$ be a Stone space, $C \subseteq S$ closed, and $D \in \operatorname{clop}(C)$. Then*

$$\operatorname{rk}_{\mathrm{P}}(D/\operatorname{clop}(D)) = \operatorname{rk}_{\mathrm{P}}(D/\operatorname{clop}(C))\,.$$

As an application of these results, we show that, under a surjective continuous map, the Cantor-Bendixson rank never increases.

Lemma 7.14. *Let $f : \mathfrak{S} \to \mathfrak{T}$ be a surjective continuous map between Stone spaces. For every closed set $C \subseteq T$,*

$$\operatorname{rk}_{\mathrm{CB}}(C/\mathfrak{T}) \leq \operatorname{rk}_{\mathrm{CB}}(f^{-1}[C]/\mathfrak{S})\,.$$

Proof. We prove by induction on α that

$$\operatorname{rk}_{\mathrm{CB}}(C/\mathfrak{T}) \geq \alpha \quad \text{implies} \quad \operatorname{rk}_{\mathrm{CB}}(f^{-1}[C]/\mathfrak{S}) \geq \alpha\,.$$

For $\alpha = 0$, surjectivity of f implies that

$$\begin{aligned} \mathrm{rk}_{\mathrm{CB}}(C/\mathfrak{T}) \geq 0 \quad &\text{iff} \quad C \neq \varnothing \\ &\text{iff} \quad f^{-1}[C] \neq \varnothing \\ &\text{iff} \quad \mathrm{rk}_{\mathrm{CB}}(f^{-1}[C]/\mathfrak{S}) \geq 0 \,. \end{aligned}$$

For limit ordinals α, the claim follows immediately from the inductive hypothesis. For the successor step, suppose that $\mathrm{rk}_{\mathrm{CB}}(C/\mathfrak{T}) \geq \alpha + 1$. By Corollary 7.10, it follows that

$$\mathrm{rk}_{\mathrm{P}}(C/\mathrm{clop}(C)) \geq \alpha + 1 \,.$$

Consequently, we can find a sequence $(D_n)_{n<\omega}$ of disjoint, nonempty, clopen subsets $D_n \subseteq C$ such that $\mathrm{rk}_{\mathrm{P}}(D_n/\mathrm{clop}(C)) \geq \alpha$. Using Corollary 7.10 and Corollary 7.13, this implies that $\mathrm{rk}_{\mathrm{CB}}(D_n/\mathfrak{T}) \geq \alpha$. By inductive hypothesis, it therefore follows that

$$\mathrm{rk}_{\mathrm{CB}}(f^{-1}[D_n]/\mathfrak{S}) \geq \alpha \,.$$

Since, by Corollary 7.10, $(f^{-1}[D_n])_{n<\omega}$ is a sequence of disjoint, nonempty clopen subsets of $f^{-1}[C]$ with

$$\mathrm{rk}_{\mathrm{P}}\big(f^{-1}[D_n] \,/\, \mathrm{clop}(f^{-1}[C])\big) \geq \alpha \,,$$

it follows that

$$\mathrm{rk}_{\mathrm{P}}\big(f^{-1}[C] \,/\, \mathrm{clop}(f^{-1}[C])\big) \geq \alpha + 1 \,.$$

Hence, $\mathrm{rk}_{\mathrm{CB}}(f^{-1}[C]/\mathfrak{S}) \geq \alpha + 1$. □

B6. *Classical Algebra*

1. *Groups*

In this chapter we apply the general theory developed so far to the structures arising in classical algebra.

Definition 1.1. (a) A *monoid* is a structure $\mathfrak{M} = \langle M, \circ, e \rangle$ with a binary function $\circ$ and a constant e such that all elements $a, b, c \in G$ satisfy the following equations:

$$a \circ (b \circ c) = (a \circ b) \circ c \qquad (\textit{associativitiy})$$

$$a \circ e = a = e \circ a \qquad (\textit{neutral element})$$

Usually, we omit the symbol $\circ$ in $a \circ b$ and just write ab instead.

(b) A *group* is a structure $\mathfrak{G} = \langle G, \circ, {}^{-1}, e \rangle$ with a binary function $\circ$, a unary function ${}^{-1}$, and a constant e such that $\langle G, \circ, e \rangle$ is a monoid and, for all $a \in G$, we have

$$a \circ a^{-1} = e \qquad (\textit{inverse})$$

(c) A group $\mathfrak{G}$ is *abelian,* or *commutative,* if we further have

$$ab = ba\,, \qquad \text{for all } a, b \in G\,.$$

Remark. Every substructure of a group is again a group.

Example. (a) Let A be a set. The structure $\langle A^{<\omega}, \cdot, \langle\rangle \rangle$ of all finite sequences over A with concatenation forms a monoid.

(b) The integers with addition form a group $\langle \mathbb{Z}, +, -, 0 \rangle$.

(c) The positive rational numbers with multiplication form the group $\langle \mathbb{Q}^+, \cdot, {}^{-1}, 1 \rangle$.

Definition 1.2. Let $\mathfrak{M}$ be a Σ-structure. The *automorphism group*

$$\mathfrak{Aut}\,\mathfrak{M} = \langle \operatorname{Aut}\mathfrak{M}, \circ, {}^{-1}, \mathrm{id}_M \rangle$$

of $\mathfrak{M}$ consists of all automorphisms of $\mathfrak{M}$ with composition $\circ$ as multiplication and the identity function id_M as neutral element.

Exercise 1.1. Let $\mathfrak{G}$ be a group. Prove that $GG = G$ and $G^{-1} = G$ where

$$GG := \{\, gh \mid g, h \in G \,\} \quad \text{and} \quad G^{-1} := \{\, g^{-1} \mid g \in G \,\}.$$

Below we will show that the congruences of a group can be described in terms of certain subgroups. We start by looking more generally at equivalence relations induced by arbitrary subgroups.

Definition 1.3. Let $\mathfrak{U} \subseteq \mathfrak{G}$ be groups. We define

$$G/U := \{\, gU \mid g \in G \,\}.$$

The elements of G/U are called (left) *cosets* of $\mathfrak{U}$. The number $|G/U|$ of cosets is called the the *index* of $\mathfrak{U}$ in $\mathfrak{G}$.

Lemma 1.4. *Let* $\mathfrak{U} \subseteq \mathfrak{G}$ *be groups.*

(a) G/U *forms a partition of* G.

(b) *For all* $g, h \in G$, *we have a bijection* $\lambda : gU \to hU$ *with* $\lambda(x) := hg^{-1}x$.

Proof. (a) Since $g \in gU$, we have $G = \bigcup_g gU = \bigcup(G/U)$. If $gU \cap hU \neq \varnothing$ then there are elements $u, v \in U$ with $gu = hv$. Consequently, $h = g(uv^{-1}) \in gU$ which implies that $hU = gU$.

(b) To show that λ is surjective let $u \in U$. Then $hu = hg^{-1}gu = \lambda(gu)$ with $gu \in gU$. For injectivity, suppose that $\lambda(x) = \lambda(y)$ then $hg^{-1}x = hg^{-1}y$ and, multiplying with $(hg^{-1})^{-1}$ on the left, it follows that $x = y$. □

Theorem 1.5 (Lagrange). *If* $\mathfrak{U} \subseteq \mathfrak{G}$ *are groups then*

$$|G| = |G/U| \otimes |U|.$$

Proof. By the preceding lemma, we have $G = \bigcup(G/U)$ and $|gU| = |hU|$, for all $g, h \in U$. It follows that

$$|G| = \left|\bigcup(G/U)\right| = \sum_{gU \in G/U} |gU| = \sum_{gU \in G/U} |U| = |G/U| \otimes |U|\,. \qquad \square$$

The equivalence relation induced by the partition G/U does not need to be a congruence. Subgroups where it is one are called *normal.*

Definition 1.6. Let $\mathfrak{G}$ be a group. A subgroup $\mathfrak{N} \subseteq \mathfrak{G}$ is *normal* if we have $gN = Ng$, for all $g \in G$.

Remark. Every subgroup of an abelian group is normal.

Lemma 1.7. *If $\mathfrak{N}$ is a normal subgroup of $\mathfrak{G}$ then the relation*

$$g \approx_N h \quad :\text{iff} \quad gN = hN$$

is a congruence relation.

Proof. If $gN = g'N$ and $hN = h'N$ then

$$ghN = ghNN = gNhN = g'Nh'N = g'h'NN = g'h'N\,,$$

and $$\begin{aligned} g^{-1}N &= g^{-1}N^{-1} = (Ng)^{-1} = (gN)^{-1} = (g'N)^{-1} \\ &= (Ng')^{-1} = (g')^{-1}N^{-1} = (g')^{-1}N\,. \end{aligned} \qquad \square$$

Lemma 1.8. *Let $f : \mathfrak{G} \to \mathfrak{H}$ be a surjective homomorphism. If $\mathfrak{G}$ is a group then so is $\mathfrak{H}$.*

Proof. Let $x, y, z \in H$ and set $u := f(e)$. Since f is surjective there are elements $a, b, c \in G$ with $f(a) = x$, $f(b) = y$, and $f(c) = z$. It follows that

$$\begin{aligned} [xy]z &= [f(a)f(b)]f(c) = f(ab)f(c) = f((ab)c) \\ &= f(a(bc)) = f(a)f(bc) = f(a)[f(b)f(c)] = x[yz]\,, \end{aligned}$$

$$xu = f(a)f(e) = f(ae) = f(a) = x\,,$$

$$xf(a^{-1}) = f(a)f(a^{-1}) = f(aa^{-1}) = f(e) = u\,.$$

Consequently, the multiplication of $\mathfrak{H}$ is associative, u is its neutral element, and every element $x = f(a) \in H$ has the inverse $f(a^{-1})$. □

Corollary 1.9. *Let $\mathfrak{N}$ be a normal subgroup of $\mathfrak{G}$. Then the quotient*

$$\mathfrak{G}/\mathfrak{N} := \langle G/N, \cdot, ^{-1}, N \rangle$$

where the multiplication is defined by $gN \cdot hN = ghN$ is a group.

Proof. The function $g \mapsto gN$ is a surjective homomorphism $\mathfrak{G} \to \mathfrak{G}/\mathfrak{N}$. □

We have seen that every normal subgroup induces a congruence. The converse is given by the following lemma.

Lemma 1.10. *If $\approx$ is a congruence a of group $\mathfrak{G}$ then $[e]_\approx$ induces a normal subgroup of $\mathfrak{G}$.*

Proof. Let $\pi : \mathfrak{G} \to \mathfrak{G}/\approx$ be the canonical projection. Since $\{[e]_\approx\}$ induces a subgroup of the quotient $\mathfrak{G}/\approx$ it follows by Lemma B1.2.8 that the set $[e]_\approx = \pi^{-1}([e]_\approx)$ induces a subgroup of $\mathfrak{G}$. To show that this subgroup is normal, let $u \in [e]_\approx$ and $g \in G$. Then

$$\begin{aligned} [gug^{-1}]_\approx &= [g]_\approx [u]_\approx [g^{-1}]_\approx \\ &= [g]_\approx [e]_\approx [g^{-1}]_\approx = [geg^{-1}]_\approx = [e]_\approx \,, \end{aligned}$$

which implies that $gug^{-1} \in [e]_\approx$. Consequently, we have

$$g[e]_\approx g^{-1} \subseteq [e]_\approx \quad \text{and} \quad g[e]_\approx \subseteq [e]_\approx g \,.$$

Analogously, we can show that $g^{-1}ug \in [e]_\approx$, for all $u \in [e]_\approx$. This implies that $[e]_\approx g \subseteq g[e]_\approx$. □

Combining Lemmas 1.7 and 1.10, we obtain the following characterisation of the congruence lattice of a group.

Theorem 1.11. *Let $\mathfrak{G}$ be a group. Then $\mathfrak{Cong}(\mathfrak{G})$ is isomorphic to the lattice of all normal subgroups of $\mathfrak{G}$. The corresponding isomorphism is given by $\approx \mapsto [e]_\approx$ and its inverse is $\mathfrak{N} \mapsto \approx_N$.*

It follows that we can translate Theorems B1.4.12 and B1.4.18 into the language of normal subgroups.

Theorem 1.12. *Let $h : \mathfrak{G} \to \mathfrak{H}$ be a homomorphism between groups and set $K := h^{-1}[e]$. Then*

$$\mathfrak{G}/\mathfrak{K} \cong \operatorname{rng} h\,.$$

Theorem 1.13. *Let $\mathfrak{G}$ be a group with normal subgroups $\mathfrak{K}, \mathfrak{N} \subseteq \mathfrak{G}$ where $\mathfrak{K} \subseteq \mathfrak{N}$. Then $\mathfrak{N}/\mathfrak{K}$ is a normal subgroup of $\mathfrak{G}/\mathfrak{K}$ and*

$$(\mathfrak{G}/\mathfrak{K}) \big/ (\mathfrak{N}/\mathfrak{K}) \cong \mathfrak{G}/\mathfrak{N}\,.$$

A related statement is the following one.

Theorem 1.14. *Let $\mathfrak{G}$ be a group with subgroups $\mathfrak{U}, \mathfrak{N} \subseteq \mathfrak{G}$ where $\mathfrak{N}$ is normal. Then*

$$\mathfrak{U}\mathfrak{N}/\mathfrak{N} \cong \mathfrak{U}/(\mathfrak{U} \cap \mathfrak{N})\,.$$

Exercise 1.2. Prove the preceding theorem and formulate a generalisation to arbitrary structures and congruences.

2. *Group actions*

One important class of groups we will deal with frequently are automorphism groups. To study such groups we can make use of the fact that they consist of functions on some set.

Definition 2.1. Let Ω be a set.

(a) The *symmetric group* of Ω is the group

$$\mathfrak{Sym}\,\Omega := \langle \operatorname{Sym}\Omega, \circ, {}^{-1}, \mathrm{id}_\Omega \rangle$$

where the universe

$$\operatorname{Sym}\Omega := \{\, \alpha \in \Omega^\Omega \mid \alpha \text{ bijective} \,\}$$

consists of all permutations of Ω.

(b) An *action* of a group $\mathfrak{G}$ on Ω is a homomorphism $\alpha : \mathfrak{G} \to \mathfrak{Sym}\,\Omega$, that is, to every element $g \in G$ we associate a permutation $\alpha(g)$ of Ω. Such an action induces a map $G \times \Omega \to \Omega$. If α is understood then we usually write ga instead of $\alpha(g)(a)$, for $g \in G$ and $a \in \Omega$.

(c) If $\Omega = \bigcup_s \Omega_s$ is a many-sorted set then an action α of $\mathfrak{G}$ on Ω is a family of actions α_s of $\mathfrak{G}$ on Ω_s.

(d) Each action of $\mathfrak{G}$ on Ω induces an action of $\mathfrak{G}$ on Ω^n by

$$g\langle a_0, \dots, a_{n-1}\rangle := \langle ga_0, \dots, ga_{n-1}\rangle\,.$$

Remark. Any action of a group $\mathfrak{G}$ on a set Ω satisfies the following laws. For all $g, h \in G$ and $a \in \Omega$, we have

$$g(ha) = (gh)a \quad \text{and} \quad ea = a\,,$$

where e is the neutral element of G.

Example. Every subgroup $\mathfrak{G} \subseteq \mathfrak{Sym}\,\Omega$ induces a canonical action $\mathrm{id}_G : \mathfrak{G} \to \mathfrak{Sym}\,\Omega$. In particular, we have a canonical action of the automorphism group $\mathfrak{Aut}\,\mathfrak{A}$ on $A^{\bar{s}}$, for all $\bar{s}$.

Definition 2.2. Let $\mathfrak{G}$ be a group acting on Ω.

(a) For $F \subseteq G$ and $\bar{a} \subseteq \Omega$, we set

$$F(\bar{a}) := \{\, g\bar{a} \mid g \in F \,\}\,.$$

(b) The *orbit* of a tuple $\bar{a} \subseteq \Omega$ is the set $G(\bar{a})$.

(c) If there is some element $a \in \Omega$ with $G(a) = \Omega$ then we call the action *transitive.* The action is *oligomorphic* if, for every finite tuple of sorts $\bar{s}$, there are only finitely many different orbits on $\Omega^{\bar{s}}$.

Remark. For each $\bar{s}$, the orbits of all $\bar{s}$-tuples form a partition of $\Omega^{\bar{s}}$. In particular, the orbits of two $\bar{s}$-tuples are either equal or disjoint.

Example. Consider the action of the automorphism group on the structure $\langle \mathbb{Q}, \leq\rangle$. The orbit of $\langle 0, 1\rangle$ consist of all pairs $\langle a, b\rangle$ with $a < b$. It follows that $\mathbb{Q}^2$ is the disjoint union of the orbits of $\langle 0, 1\rangle$, $\langle 0, 0\rangle$, and $\langle 1, 0\rangle$. In fact, the automorphism group of $\langle \mathbb{Q}, \leq\rangle$ is oligomorphic.

Example. Every group $\mathfrak{G}$ acts on itself via *conjugation*. This action is defined by

$$\alpha(g)(h) := ghg^{-1}.$$

The orbits of α on G are called the *conjugacy classes* of $\mathfrak{G}$.

We can characterise normal subgroups of $\mathfrak{G}$ in terms of α. A subgroup $\mathfrak{N} \subseteq \mathfrak{G}$ is normal if and only if N is a union of conjugacy classes.

($\Rightarrow$) Suppose that $\mathfrak{N}$ is a normal subgroup. By definition this means that $gN = Ng$, for all $g \in G$. Consequently, we have $gNg^{-1} = Ngg^{-1} = N$ which implies that $\alpha(g)(u) \in N$, for all $u \in N$. Hence, N is a union of orbits of α.

($\Leftarrow$) Let $g \in G$. By assumption we have $gNg^{-1} = N$. Hence, $gN = gNg^{-1}g = Ng$ and $\mathfrak{N}$ is normal.

Definition 2.3. Let $\mathfrak{G}$ be a group acting on Ω and let $X \subseteq \Omega$.

(a) The *pointwise stabiliser* of X is the set

$$\mathfrak{G}_{(X)} := \{\, g \in G \mid gx = x \text{ for all } x \in X \,\}.$$

(b) Its *setwise stabiliser* is the set

$$\mathfrak{G}_{\{X\}} := \{\, g \in G \mid gX = X \,\}.$$

Remark. $\mathfrak{G}_{(X)}$ and $\mathfrak{G}_{\{X\}}$ are subgroups of $\mathfrak{G}$ with $\mathfrak{G}_{(X)} \subseteq \mathfrak{G}_{\{X\}} \subseteq \mathfrak{G}$.

We can use the following lemmas to compute the size or the number of orbits.

Lemma 2.4. *Let $\mathfrak{G}$ be a group acting on Ω and let $a \in \Omega$. Then*

$$|G| = |G(a)| \otimes |G_{(a)}|.$$

Proof. By Theorem 1.5 it is sufficient to prove that $|G(a)| = |G/G_{(a)}|$. We define a function $\mu : G/G_{(a)} \to G(a)$ by

$$\mu(gG_{(a)}) := ga.$$

First, let us show that μ is well-defined. Suppose that $gG_{(a)} = hG_{(a)}$. Then there is some $u \in G_{(a)}$ with $g = hu$. Hence,

$$\mu(gG_{(a)}) = ga = hua = ha = \mu(hG_{(a)}) .$$

Furthermore, μ is surjective since, for every $b \in G(a)$ there is some $g \in G$ with $b = ga$. Hence, $b = \mu(gG_{(a)})$. Therefore, it remains to prove that μ is injective. Suppose that $\mu(gG_{(a)}) = \mu(hG_{(a)})$. Then $ga = ha$ implies $h^{-1}ga = a$. Hence, $h^{-1}g \in G_{(a)}$ and

$$gG_{(a)} = hh^{-1}gG_{(a)} = hG_{(a)} . \qquad \square$$

Lemma 2.5. *Let $\mathfrak{G}$ be a group acting on Ω and let $a \in \Omega$. Then $G_{(ga)} = gG_{(a)}g^{-1}$.*

Proof. We have

$$\begin{aligned} h \in G_{(ga)} \quad &\text{iff} \quad hga = ga \\ &\text{iff} \quad g^{-1}hga = a \\ &\text{iff} \quad g^{-1}hg \in G_{(a)} \quad \text{iff} \quad h \in gG_{(a)}g^{-1}. \end{aligned} \qquad \square$$

Corollary 2.6. *Let $\mathfrak{G}$ be a group acting on Ω and $a, b \in \Omega$. If $G(a) = G(b)$ then $|G_{(a)}| = |G_{(b)}|$.*

Proof. Let $g \in G$ be an element with $gb = a$. The function $G_{(a)} \to G_{(b)} : h \mapsto ghg^{-1}$ is bijective. $\square$

Lemma 2.7 (Burnside). *Let $\mathfrak{G}$ be a group acting on Ω and let κ be the number of orbits. Then*

$$\kappa \otimes |G| = \sum_{g \in G} |\operatorname{fix} g| \quad \textit{where} \quad \operatorname{fix} g := \{ a \in \Omega \mid ga = a \} .$$

Proof. For each orbit of $\mathfrak{G}$, fix one representative $a_i \in \Omega$, $i < \kappa$. It follows that

$$\begin{aligned}\kappa \otimes |G| &= \sum_{i<\kappa} |G| = \sum_{i<\kappa} |G(a_i)| \otimes |G_{(a_i)}| = \sum_{i<\kappa} \sum_{b \in G(a_i)} |G_{(a_i)}| \\ &= \sum_{i<\kappa} \sum_{b \in G(a_i)} |G_{(b)}| = \sum_{b \in \Omega} |G_{(b)}| \\ &= \big|\big\{ \langle g, b \rangle \in G \times \Omega \;\big|\; gb = b \big\}\big| = \sum_{g \in G} |\operatorname{fix} g| .\end{aligned}$$

□

Corollary 2.8. *If $\mathfrak{G}$ is a finite group acting on Ω then the number of orbits is*

$$\frac{1}{|G|} \sum_{g \in G} |\operatorname{fix} g| .$$

Let us collect two combinatorial results about groups and their subgroups.

Lemma 2.9 (B. H. Neumann). *Suppose that $\mathfrak{H}_0, \dots, \mathfrak{H}_{n-1}$ are subgroups of a group $\mathfrak{G}$ and $a_0, \dots, a_{n-1} \in G$ elements such that*

$$G = a_0 H_0 \cup \dots \cup a_{n-1} H_{n-1} .$$

but $\quad G \neq \bigcup_{i \in I} a_i H_i \,, \quad$ *for every proper subset $I \subset [n]$.*

Then $|G / \bigcap_i H_i| \leq n!$. In particular, $|G/H_i|$ is finite for all i.

Proof. Let $\mathfrak{H} := \bigcap_i \mathfrak{H}_i$. We claim that

$$|\textstyle\bigcap_{i \in I} H_i / H| \leq (n - |I|)! \,, \quad \text{for all nonempty } I \subseteq [n] .$$

For $I = \{i\}$, it then follows that every H_i is the union of at most $(n-1)!$ cosets of $\mathfrak{H}$. Hence, G can be written as union of $n!$ such cosets, i.e., $|G/H| \leq n!$.

We prove the above claim by induction on $n - |I|$. For $I = [n]$, we have $|H/H| = 1$. Suppose that $|I| < n$ and set $\mathfrak{F} := \bigcap_{i \in I} \mathfrak{H}_i$. By assumption there is some element $g \in G \setminus \bigcup_{i \in I} a_i H_i$. Hence, for all $i \in I$, we have $a_i H_i \cap g H_i = \varnothing$. This implies that

$$a_i H_i \cap gF = \varnothing \quad \text{and} \quad g^{-1} a_i H_i \cap F = \varnothing .$$

For every $i < n$, we either have

$$g^{-1} a_i H_i \cap F = \varnothing$$

or there is some $h_i \in G$ with

$$g^{-1} a_i H_i \cap F = h_i (F \cap H_i) .$$

For $i \in I$, we have seen that the intersection is empty. Therefore, F is the union of at most $n - |I|$ sets of the form $h_i(F \cap H_i)$ with $i \notin I$. By inductive hypothesis, we can write each of these as union of at most $(n - |I| - 1)!$ cosets of $\mathfrak{H}$. Therefore, $|F/H| \leq (n - |I|)!$. □

Corollary 2.10 (П. M. Neumann). *Let $\mathfrak{M}$ be a Σ-structure and $\bar{a} \in M^{<\omega}$. If no a_i lies in a finite orbit of $\mathfrak{Aut}\,\mathfrak{M}$ then the orbit of $\bar{a}$ under $\mathfrak{Aut}\,\mathfrak{M}$ contains an infinite set of pairwise disjoint tuples.*

Proof. Let $C \subseteq M$ be finite. We claim that there is some $g \in \operatorname{Aut}\mathfrak{M}$ such that $g\bar{a} \cap C = \varnothing$. For a contradiction, suppose otherwise. For every $c \in C$ and each $i < n$, choose, if possible, some element $g_{ic} \in \operatorname{Aut}\mathfrak{M}$ with $g_{ic} a_i = c$. Let $\mathfrak{H}_i := (\mathfrak{Aut}\,\mathfrak{M})_{(a_i)}$. By assumption, every $g \in \operatorname{Aut}\mathfrak{M}$ is contained in some coset $g_{ic} H_i$. Hence, we can apply B. H. Neumann's lemma and it follows that at least one $\mathfrak{H}_i$ has finite index in $\mathfrak{Aut}\,\mathfrak{M}$. Therefore, the orbit of a_i under $\mathfrak{Aut}\,\mathfrak{M}$ is finite. Contradiction. □

When studying group actions it is helpful to introduce a topology on the group.

Definition 2.11. A *topological group* is a group $\mathfrak{G}$ equipped with a topology such that the group multiplication $\cdot : G \times G \to G$ and its inverse ${}^{-1} : G \to G$ are continuous.

Example. The additive group of the real vector space $\mathbb{R}^n$ is topological in the usual topology.

Each action induces a canonical topology on its group.

Definition 2.12. Let $\mathfrak{G}$ be a group acting on Ω. For finite tuples $\bar{a}, \bar{b} \in \Omega^n$, we set

$$\langle \bar{a} \mapsto \bar{b} \rangle := \{ g \in G \mid g\bar{a} = \bar{b} \} .$$

Subsets $O \subseteq G$ of the form $O = \langle \bar{a} \mapsto \bar{b} \rangle$ are called *basic open*.

Lemma 2.13. *Let $\mathfrak{G}$ be a group acting on Ω.*

(a) *The family of all basic open sets induces a topology on $\mathfrak{G}$.*

(b) *$\mathfrak{G}$ equipped with this topology forms a topological group.*

(c) *A subgroup $\mathfrak{H} \subseteq \mathfrak{G}$ is open if and only if there is some finite tuple $\bar{a} \in \Omega^{<\omega}$ with $G_{(\bar{a})} \subseteq H$.*

(d) *A subset $F \subseteq G$ is closed if and only if, whenever $g \in G$ is an element such that, for all finite tuples $\bar{a} \subseteq \Omega$, there is some element $h \in F$ with $g\bar{a} = h\bar{a}$, then we have $g \in F$.*

(e) *A subset $F \subseteq G$ is dense in G if and only if the orbits of G and F on Ω^n are the same, for all $n < \omega$.*

Proof. (a) We have $\langle \bar{a}_0 \mapsto \bar{b}_0 \rangle \cap \langle \bar{a}_1 \mapsto \bar{b}_1 \rangle = \langle \bar{a}_0\bar{a}_1 \mapsto \bar{b}_0\bar{b}_1 \rangle$. Therefore, we only have to show that every $g \in G$ is contained in some basic open set. Fix an arbitrary element $a \in \Omega$ and let $b := ga$. Then $g \in \langle a \mapsto b \rangle$.

(b) If $g \in \langle \bar{a} \mapsto \bar{b} \rangle$ then $g^{-1} \in \langle \bar{b} \mapsto \bar{a} \rangle$. Hence, $^{-1}$ is continuous. Similarly, $gh \in \langle \bar{a} \mapsto \bar{b} \rangle$ implies $g\bar{c} = \bar{b}$ where $\bar{c} := h\bar{a}$. Consequently, we have $g \in \langle \bar{c} \mapsto \bar{b} \rangle$, $h \in \langle \bar{a} \mapsto \bar{c} \rangle$, and $\langle \bar{c} \mapsto \bar{b} \rangle \cdot \langle \bar{a} \mapsto \bar{c} \rangle \subseteq \langle \bar{a} \mapsto \bar{b} \rangle$.

(c) If $G_{(\bar{a})} \subseteq H$ then

$$H = \bigcup_{h \in H} hG_{(\bar{a})} = \bigcup_{h \in H} \langle \bar{a} \mapsto h\bar{a} \rangle$$

is open. Conversely, if H is open then it contains some basic open set $\langle \bar{a} \mapsto \bar{b} \rangle$. Fixing some $h \in \langle \bar{a} \mapsto \bar{b} \rangle \subseteq H$ we have

$$G_{(\bar{a})} = \langle \bar{a} \mapsto \bar{a} \rangle = h^{-1} \langle \bar{a} \mapsto \bar{b} \rangle \subseteq h^{-1} H = H \,.$$

(d) F is closed if and only if it contains all elements $g \in G$ such that

$$F \cap \langle \bar{a} \mapsto \bar{b} \rangle \neq \varnothing \,, \qquad \text{for all basic open set with } g \in \langle \bar{a} \mapsto \bar{b} \rangle \,.$$

This is equivalent to (d).

(e) F is dense if and only if every nonempty basic open set $\langle \bar{a} \mapsto \bar{b} \rangle$ has a nonempty intersection with F. Therefore, F is dense iff, for every $g \in G$ with $g\bar{a} = \bar{b}$, there is some $h \in F$ mapping $\bar{a}$ to $\bar{b}$. □

We can characterise automorphism groups in topological terms.

Lemma 2.14. *Let $\mathfrak{G} \subseteq \mathfrak{Sym}\,\Omega$. A subgroup $\mathfrak{H} \subseteq \mathfrak{G}$ is closed in $\mathfrak{G}$ if and only if there is some structure $\mathfrak{M}$ with universe Ω such that $H = G \cap \mathrm{Aut}\,\mathfrak{M}$.*

In particular, a subgroup $\mathfrak{H} \subseteq \mathfrak{Sym}\,\Omega$ is of the form $\mathfrak{Aut}\,\mathfrak{M}$ if and only if it is closed.

Proof. ($\Rightarrow$) Let $\mathfrak{M}$ be the structure with universe Ω that, for each finite tuple $\bar{s}$ of sorts and every orbit $\Delta \subseteq \Omega^{\bar{s}}$, has a relation $R^{\mathfrak{M}}_{\Delta} := \Delta$ of type $\bar{s}$. Since every element of H maps R_Δ into R_Δ we have $H \subseteq \mathrm{Aut}\,\mathfrak{M}$. Hence, $H \subseteq G$ implies $H \subseteq G \cap \mathrm{Aut}\,\mathfrak{M}$.

For the converse, let $g \in G \cap \mathrm{Aut}\,\mathfrak{M}$. If $\bar{a} \in R_\Delta$ then $g\bar{a} \in R_\Delta$. Hence, there is some $h \in H$ mapping $\bar{a}$ to $g\bar{a}$. Since H is closed in G it follows by Lemma 2.13 (d) that $g \in H$.

($\Leftarrow$) Let $H = G \cap \mathrm{Aut}\,\mathfrak{M}$. To show that H is closed in G we apply Lemma 2.13 (d). Let $g \in G$ and suppose that, for every finite tuple $\bar{a} \in \Omega$, there is some $h \in H$ with $h\bar{a} = g\bar{a}$. Let $\varphi(\bar{x})$ be an atomic formula and $\bar{a} \in \Omega^n$. Choose $h \in H$ such that $h\bar{a} = g\bar{a}$. Since $H \subseteq \mathrm{Aut}\,\mathfrak{M}$ it follows that

$$\mathfrak{M} \vDash \varphi(\bar{a}) \quad \text{iff} \quad \mathfrak{M} \vDash \varphi(h\bar{a}) \quad \text{iff} \quad \mathfrak{M} \vDash \varphi(g\bar{a}) \,.$$

Hence, $g \in \mathrm{Aut}\,\mathfrak{M}$ which implies that $g \in H$. □

Exercise 2.1. Let $\mathfrak{A}$ be a countable structure with countable signature such that

$$|\mathrm{Aut}\langle\mathfrak{A}, \bar{a}\rangle| > 1\,, \quad \text{for all } \bar{a} \in A^{<\omega}\,.$$

Prove that $|\mathrm{Aut}\,\mathfrak{A}| = 2^{\aleph_0}$.

3. *Rings*

Let us consider what happens if we add a second binary operation to an abelian group.

Definition 3.1. (a) A structure $\mathfrak{R} = \langle R, +, -, \cdot, 0, 1\rangle$ is a *ring* if the reduct $\langle R, +, -, 0\rangle$ is an abelian group, $\langle R, \cdot, 1\rangle$ is a monoid, and all elements $a, b, c \in R$ satisfy the following distributive laws:

$$a \cdot (b + c) = a \cdot b + a \cdot c\,,$$
$$(a + b) \cdot c = a \cdot c + b \cdot c\,.$$

Usually we omit the dot and write ab instead of $a \cdot b$.

(b) A ring $\mathfrak{R}$ is *commutative* if we further have

$$a \cdot b = b \cdot a\,, \quad \text{for all } a, b \in R\,.$$

(c) A ring $\mathfrak{R}$ is a *skew field* if $0 \neq 1$ and, for every $a \in R$ with $a \neq 0$, there is some element $a^{-1} \in R$ such that

$$a \cdot a^{-1} = 1 = a^{-1} \cdot a\,.$$

A commutative skew field is called a *field.*

Example. (a) The integers $\langle \mathbb{Z}, +, -, \cdot, 0, 1\rangle$ form a commutative ring.

(b) The rationals $\langle \mathbb{Q}, +, -, \cdot, 0, 1\rangle$ form a field.

(c) Let $\mathfrak{V}$ be a vector space. The set $\mathrm{Lin}(\mathfrak{V}, \mathfrak{V})$ of all linear maps $h : \mathfrak{V} \to \mathfrak{V}$ forms a ring where addition is defined component wise:

$$(g + h)(x) := g(x) + h(x)\,,$$

and multiplication is composition:

$$(g \cdot h)(x) := g(h(x)) .$$

This ring is not commutative.

An important example of rings are polynomial rings. Here we present only their basic properties. In Section 5 we will study polynomial rings over a field in more detail.

Definition 3.2. Let $\mathfrak{R}$ be a ring.

(a) The ring $\mathfrak{R}[[x]]$ of *formal power series* over $\mathfrak{R}$ has the universe

$$R[[x]] := R^\omega .$$

For $s, t \in R[[x]]$, we define addition and multiplication by

$$(s + t)(n) := s(n) + t(n) \quad \text{and} \quad (s \cdot t)(n) := \sum_{i=0}^{n} s(i)t(n - i) .$$

We also define a *derivation* operation on $\mathfrak{R}[[x]]$ by

$$s'(n) := (n + 1)s(n + 1) .$$

Usually, elements $s \in R[[x]]$ are written more suggestively in the form

$$s = \sum_{n<\omega} a_n x^n \quad \text{where} \quad a_n := s(n) .$$

The numbers a_n are called the *coefficients* of s. In this notation the above definitions take the following form:

$$\sum_{n<\omega} a_n x^n + \sum_{n<\omega} b_n x^n := \sum_{n<\omega} (a_n + b_n) x^n ,$$

$$\sum_{n<\omega} a_n x^n \cdot \sum_{n<\omega} b_n x^n := \sum_{n<\omega} \Big(\sum_{i=0}^{n} a_i b_{n-i} \Big) x^n ,$$

$$\Big(\sum_{n<\omega} a_n x^n \Big)' := \sum_{n<\omega} a_{n+1} (n + 1) x^n .$$

(b) The *polynomial ring* over $\mathfrak{R}$ is the subring $\mathfrak{R}[x] \subseteq \mathfrak{R}[[x]]$ of all formal power series $\sum_{n<\omega} a_n x^n$ where $a_n = 0$ for all but finitely many n. Elements $p \in R[x]$ are called *polynomials*. Omitting zero terms we can write them as finite sums

$$p = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 ,$$

where $a_i := p(i)$ and n is an arbitrary number such that $p(i) = 0$, for $i > n$.

(c) The *degree* of a polynomial $\sum_i a_i x^i \in R[x]$ is the largest number n with $a_n \neq 0$. We denote it by $\deg p$. If all coefficients a_i are equal to 0 then we set $\deg p := -\infty$.

(d) We can iterate the construction of polynomial rings to obtain rings $R[x_0, x_1, \dots, x_{n-1}] := R[x_0][x_1] \dots [x_{n-1}]$.

Remark. Let $\mathfrak{Ring}$ be the category of all rings with homomorphisms. We can turn the operation $\mathfrak{R} \mapsto \mathfrak{R}[x]$ into a functor $F : \mathfrak{Ring} \to \mathfrak{Ring}$ if, for homomorphisms $h : \mathfrak{R} \to \mathfrak{S}$, we define

$$F(h)(\textstyle\sum_n a_n x^n) := \sum_n h(a_n) x^n .$$

Remark. Let R be a commutative ring and $p, q \in R[x]$. A direct calculation shows that we have

$$(p + q)' = p' + q' \quad \text{and} \quad (pq)' = pq' + p'q .$$

Polynomial rings can be regarded as a free extension of a ring by a single new element x.

Lemma 3.3. *Let $\mathfrak{R}$ and $\mathfrak{S}$ be rings. For each homomorphism $h_0 : \mathfrak{R} \to \mathfrak{S}$ and every element $a \in S$, there exists a unique homomorphism $h : \mathfrak{R}[x] \to \mathfrak{S}$ with $h(x) = a$ and $h \restriction R = h_0$.*

Proof. For $p = c_n x^n + \cdots + c_1 x + c_0$, we define

$$h(p) := h_0(c_n) a^n + \cdots + h_0(c_1) a + h_0(c_0) .$$

It is straightforward to check that h is a homomorphism. For uniqueness, suppose that g is another homomorphism such that $g(x) = a$ and $g \upharpoonright R = h_0$. For every polynomial $p = c_n x^n + \dots + c_1 x + c_0$, we have

$$\begin{aligned} g(p) &= g(c_n)g(x)^n + \dots + g(c_1)g(x) + g(c_0) \\ &= h_0(c_n)a^n + \dots + h_0(c_1)a + h_0(c_0) = h(p)\,. \end{aligned}$$

Hence, $g = h$. □

As for groups we can characterise congruences of rings in terms of certain subrings.

Definition 3.4. Let $\mathfrak{R}$ be a ring.

(a) A *left ideal* of $\mathfrak{R}$ is a subset $\mathfrak{a} \subseteq R$ such that

$$\begin{aligned} a + b &\in \mathfrak{a}\,, &&\text{for all } a, b \in \mathfrak{a}\,, \\ ra &\in \mathfrak{a}\,, &&\text{for all } a \in \mathfrak{a} \text{ and every } r \in R\,. \end{aligned}$$

(b) A (*two-sided*) *ideal* of $\mathfrak{R}$ is a subset $\mathfrak{a} \subseteq R$ such that

$$\begin{aligned} a + b &\in \mathfrak{a}\,, &&\text{for all } a, b \in \mathfrak{a}\,, \\ ras &\in \mathfrak{a}\,, &&\text{for all } a \in \mathfrak{a} \text{ and all } r, s \in R\,. \end{aligned}$$

(c) We denote the set of all ideals of $\mathfrak{R}$ ordered by inclusion by

$$\mathfrak{Idl}(\mathfrak{R}) := \langle \mathrm{Idl}(\mathfrak{R}), \subseteq \rangle\,.$$

(d) Let $\bar{a} \subseteq R$. The ideal *generated* by $\bar{a}$ is

$$(\bar{a}) := \bigcap \{\, \mathfrak{a} \subseteq R \mid \mathfrak{a} \text{ an ideal with } \bar{a} \subseteq \mathfrak{a} \,\}\,.$$

Remark. Clearly, every two-sided ideal is also a left ideal. The converse does not hold in general, but for commutative rings both notions coincide.

Example. Let $\mathfrak{Z} = \langle \mathbb{Z}, +, -, \cdot, 0, 1 \rangle$ be the ring of integers. A subset $\mathfrak{a} \subseteq \mathbb{Z}$ is an ideal if and only if it is of the form $m\mathbb{Z}$, for some $m \in \mathbb{N}$.

Exercise 3.1. Prove that

$$(a_0, \dots, a_{n-1}) = \{ r_0 a_0 s_0 + \dots + r_{n-1} a_{n-1} s_{n-1} \mid \bar{r}, \bar{s} \subseteq R \} .$$

Lemma 3.5. *Let $\mathfrak{R}$ be a ring.*

(a) *If $h : \mathfrak{R} \to \mathfrak{S}$ is a surjective homomorphism then $\mathfrak{S}$ is also a ring.*

(b) *If $h : \mathfrak{R} \to \mathfrak{S}$ is a homomorphism into a ring $\mathfrak{S}$, then $h^{-1}[0]$ is an ideal of $\mathfrak{R}$.*

(c) *If $\mathfrak{a}$ is an ideal of $\mathfrak{R}$, then the relation*

$$r \approx_{\mathfrak{a}} s \quad :\text{iff} \quad r - s \in \mathfrak{a}$$

is a congruence of $\mathfrak{R}$.

Proof. (a) For all elements $a, b, c \in S$, there are elements $x \in h^{-1}(a)$, $y \in h^{-1}(b)$, and $z \in h^{-1}(c)$. Since h is a homomorphism it follows that every equation satisfied by x, y, and z is also satisfied by a, b, and c.

(b) Let $a, b \in h^{-1}[0]$ and $r, s \in R$. Then

$$h(a + b) = h(a) + h(b) = 0 + 0 = 0 ,$$

and $$h(ras) = h(r) \cdot h(a) \cdot h(s) = h(r) \cdot 0 \cdot h(s) = 0 .$$

(c) First, we prove that $\approx_{\mathfrak{a}}$ is an equivalence relation. Let $r, s, t \in R$. The relation $\approx_{\mathfrak{a}}$ is reflexive since $r - r = 0 \in \mathfrak{a}$. It is symmetric since $r - s \in \mathfrak{a}$ implies $s - r = (-1) \cdot (r - s) \in \mathfrak{a}$. Finally, it is transitive since $r - s,\ s - t \in \mathfrak{a}$ implies $r - t = (r - s) + (s - t) \in \mathfrak{a}$.

It remains to show that $\approx_{\mathfrak{a}}$ is a congruence. Suppose that $r \approx_{\mathfrak{a}} r'$ and $s \approx_{\mathfrak{a}} s'$. Then

$$(r + s) - (r' + s') = (r - r') + (s - s') \in \mathfrak{a} ,$$

and $$rs - r's' = rs - rs' + rs' - r's' = r(s - s') + (r - r')s' \in \mathfrak{a} .$$ □

Theorem 3.6. *Let $\mathfrak{R}$ be a ring. The function $\mathfrak{Idl}(\mathfrak{R}) \to \mathfrak{Cong}(\mathfrak{R}) : \mathfrak{a} \mapsto \approx_{\mathfrak{a}}$ is an isomorphism.*

Proof. By definition, $\mathfrak{a} \subseteq \mathfrak{b}$ implies $\approx_{\mathfrak{a}} \subseteq \approx_{\mathfrak{b}}$. Hence, $h : \mathfrak{a} \mapsto \approx_{\mathfrak{a}}$ is a homomorphism and it remains to find a homomorphism $g : \mathfrak{Cong}(\mathfrak{R}) \to \mathfrak{Idl}(\mathfrak{R})$ that is inverse to h. For $\sim \in \mathrm{Cong}(\mathfrak{R})$, we define

$$g(\sim) := [0]_{\sim} .$$

Then $\sim \subseteq \approx$ implies $g(\sim) \subseteq g(\approx)$. Furthermore,

$$g(h(\mathfrak{a})) = g(\approx_{\mathfrak{a}}) = [0]_{\approx_{\mathfrak{a}}} = \mathfrak{a} ,$$

and $\quad h(g(\sim)) = h([0]_{\sim}) = \approx_{[0]_{\sim}} = \sim .$ ☐

Definition 3.7. Let $\mathfrak{R}$ be a ring.
(a) For an ideal $\mathfrak{a}$ of $\mathfrak{R}$, we set

$$\mathfrak{R}/\mathfrak{a} := \mathfrak{R}/\approx_{\mathfrak{a}} .$$

(b) The *kernel* of a homomorphism $h : \mathfrak{R} \to \mathfrak{S}$ is the ideal

$$\mathrm{Ker}\, h := h^{-1}[0] \quad (= [0]_{\ker h}) .$$

To every ring we can assign a topological space in much the same way as we associated Stone spaces with boolean algebras.

Definition 3.8. Let $\mathfrak{R}$ be a ring.
(a) An ideal $\mathfrak{p}$ of $\mathfrak{R}$ is *prime* if $\mathfrak{p} \neq R$ and

$$ab \in \mathfrak{p} \quad \text{implies} \quad a \in \mathfrak{p} \text{ or } b \in \mathfrak{p} , \quad \text{for all } a, b \in R .$$

(b) The *spectrum* of $\mathfrak{R}$ is the set $\mathrm{spec}(\mathfrak{R})$ of all prime ideals. We endow $\mathrm{spec}(\mathfrak{R})$ with a topology by taking as closed sets all sets of the form

$$\langle X \rangle := \{ \mathfrak{p} \in \mathrm{spec}(\mathfrak{R}) \mid X \subseteq \mathfrak{p} \} , \quad \text{for } X \subseteq R .$$

Exercise 3.2. Prove that $\mathrm{spec} : \mathfrak{Ring} \to \mathfrak{Top}$ is a contravariant functor.

4. Modules

Instead of a group acting on a set we can consider a ring acting on an abelian group. This leads to the notion of a module.

Definition 4.1. Let $\mathfrak{R}$ be a ring.

(a) An $\mathfrak{R}$-*module* $\mathfrak{M}$ consists of an abelian group $\mathfrak{M} = \langle M, +, -, 0\rangle$ and an action $R \times M \to M$ satisfying

$$\begin{aligned} r(sa) &= (rs)a\,, \\ r(a+b) &= ra + rb\,, \qquad \text{for all } r, s \in R \text{ and } a, b \in M\,. \\ (r+s)a &= ra + sa\,, \end{aligned}$$

The action $R \times M \to M$ is called *scalar multiplication.*

(b) A *vector space* is an $\mathfrak{R}$-module where the ring $\mathfrak{R}$ is a skew field.

(c) We regard $\mathfrak{R}$-modules as one-sorted structures

$$\mathfrak{M} = \langle M, +, -, 0, (\lambda_r)_{r\in R}\rangle$$

where $\lambda_r : a \mapsto ra$ are the scalar multiplication maps. When we talk about substructures or homomorphisms of modules we always have this signature in mind.

(d) We denote by $\mathfrak{Mod}_{\mathfrak{R}}$ the category of all $\mathfrak{R}$-modules and homomorphisms.

Example. (a) We can turn every abelian group $\mathfrak{A}$ into a $\mathbb{Z}$-module by defining

$$\begin{aligned} 0a &:= 0\,, \\ (n+1)a &:= na + a\,, \qquad \text{for } n \in \mathbb{N} \text{ and } a \in A\,. \\ (-n)a &:= -(na)\,, \end{aligned}$$

(b) Every ring $\mathfrak{R}$ is an $\mathfrak{R}$-module for the canonical action $\alpha(r)(a) := ra$ given by multiplication.

(c) The derivation map $\mathfrak{R}[x] \to \mathfrak{R}[x] : p \mapsto p'$ is a homomorphism of $\mathfrak{R}$-modules. It is not a ring homomorphism.

We can turn the set of all homomorphisms $\mathfrak{M} \to \mathfrak{N}$ into an $\mathfrak{R}$-module by defining addition and scalar multiplication pointwise.

Exercise 4.1. If $\mathfrak{M}$ and $\mathfrak{N}$ are $\mathfrak{R}$-modules then so is $\mathrm{Mod}_{\mathfrak{R}}(\mathfrak{M}, \mathfrak{N})$.

For $\mathfrak{N} = \mathfrak{M}$ we not only get a module but even a ring.

Definition 4.2. The *endomorphism ring* $\mathfrak{End}(\mathfrak{M})$ of an $\mathfrak{R}$-module $\mathfrak{M}$ is the ring with universe

$$\mathrm{End}(\mathfrak{M}) := \mathrm{Mod}_{\mathfrak{R}}(\mathfrak{M}, \mathfrak{M})$$

where addition and multiplication are defined by

$$(g + h)(x) := g(x) + h(x) \quad \text{and} \quad (g \cdot h)(x) := g(h(x)) .$$

Lemma 4.3. $\mathfrak{End}(\mathfrak{M})$ *is a ring.*

Exercise 4.2. Prove the lemma.

We have seen above that congruences of groups and rings can be described in terms on certain substructures. For modules, the situation is much simpler. Every submodule corresponds to a congruence.

Theorem 4.4. *Let $\mathfrak{M}$ be an $\mathfrak{R}$-module. The function*

$$\mathfrak{Sub}(\mathfrak{M}) \to \mathfrak{Cong}(\mathfrak{M}) : U \mapsto \{ \langle a, b \rangle \mid a - b \in U \}$$

is an isomorphism. Its inverse is given by the map $\approx \, \mapsto [0]_{\approx}$.

Exercise 4.3. Prove the preceding theorem.

Lemma 4.5. *Let $\mathfrak{M}$ be an $\mathfrak{R}$-module. Then $\mathfrak{Sub}(\mathfrak{M})$ is a modular lattice.*

Proof. Let $\mathfrak{K}, \mathfrak{L} \subseteq \mathfrak{M}$. It is straightforward to check that

$$\mathfrak{K} \sqcap \mathfrak{L} = \mathfrak{K} \cap \mathfrak{L} := \mathfrak{M}|_{K \cap L} \quad \text{and} \quad \mathfrak{K} \sqcup \mathfrak{L} = \mathfrak{K} + \mathfrak{L} := \mathfrak{M}|_{K+L} .$$

Hence, $\mathfrak{Sub}(\mathfrak{M})$ is a lattice. To show that it is modular it is sufficient to prove that

$$\mathfrak{K} \subseteq \mathfrak{L} \quad \text{implies} \quad \mathfrak{L} \cap (\mathfrak{K} + \mathfrak{N}) \subseteq \mathfrak{K} + (\mathfrak{L} \cap \mathfrak{N}) .$$

Let $a \in L \cap (K + N)$. Then there are elements $b \in K$ and $c \in N$ such that $a = b + c$. Since $a \in L$ and $b \in K \subseteq L$ it follows that $c = a - b \in L$. Hence, $c \in L \cap N$ and we have $a = b + c \in K + (L \cap N)$. □

Since congruences of modules are simpler than those of rings, it is frequently worthwhile to regard rings as modules. The following observation shows that we can study the left ideals of a ring in this way. For the proof, it is sufficient to note that the closure conditions of a left ideal and those of a submodule coincide.

Lemma 4.6. *Let $\mathfrak{R}$ be a ring. A subset $\mathfrak{a} \subseteq R$ is a left ideal of $\mathfrak{R}$ if and only if it is a submodule of $\mathfrak{R}$.*

Let us consider products of modules. We will show below that we can decompose every vector space over a skew field $\mathfrak{S}$ as a product of copies of $\mathfrak{S}$.

Lemma 4.7. *If $\mathfrak{M}_i$, for $i \in I$, are $\mathfrak{R}$-modules then so is their direct product $\prod_{i \in I} \mathfrak{M}_i$.*

Definition 4.8. Let $(\mathfrak{M}_i)_{i \in I}$ be a family of $\mathfrak{R}$-modules. The *direct sum* $\bigoplus_{i \in I} \mathfrak{M}_i$ is the submodule of $\prod_{i \in I} \mathfrak{M}_i$ consisting of all sequence $a \in \prod_i M_i$ such that $a(i) = 0$, for all but finitely many i.

The *direct power* of a module $\mathfrak{M}$ is the direct sum $\mathfrak{M}^{(I)} := \bigoplus_{i \in I} \mathfrak{M}$ of I copies of $\mathfrak{M}$.

Remark. In the category $\mathfrak{Mod}_{\mathfrak{R}}$ the direct product $\prod_i \mathfrak{M}_i$ and the direct sum $\bigoplus_i \mathfrak{M}_i$ play the role of, respectively, product and coproduct.

That is, for every family of homomorphisms $h_i : \mathfrak{N} \to \mathfrak{M}_i$, $i \in I$, there is a unique homomorphism $g : \mathfrak{N} \to \prod_i \mathfrak{M}_i$ such that $h_i = \mathrm{pr}_i \circ g$ where $\mathrm{pi}_i : \prod_j \mathfrak{M}_j \to \mathfrak{M}_i$ is the i-th projection.

Similarly, for every family of homomorphisms $h_i : \mathfrak{M}_i \to \mathfrak{N}$, $i \in I$, there is a unique homomorphism $g : \bigoplus_i \mathfrak{M}_i \to \mathfrak{N}$ such that $h_i = g \circ \mathrm{in}_i$ where $\mathrm{in}_i : \mathfrak{M}_i \to \bigoplus_j \mathfrak{M}_j$ is the i-th injection.

To conclude this section we take a look at the structure of vector spaces, which is particularly simple. We will show below that every vector space over a skew field $\mathfrak{S}$ is isomorphic to a direct power of $\mathfrak{S}$.

Definition 4.9. Let $\mathfrak{V}$ be a vector space over a skew field $\mathfrak{S}$.

(a) A set $X \subseteq V$ is *linearly dependent* if there are pairwise distinct elements $a_0, \dots, a_{n-1} \in X$ and nonzero scalars $s_0, \dots, s_{n-1} \in S \setminus \{0\}$, such that

$$s_0 a_0 + \dots + s_{n-1} a_{n-1} = 0 .$$

Otherwise, X is called *linearly independent.*

(b) A *basis* of $\mathfrak{V}$ is a linearly independent subset $B \subseteq V$ generating $\mathfrak{V}$.

Lemma 4.10. *Let $\mathfrak{V}$ be a vector space over a skew field $\mathfrak{S}$, $a \in V$, and suppose that $I \subseteq V$ is linearly independent. Then $I \cup \{a\}$ is linearly independent if and only if $a \notin \langle\!\langle I \rangle\!\rangle_{\mathfrak{V}}$.*

Proof. ($\Rightarrow$) If $a \in \langle\!\langle I \rangle\!\rangle_{\mathfrak{V}}$ then there are elements $b_0, \dots, b_{n-1} \in I$ and scalars $s_0, \dots, s_{n-1} \in S$ such that

$$a = s_0 b_0 + \dots + s_{n-1} b_{n-1} .$$

Omitting all terms $s_i b_i$ that are zero, we may assume that $s_i \neq 0$, for all i. Consequently,

$$s_0 b_0 + \dots + s_{n-1} b_{n-1} - a = 0$$

and $I \cup \{a\}$ is linearly dependent.

($\Leftarrow$) Suppose that $I \cup \{a\}$ is linearly dependent. Then there are elements $b_0, \dots, b_{n-1} \in I$ and nonzero scalars $r, s_0, \dots, s_{n-1} \in S$ such that

$$ra + s_0 b_0 + \dots + s_{n-1} b_{n-1} = 0 .$$

(This sum must contain a term with a since I is independent.) Consequently,

$$a = -r^{-1}s_0 b_0 - \cdots - r^{-1}s_{n-1}b_{n-1} \in \langle\!\langle I \rangle\!\rangle_{\mathfrak{V}} . \qquad \square$$

Lemma 4.11. *Every vector space has a basis.*

Proof. Suppose that $\mathfrak{V}$ is a vector space over $\mathfrak{S}$. Let $\mathcal{I}$ be the set of all linearly independent sets $I \subseteq V$. The partial order $\langle \mathcal{I}, \subseteq \rangle$ is inductive. Consequently, it has a maximal element B. We claim that B is a basis. Suppose otherwise. Then there is some vector $a \in V \setminus \langle\!\langle B \rangle\!\rangle_{\mathfrak{V}}$. By Lemma 4.10, it follows that $B \cup \{a\}$ is linearly independent. This contradicts the maximality of B. $\square$

Theorem 4.12. *Let $\mathfrak{V}$ be an $\mathfrak{S}$-vector space with basis B. There exists an isomorphism*

$$h : \mathfrak{S}^{(B)} \to \mathfrak{V} : (s_b)_{b \in B} \mapsto \sum_{b \in B} s_b b .$$

Proof. It is straightforward to check that h is a homomorphism. We claim that it is bijective. For surjectivity, fix $a \in V$. Since $V = \langle\!\langle B \rangle\!\rangle_{\mathfrak{V}}$ there are elements $b_0, \dots, b_{n-1} \in B$ and scalars $s_0, \dots, s_{n-1} \in S$ such that

$$a = s_0 b_0 + \cdots + s_{n-1} b_{n-1} .$$

Hence, $a \in \operatorname{rng} h$.

It remains to prove that h is injective. Suppose that $h(s_b)_b = h(s'_b)_b$. We have

$$\sum_{b \in B} (s_b - s'_b) b = \sum_{b \in B} s_b b - \sum_{b \in B} s'_b b = h(s_b)_b - h(s'_b)_b = 0 .$$

(Note that these sums are defined since $(s_b)_b, (s'_b)_b \in S^{(B)}$.) Since B is linearly independent it follows that $s_b - s'_b = 0$, for all b. Consequently, $(s_b)_b = (s'_b)_b$. $\square$

Every vector space is freely generated by its basis.

Lemma 4.13. *Let $\mathfrak{V}$ and $\mathfrak{W}$ be $\mathfrak{S}$-vector spaces and suppose that B is a basis of $\mathfrak{V}$. For every map $h_0 : B \to W$, there exists a unique homomorphism $h : \mathfrak{V} \to \mathfrak{W}$ such that $h \restriction B = h_0$.*

Proof. By Theorem 4.12, we can find, for every $a \in V$, a unique sequence $(s_b)_b \in S^{(B)}$ such that $a = \sum_b s_b b$. We define $h(a) := \sum_b s_b h_0(b)$.

Then $h \restriction B = h_0$ and we have

$$h(a+b) = h(a) + h(b) \quad \text{and} \quad h(sa) = sh(a)\,.$$

Hence, h is a homomorphism. It is obviously unique. □

Lemma 4.14 (Exchange Lemma)**.** *Let $\mathfrak{V}$ be a vector space over a skew field $\mathfrak{S}$, suppose that $I \subseteq V$ is linearly independent, and let $I_0 \subseteq I$. For every element $a \in \langle\!\langle I \rangle\!\rangle_{\mathfrak{V}} \setminus \langle\!\langle I_0 \rangle\!\rangle_{\mathfrak{V}}$, there exists ane element $b \in I \setminus I_0$ such that $(I \setminus \{b\}) \cup \{a\}$ is linearly independent and $b \in \langle\!\langle (I \setminus \{b\}) \cup \{a\} \rangle\!\rangle_{\mathfrak{V}}$.*

Proof. Since $I \cup \{a\}$ is dependent it follows by Lemma 4.10 that there are elements $b_0, \dots, b_{n-1} \in I$ and scalars $s_0, \dots, s_{n-1} \in S$ such that

$$a = s_0 b_0 + \dots + s_{n-1} b_{n-1}\,.$$

We choose these elements such that the number n is minimal. It particular this implies that $s_i \neq 0$, for all i.

Since the set $I_0 \cup \{a\}$ is independent we have $b_i \in I \setminus I_0$, for some i. By renumbering the elements we may assume that $b_0 \in I \setminus I_0$. We claim that b_0 is the desired element.

First of all,

$$b_0 = s_0^{-1} a - s_0^{-1} s_1 b_1 - \dots - s_0^{-1} s_{n-1} b_{n-1}$$

implies that $b_0 \in \langle\!\langle (I \setminus b_0) \cup \{a\} \rangle\!\rangle_{\mathfrak{V}}$. Hence, it remains to prove that $(I \setminus b_0) \cup \{a\}$ is linearly independent.

For a contradiction, suppose otherwise. Then Lemma 4.10 implies that $a \in \langle\!\langle I \smallsetminus \{b_0\} \rangle\!\rangle_{\mathfrak{V}}$. Since $\langle\!\langle \cdot \rangle\!\rangle_{\mathfrak{V}}$ is a closure operator it follows that

$$b_0 \in \langle\!\langle (I \smallsetminus \{b_0\}) \cup \{a\} \rangle\!\rangle_{\mathfrak{V}} \subseteq \langle\!\langle \langle\!\langle I \smallsetminus \{b_0\} \rangle\!\rangle_{\mathfrak{V}} \rangle\!\rangle_{\mathfrak{V}} = \langle\!\langle I \smallsetminus \{b_0\} \rangle\!\rangle_{\mathfrak{V}} .$$

Hence, $I = (I \smallsetminus \{b_0\}) \cup \{b_0\}$ is linearly dependent. Contradiction. □

Theorem 4.15. *Let $\mathfrak{V}$ be a vector space over the skew field $\mathfrak{S}$. If $\mathfrak{V}$ has a finite basis then all bases of $\mathfrak{V}$ have the same cardinality.*

Proof. Let B and C be two bases of $\mathfrak{V}$ and suppose that B is finite. We prove by induction on $|B \smallsetminus C|$ that $|B| = |C|$.

First, suppose that $B \subseteq C$. If there is some element $c \in C \smallsetminus B$ then $B \cup \{c\}$ is linearly independent. By Lemma 4.10, it follows that $c \notin \langle\!\langle B \rangle\!\rangle_{\mathfrak{V}} = V$. A contradiction. Consequently, $C = B$.

For the inductive step, suppose that there is some element $b \in B \smallsetminus C$. Let $I := B \cap C$. By Lemma 4.14, we can find a vector $c \in C \smallsetminus I$ such that $C' := (C \smallsetminus \{c\}) \cup \{b\}$ is linearly independent and $\langle\!\langle C' \rangle\!\rangle_{\mathfrak{V}} = \langle\!\langle C \rangle\!\rangle_{\mathfrak{V}} = V$. Hence, C' is a basis of $\mathfrak{V}$ and it follows by inductive hypothesis that $|C| = |C'| = |B|$. □

Remark. The preceding theorem holds also for vector spaces with infinite bases. We postpone the proof to Section F1.1 where we will prove the corresponding result in a more general setting.

Definition 4.16. Let $\mathfrak{V}$ be a vector space. The *dimension* $\dim \mathfrak{V}$ of $\mathfrak{V}$ is the minimal cardinality of a basis of $\mathfrak{V}$.

Theorem 4.17. *Let $\mathfrak{V}$ and $\mathfrak{W}$ be $\mathfrak{S}$-vector spaces. Then $\mathfrak{V} \cong \mathfrak{W}$ if and only if* $\dim \mathfrak{V} = \dim \mathfrak{W}$.

Proof. $(\Rightarrow)$ is trivial. For $(\Leftarrow)$, suppose that B and C are bases of, respectively, $\mathfrak{V}$ and $\mathfrak{W}$ such that $|B| = |C|$. Then $\mathfrak{V} \cong \mathfrak{S}^{(B)} \cong \mathfrak{S}^{(C)} \cong \mathfrak{W}$. □

Lemma 4.18. *Let $\mathfrak{V}$ be a vector space and $n < \omega$. Then we have* $\dim \mathfrak{V} \geq n$ *if and only if there exists a strictly increasing chain*

$$\{0\} = \mathfrak{U}_0 \subset \dots \subset \mathfrak{U}_n = \mathfrak{V}$$

of subspaces of $\mathfrak{V}$.

Proof. ($\Rightarrow$) Let B be a basis of $\mathfrak{V}$. By assumption, $|B| \geq n$. Choose n distinct elements $b_0, \dots, b_{n-1} \in B$ and set

$$\mathfrak{U}_k := \langle\!\langle b_0, \dots, b_{k-1} \rangle\!\rangle_{\mathfrak{V}} .$$

We claim that $\mathfrak{U}_0 \subset \dots \subset \mathfrak{U}_n$. For a contradiction, suppose that $\mathfrak{U}_{k+1} = \mathfrak{U}_k$, for some k. Then

$$b_k \in \mathfrak{U}_k = \langle\!\langle b_0, \dots, b_{k-1} \rangle\!\rangle_{\mathfrak{V}} .$$

By Lemma 4.10 it follows that $\{b_0, \dots, b_{k-1}, b_k\}$ is linearly dependent. Contradiction.

($\Leftarrow$) Suppose that $\{0\} = \mathfrak{U}_0 \subset \dots \subset \mathfrak{U}_n = \mathfrak{V}$. For every $k < n$, choose some element $b_k \in U_{k+1} \setminus U_k$. Let m be the maximal number such that the set $\{b_0, \dots, b_{m-1}\}$ is linearly independent. Since $m \leq \dim \mathfrak{V}$ it is sufficient to prove that $m = n$.

For a contradiction, suppose otherwise. Then $\{b_0, \dots, b_{m-1}, b_m\}$ is linearly dependent and, by Lemma 4.10, it follows that

$$b_m \in \langle\!\langle b_0, \dots, b_{m-1} \rangle\!\rangle_{\mathfrak{V}} \subseteq \mathfrak{U}_m .$$

But $b_m \in U_{m+1} \setminus U_m$. Contradiction. □

5. Fields

We have seen in the previous section that modules over fields are better behaved than modules over arbitrary rings. In this section we study further properties particular to fields. The first and largest part of the section is devoted to constructions turning rings into fields. In particular, we will study quotients of polynomial rings. In the second part we use this machinery to investigate extensions of fields.

Definition 5.1. Let $\mathfrak{R}$ be a ring.

(a) An ideal $\mathfrak{a} \subseteq R$ is *maximal* if $\mathfrak{a} \neq R$ and there is no ideal $\mathfrak{b}$ with $\mathfrak{a} \subset \mathfrak{b} \subset R$.

(b) An element $a \in R$ is a *unit* if there is some $b \in R$ such that $ab = 1 = ba$.

(c) An element $a \in R$ is a *zero-divisor* if $a \neq 0$ and there exists some element $b \neq 0$ such that $ab = 0$ or $ba = 0$.

(d) $\mathfrak{R}$ is an *integral domain* if it is commutative and it contains no zero-divisors.

Remark. (a) Every field is an integral domain. (b) A zero-divisor is never a unit. (c) A ring is a skew field if and only if every element but 0 is a unit.

Exercise 5.1. Let $\mathfrak{R}$ and $\mathfrak{S}$ be commutative rings. Show that the direct product $\mathfrak{R} \times \mathfrak{S}$ is never an integral domain.

Exercise 5.2. Prove that every maximal ideal is prime.

In the same way as $\mathbb{Q}$ is obtained from $\mathbb{Z}$, we can associate a field with every integral domain.

Definition 5.2. Let $\mathfrak{R}$ be an integral domain. The *field of fractions* of $\mathfrak{R}$ is the ring $\mathrm{FF}(\mathfrak{R})$ consisting of all pairs $\langle r, s\rangle \in R^2$ with $s \neq 0$. We write such pairs as fractions r/s.

Two fractions r/s and r'/s' are considered to be equal if $rs' = r's$. Addition and multiplication is defined by the usual formulae

$$r/s + r'/s' := (rs' + r's)/ss' \quad \text{and} \quad r/s \cdot r'/s' := rr'/ss' .$$

Lemma 5.3. *Let $\mathfrak{R}$ be an integral domain. Then $\mathrm{FF}(\mathfrak{R})$ is a field.*

Exercise 5.3. Prove the preceding lemma.

Lemma 5.4. *Let $\mathfrak{R}$ be an integral domain and $\mathfrak{K}$ a field. For every embedding $h_0 : \mathfrak{R} \to \mathfrak{K}$, there exists a unique embedding $h : \mathrm{FF}(\mathfrak{R}) \to \mathfrak{K}$ with $h \restriction R = h_0$.*

Proof. We define $h(r/s) := h_0(r) \cdot h_0(s)^{-1}$. It is straightforward to check that h is an embedding and that this is the only possible choice to define h. □

Theorem 5.5. *A ring $\mathfrak{R}$ is an integral domain if and only if $\mathfrak{R}$ can be embedded into some field $\mathfrak{K}$.*

Proof. Every integral domain $\mathfrak{R}$ can be embedded into the field $\mathrm{FF}(\mathfrak{R})$. Conversely, suppose that $\mathfrak{R} \subseteq \mathfrak{K}$, for some field $\mathfrak{K}$. Since $\mathfrak{K}$ is an integral domain, so is $\mathfrak{R}$. □

We can construct integral domains by taking quotients by prime ideals.

Lemma 5.6. *Let $\mathfrak{R}$ be a commutative ring and $\mathfrak{a} \subseteq R$ an ideal. The quotient $\mathfrak{R}/\mathfrak{a}$ is an integral domain if and only if $\mathfrak{a}$ is prime.*

Proof. Let $\pi : \mathfrak{R} \to \mathfrak{R}/\mathfrak{a}$ be the canonical projection.

($\Rightarrow$) To show that $\mathfrak{a}$ is prime consider elements $a, b \in R$ with $ab \in \mathfrak{a}$. Then $\pi(ab) = 0$. Since $\mathfrak{R}/\mathfrak{a}$ is an integral domain it follows that $\pi(a) = 0$ or $\pi(b) = 0$. Hence, $a \in \mathfrak{a}$ or $b \in \mathfrak{a}$.

($\Leftarrow$) Suppose that $\pi(a)\pi(b) = 0$. Then $ab \in \mathfrak{p}$. Since $\mathfrak{p}$ is prime it follows that $a \in \mathfrak{p}$ or $b \in \mathfrak{p}$. Hence, $\pi(a) = 0$ or $\pi(b) = 0$. □

In a similar way we can characterise ideals $\mathfrak{a}$ such that $\mathfrak{R}/\mathfrak{a}$ is a field.

Definition 5.7. A structure $\mathfrak{A}$ is *simple* if $\mathrm{Cong}_w(\mathfrak{A}) = \{\bot, \top\}$.

Example. A ring $\mathfrak{R}$ is simple if and only if $\{0\}$ and R are its only ideals.

Exercise 5.4. Let $\mathfrak{R}$ be a ring. Prove that an ideal $\mathfrak{m}$ of $\mathfrak{R}$ is maximal if and only if the quotient $\mathfrak{R}/\mathfrak{m}$ is simple.

Lemma 5.8. *A commutative ring $\mathfrak{R}$ is a field if and only if it is simple.*

Proof. ($\Rightarrow$) Let $\mathfrak{R}$ be a field and $\mathfrak{a}$ an ideal of $\mathfrak{R}$. Suppose that $\mathfrak{a} \neq \{0\}$ and choose a nonzero element $a \in \mathfrak{a}$. Since $\mathfrak{R}$ is a field it follows that $1 = a^{-1}a \in \mathfrak{a}$. Hence, $\mathfrak{a} = R$.

($\Leftarrow$) The set $\mathfrak{a} := \{ a \in R \mid a \text{ is not a unit} \}$ is an ideal of R. Since $1 \notin \mathfrak{a}$ it follows that $\mathfrak{a} = \{0\}$. Consequently, every nonzero element of R is a unit and $\mathfrak{R}$ is a field. □

Corollary 5.9. *Let $\mathfrak{R}$ be a commutative ring and $\mathfrak{a} \subseteq R$ an ideal. The quotient $\mathfrak{R}/\mathfrak{a}$ is a field if and only if $\mathfrak{a}$ is maximal.*

Proof. By Theorem B1.4.19, each ideal of $\mathfrak{R}/\mathfrak{a}$ corresponds to an ideal $\mathfrak{b}$ of $\mathfrak{R}$ with $\mathfrak{a} \subseteq \mathfrak{b}$. Hence, $\mathfrak{R}/\mathfrak{a}$ is simple if and only if $\mathfrak{a}$ is maximal. Consequently, the claim follows from Lemma 5.8. □

Exercise 5.5. Show that every homomorphism between fields is an embedding.

The main part of this section is concerned with extensions of fields and ways to construct them. First we take a look at the subfields of a given fields.

Definition 5.10. Let $\mathfrak{K}$ be a field

(a) The *characteristic* of $\mathfrak{K}$ is the least number $n > 0$ such that

$$\underbrace{1 + \cdots + 1}_{n \text{ times}} = 0\,.$$

If there is no such number then we define the characteristic to be 0.

(b) The *subfield generated* by a subset $X \subseteq K$ is the set

$$\{\, ab^{-1} \mid a, b \in \langle\!\langle X \rangle\!\rangle_{\mathfrak{K}} \,\}\,.$$

(c) The *prime field* of $\mathfrak{K}$ is the subfield generated by $\varnothing$.

Example. (a) The prime field of $\mathbb{R}$ is $\mathbb{Q}$.

(b) Let p be a prime number. The ring $\mathbb{Z}/(p)$ of all integers modulo p is a field of characteristic p.

Exercise 5.6. Let $\mathfrak{K}$ be a field of characteristic $m > 0$. Prove that m is a prime number.

Lemma 5.11. *Let $\mathfrak{K}$ be a field with prime field $\mathfrak{K}_0$.*

(a) *$\mathfrak{K}$ has characteristic 0 if and only if $\mathfrak{K}_0 \cong \mathbb{Q}$.*

(b) *$\mathfrak{K}$ has characteristic $p > 0$ if and only if $\mathfrak{K}_0 \cong \mathbb{Z}/(p)$.*

Definition 5.12. (a) An embedding $h : \mathfrak{K} \to \mathfrak{L}$ of fields is called a *field extension.*

(b) Let $h : \mathfrak{K} \to \mathfrak{L}$ be a field extension. We can regard $\mathfrak{L}$ as a $\mathfrak{K}$-vector space by defining

$$\lambda a := h(\lambda) \cdot a\,, \quad \text{for } \lambda \in K \text{ and } a \in L\,.$$

The *dimension* of the extension h is the dimension of this vector space.

(c) If $\mathfrak{K} \to \mathfrak{L}$ is a field extension and $\bar{a} \subseteq L$, then we denote the subfield of $\mathfrak{L}$ generated by $K \cup \bar{a}$ by $\mathfrak{K}(\bar{a})$.

Example. The subfield of $\mathbb{R}$ generated by $\sqrt{2}$ is

$$K := \{\, a + b\sqrt{2} \mid a, b \in \mathbb{Q} \,\}\,.$$

The field extension $\mathbb{Q} \to \mathfrak{K}$ has dimension 2.

One way to obtain an extension of a field $\mathfrak{K}$ is by considering its polynomial ring $\mathfrak{K}[x]$. We can obtain a field extending $\mathfrak{K}$ by either forming the field of fractions $\mathrm{FF}(\mathfrak{K}[x])$, or by taking a suitable quotient $\mathfrak{K}[x]/\mathfrak{p}$. We start by taking a closer look at polynomial rings of fields.

Lemma 5.13. *Let $\mathfrak{R}$ be an integral domain and $p, q \in R[x]$ polynomials.*

$$\deg(pq) = \deg p + \deg q\,.$$

Proof. Let $m := \deg p$ and $n := \deg q$ and suppose that

$$p = a_m x^m + \dots + a_0 \quad \text{and} \quad q = b_n x^n + \dots + b_0\,.$$

If $p = 0$ or $q = 0$ then $\deg(pq) = \deg 0 = -\infty$ and we are done. Hence, suppose that p and q are nonzero. Then

$$pq = \sum_{k=0}^{m+n} \Big(\sum_{i=0}^{k} a_i b_{k-i} \Big) x^i = a_m b_n x^{m+n} + \sum_{k=0}^{m+n-1} \Big(\sum_{i=0}^{k} a_i b_{k-i} \Big) x^i$$

(where $a_i := 0$, for $i > m$, and $b_i := 0$, for $i > n$). By assumption $a_m \neq 0$ and $b_n \neq 0$. Since $\mathfrak{R}$ is an integral domain it follows that $a_m b_n \neq 0$. Hence, $\deg pq = m + n$. ☐

Lemma 5.14. *Let $\mathfrak{K}$ be a field.*

(a) *For all polynomials $p, q \in K[x]$ with $p \neq 0$, there exist polynomials $r, s \in K[x]$ such that*

$$q = rp + s \quad \textit{and} \quad \deg s < \deg p\,.$$

(b) *For every ideal $\mathfrak{a} \subseteq K[x]$, there exists a polynomial $p \in K[x]$ such that $(p) = \mathfrak{a}$.*

Proof. (a) Suppose that

$$p = a_m x^m + \cdots + a_0 \quad \text{and} \quad q = b_n x^n + \cdots + b_0\,,$$

where $a_m \neq 0$ and $b_n \neq 0$. We prove the claim by induction on n. If $m > n$ we can take $r := 0$ and $s := q$. Hence, we may assume that $m \leq n$. Setting

$$r' := a_m^{-1} b_n x^{n-m} \quad \text{and} \quad s' := q - r'p$$

it follows that $q = r'p + s'$ and the degree of s' is less than n. By inductive hypothesis, there are polynomials r'' and s'' such that $s' = r''p + s''$ and the degree of s'' is less than n. Consequently, we obtain the desired polynomials by setting $r := r' + r''$ and $s := s''$.

(b) If $\mathfrak{a} = \{0\} = (0)$ then there is nothing to do. Hence, suppose that $\mathfrak{a}$ contains some nonzero polynomial. Choose a nonzero polynomial $p \in \mathfrak{a}$ of minimal degree. We claim that $(p) = \mathfrak{a}$. Clearly, we have $(p) \subseteq \mathfrak{a}$. For the converse, let $\mathfrak{q} \in \mathfrak{a}$. By (a), there are polynomials $r, s \in K[x]$ such that $q = rp + s$ and $\deg s < \deg p$. Since $s = q - rp \in \mathfrak{a}$ it follows, by choice of p, that $s = 0$. Hence, $q = rp \in (p)$. ☐

Definition 5.15. Let $\mathfrak{R}$ be a ring, $p \in R[x]$ a polynomial, and $a \in R$.

(a) We define

$$p[a] := h_a(p)\,,$$

where $h_a : \mathfrak{R}[x] \to \mathfrak{R}$ is the unique homomorphism such that $h_a(x) = a$ and $h_a \upharpoonright R = \mathrm{id}$. The *polynomial function* associated with p is the function

$$p[x] : \mathfrak{R} \to \mathfrak{R} : a \mapsto p[a] .$$

(b) We say that a is a *root* of p if $p[a] = 0$.

Lemma 5.16. *Let $\mathfrak{K}$ be a field and $p \in K[x]$ a nonzero polynomial of degree n.*

(a) *If a is a root of p then $p = q \cdot (x - a)$, for some $q \in K[x]$.*

(b) *p has at most n roots in K.*

Proof. (a) We can use Lemma 5.14 to find polynomials q, r such that $p = q(x-a) + r$ and $\deg r < \deg(x - a) = 1$. Hence, $r \in K$ and it follows that

$$0 = p[a] = q[a](a - a) + r[a] = r[a] = r .$$

Consequently, $p = q(x - a)$.

(b) Let $a_0, \dots, a_{m-1}$ be an enumeration of all roots of p. By (a), we have $p = q(x - a_0)\cdots(x - a_{m-1})$. Therefore, the degree of p is at least m. □

Definition 5.17. Let $\mathfrak{R}$ be a ring. A nonzero polynomial $p \in R[x]$ is *irreducible* if p is not a unit and there is no factorisation $p = qr$ with $q, r \in R[x]$ such that neither q nor r is a unit.

Lemma 5.18. *Let $\mathfrak{K}$ be a field. A polynomial $p \in K[x]$ is irreducible if and only if the ideal (p) is maximal.*

Proof. ($\Rightarrow$) Suppose that $\mathfrak{a} \subseteq K[x]$ is an ideal with $(p) \subset \mathfrak{a}$. Fix some $q \in \mathfrak{a} \setminus (p)$. By Lemma 5.14, there is some polynomial r with $(r) = (p, q)$. In particular, $p = sr$, for some $s \in K[x]$. Since p is irreducible it follows that one of r or s is a unit. If r is a unit then we have $\mathfrak{a} \supseteq (p, q) = (r) = K[x]$. Otherwise, $r = s^{-1}p$ implies that $(r) = (p) \subset (p, q)$. Contradiction.

($\Leftarrow$) Let (p) be maximal and suppose that $p = qr$, for some $q, r \in K[x]$. Then $(p) \subseteq (q)$ and $(p) \subseteq (r)$. By maximality of (p) it follows that either $(q) = (p)$ or $(q) = K[x]$. In the latter case q is a unit and we are done. Hence, suppose that $(q) = (p)$. Similarly, we may assume that $(r) = (p)$. Consequently, there are units $u, v \in K[x]$ such that $q = up$ and $r = vp$. It follows that $p = qr = uvp^2$. This is only possible if $\deg p \leq 0$. Hence, $p \in K$. Contradiction. □

Lemma 5.19. *Let $\mathfrak{K}$ be a field. For every nonzero polynomial $p \in K[x]$, there exists a factorisation $p = cq_0 \cdots q_{m-1}$ where $c \in K$ and $q_0, \dots, q_{m-1} \in K[x]$ are irreducible.*

Proof. We prove the claim by induction on $\deg p$. If $p \in K$ or p is already irreducible then there is nothing to do. Otherwise, we can find polynomials $q, r \in K[x]$ of degree at least 1 such that $p = qr$. Since

$$\deg q = \deg p - \deg r < \deg p$$

we can use the inductive hypothesis to find a factorisation $q = cq_0 \cdots q_{l-1}$ of q into irreducible polynomials. In the same way we obtain such a factorisation $r = dr_0 \cdots r_{m-1}$ for r. It follows that $p = cdq_0 \cdots q_{l-1} r_0 \cdots r_{m-1}$. □

Lemma 5.20. *Let $\mathfrak{K}$ be a field and suppose that $p \in K[x]$ is an irreducible polynomial of degree n.*

(a) *$\mathfrak{K}[x]/(p)$ is a field.*

(b) *The field extension $\mathfrak{K} \to \mathfrak{K}[x]/(p)$ has dimension n.*

(c) *p has a root in $\mathfrak{K}[x]/(p)$.*

Proof. Let $\pi : \mathfrak{K}[x] \to \mathfrak{K}[x]/(p)$ be the canonical projection.

(a) follows from Lemma 5.18 and Corollary 5.9.

(c) $p[\pi(x)] = \pi(p) = 0$.

(b) We claim that $1, \pi(x), \dots, \pi(x^{n-1})$ form a basis of $\mathfrak{K}[x]/(p)$. First, let us show that these elements generate the $\mathfrak{K}$-vector space $\mathfrak{K}[x]/(p)$. For every $q \in K[x]$, we can use Lemma 5.14 to find polynomials $r, s \in$

$K[x]$ such that $q = rp + s$ and the degree of s is less than n. Hence, $s = a_{n-1}x^{n-1} + \cdots + a_0$, for some $a_0, \ldots, a_{n-1} \in K$, and

$$\pi(q) = \pi(s) = a_{n-1}\pi(x^{n-1}) + \cdots + a_1\pi(x) + a_0 \,.$$

It remains to prove that $1, \pi(x), \ldots, \pi(x^{n-1})$ are linearly independent. For a contradiction, suppose that there are nonzero coefficients $a_0, \ldots, a_{n-1} \in K$ such that

$$a_0 + a_1\pi(x) + \cdots + a_{n-1}\pi(x^{n-1}) = 0 \,.$$

Then there is some $b \in K[x]$ such that

$$a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} = bp \,.$$

But the degree of the polynomial on the left hand side is between 0 and $n-1$, while the degree of bp is either $-\infty$ or at least n. Contradiction. □

With the help of polynomial rings we can study field extensions.

Definition 5.21. Let $\mathfrak{K}$ be a field and $U \subseteq K$ a subring.

(a) A subset $X \subseteq K$ is *algebraically dependent* over U if there exist elements $a_0, \ldots, a_{n-1} \in X$ and a polynomial $p \in U[x_0, \ldots, x_{n-1}]$ such that $p[a_0, \ldots, a_{n-1}] = 0$. We call X *algebraically independent* over U if it is not algebraically dependent over U.

(b) A *transcendence basis* of $\mathfrak{K}$ over U is a maximal subset $I \subseteq K$ that is algebraically independent over U. The cardinality of a transcendence basis is called the *transcendence degree* of $\mathfrak{K}$ over U.

(d) An element $a \in K$ is *algebraic* over U if $\{a\}$ is algebraically dependent over U. Otherwise, a is *transcendental* over U. A field extension $h : \mathfrak{K} \to \mathfrak{L}$ is *algebraic* if every element $a \in L \setminus \operatorname{rng} h$ is algebraic over $\operatorname{rng} h$. Similarly, we call h *transcendental* if every $a \in L \setminus \operatorname{rng} h$ is transcendental over $\operatorname{rng} h$.

(e) The field $\mathfrak{K}$ is *algebraically closed* if every polynomial $p \in K[x]$ has a root in $\mathfrak{K}$.

Remark. The partial order of all algebraically independent subsets of a field $\mathfrak{K}$ has finite character and, consequently, it is inductively ordered. Hence, every field has a transcendence basis.

Lemma 5.22. *Let $h : \mathfrak{K} \to \mathfrak{L}$ be a field extension and $a \in L$ an element.*

(a) *If a is transcendental over K then*

$$\mathfrak{K}(a) \cong \mathrm{FF}(\mathfrak{K}[x]) .$$

(b) *If a is algebraic over K then there exists an irreducible polynomial $p \in \mathfrak{K}[x]$ such that*

$$\mathfrak{K}(a) \cong \mathfrak{K}[x]/(p) .$$

Proof. (a) There exists a unique embedding $h_0 : \mathfrak{K}[x] \to \mathfrak{L}$ with $h_0 \restriction K = \mathrm{id}$ and $h_0(x) = a$. Let $h : \mathrm{FF}(\mathfrak{K}[x]) \to \mathfrak{L}$ be the unique embedding with $h \restriction K[x] = h_0$. We claim that h is surjective. Every element of $\mathfrak{K}(a)$ is of the form bc^{-1}, for $b, c \in \langle\langle K \cup \{a\} \rangle\rangle_{\mathfrak{L}}$. Fix polynomials $p, q \in K[x]$ such that $b = h_0(p)$ and $c = h_0(q)$. Then $bc^{-1} = h_0(p) \cdot h_0(q)^{-1} = h(p/q)$.

(b) By Lemma 3.3, there exists a homomorphism $h : \mathfrak{K}[x] \to \mathfrak{K}(a)$ with $h(x) = a$ and $h \restriction K = \mathrm{id}$. Note that h is surjective since $K \cup \{a\} \subseteq \mathrm{rng}\, h$. The kernel $\mathrm{Ker}\, h$ is an ideal of $\mathfrak{K}[x]$. By Lemma 5.14, there exists a polynomial $p \in K[x]$ such that $\mathrm{Ker}\, h = (p)$. Let $\pi : \mathfrak{K}[x] \to \mathfrak{K}[x]/(p)$ be the canonical projection. By Theorem B1.4.12, there exists an isomorphism $g : \mathfrak{K}[x]/(p) \to \mathrm{rng}\, h = \mathfrak{K}(a)$ such that $h = g \circ \pi$. ☐

Definition 5.23. We call the polynomial p from statement (b) of the preceding lemma the *minimal polynomial* of a.

Lemma 5.24. *Let $\mathfrak{K} \to \mathfrak{L}$ be an extension of fields of characteristic* 0. *Suppose that $p \in K[x]$ is an irreducible polynomial (in $K[X]$) that can be factorised in $L[x]$ as*

$$p = (x - a)^n q , \quad \textit{for } a \in L,\ q \in L[x],\ n < \omega .$$

Then $n \leq 1$.

Proof. Note that $p' \notin (p)$ because $\deg p' < \deg p$. Hence, $(p) \subset (p, p')$. Since the polynomial p is irreducible, the ideal (p) is maximal and it follows that $(p, p') = K[X] = (1)$. Hence, there are $r, s \in K[x]$ such that $rp + sp' = 1$. Consequently,

$$r(x-a)^n q + s[n(x-a)^{n-1} q + (x-a)^n q'] = 1\,.$$

Setting $t := rq(x-a) + nsq + sq'(x-a)$ we obtain a polynomial such that $(x-a)^{n-1} t = 1$. This implies that $0 = \deg 1 = \deg (x-a)^{n-1} t \geq n-1$. □

Algebraically closed fields are particularly well-behaved. As we will prove below, they are uniquely determined by their characteristic and their transcendence degree.

Lemma 5.25. *Let $\mathfrak{K}$ be an algebraically closed field of transcendence degree κ. Then $|K| = \kappa \oplus \aleph_0$.*

Proof. Let $I \subseteq K$ be a transcendence basis of $\mathfrak{K}$ over $\varnothing$. Then $|K| \geq |I| = \kappa$. Furthermore, we have $|K| \geq \aleph_0$ since, if $K = \{a_0, \dots, a_{n-1}\}$ were finite, we could find a polynomial

$$p := (x-a_0)\cdots(x-a_{n-1}) + 1$$

without root in K. Hence, K would not be algebraically closed.

Therefore, we have $|K| \geq \kappa \oplus \aleph_0$ and it remains to prove the converse. For every element $a \in K \setminus I$, the set $I \cup \{a\}$ is algebraically dependent. Hence, there are elements $b_0, \dots, b_{n-1} \in I$ and a polynomial $p \in \mathbb{Q}[x, y_0, \dots, y_{n-1}]$ such that

$$p[a, b_0, \dots, b_{n-1}] = 0\,.$$

Setting $f(a) := \langle p, \bar{b} \rangle$ we obtain a function

$$f : K \setminus I \to \bigcup_{n<\omega} (\mathbb{Q}[x, \bar{y}] \times I^n)\,.$$

For every pair $\langle p, \bar{b} \rangle$, there are only finitely many elements $a \in K$ with $f(a) = \langle p, \bar{b} \rangle$ since $p[x, \bar{b}]$ has at most $\deg p < \aleph_0$ roots in K. It follows that

$$\begin{aligned} |K| &= \sum_{\langle p, \bar{b} \rangle \in \operatorname{rng} f} f^{-1}(\langle p, \bar{b} \rangle) \\ &\leq \aleph_0 \otimes |\operatorname{rng} f| = \aleph_0 \otimes (\aleph_0 \otimes \kappa^{<\omega}) \leq \aleph_0 \oplus \kappa . \end{aligned}$$ □

Lemma 5.26. *For every field $\mathfrak{K}$, there exists an extension $\mathfrak{K} \to \mathfrak{L}$ such that every polynomial in $K[x]$ of degree at least 1 has a root in L.*

Proof. We have seen in Lemma 5.20 that, if $p \in K[x]$ is a polynomial and q an irreducible factor of p, then the field $\mathfrak{K}[x]/(q)$ is an extension of $\mathfrak{K}$ in which p has the root x.

Fix an enumeration $(p_\alpha)_{\alpha<\kappa}$ of $K[x]$. We construct a chain $(\mathfrak{L}_\alpha)_{\alpha<\kappa}$ of fields $\mathfrak{L}_\alpha \supseteq \mathfrak{K}$ such that p_α has a root in $\mathfrak{L}_{\alpha+1}$. We set $\mathfrak{L}_0 := \mathfrak{K}$ and $\mathfrak{L}_\delta := \bigcup_{\alpha<\delta} \mathfrak{L}_\alpha$, for limit ordinals δ. For the successor step we define $\mathfrak{L}_{\alpha+1} := \mathfrak{L}_\alpha[x]/(q_\alpha)$ where q_α is an irreducible factor of p_α. The union $\mathfrak{L} := \bigcup_{\alpha<\kappa} \mathfrak{L}_\alpha$ is the desired extension of $\mathfrak{K}$. □

Proposition 5.27. *Every field $\mathfrak{K}$ has an extension $\mathfrak{K} \to \mathfrak{L}$ where $\mathfrak{L}$ is algebraically closed.*

Proof. By the preceding lemma, we can construct a chain $(\mathfrak{L}_n)_{n<\omega}$ as follows. $\mathfrak{L}_0 := \mathfrak{K}$ and $\mathfrak{L}_{n+1}$ is some extension of $\mathfrak{L}_n$ such that every polynomial in $L_n[x]$ has a root in L_{n+1}. The union $\mathfrak{L} := \bigcup_{n<\omega} \mathfrak{L}_n$ is algebraically closed since, if $p \in L[x]$ then $p \in L_n[x]$, for some n, and p has a root in $\mathfrak{L}_{n+1} \subseteq \mathfrak{L}$. □

The previous proposition tells us that every field has an algebraically closure. In the following lemmas we prove that it is unique.

Lemma 5.28. *Let $\mathfrak{K}_0 \to \mathfrak{L}_0$ and $\mathfrak{K}_1 \to \mathfrak{L}_1$ be field extensions with algebraically closed fields $\mathfrak{L}_0$ and $\mathfrak{L}_1$. If $\mathfrak{L}_1$ and $\mathfrak{L}_2$ have the same transcendence degree over, respectively, $\mathfrak{K}_0$ and $\mathfrak{K}_1$, then we can find, for every element $a \in L_0$ and every isomorphism $\pi : \mathfrak{K}_0 \to \mathfrak{K}_1$, and element $b \in L_1$ and an isomorphism $\sigma : \mathfrak{K}_0(a) \to \mathfrak{K}_1(b)$ such that $\sigma \restriction K_0 = \pi$.*

Proof. First, we consider the case that a is algebraic over K_0. Let p be the minimal polynomial. We can extend π to an isomorphism $\pi' : \mathfrak{K}_0[x] \to \mathfrak{K}_1[x]$. Let $q := \pi'(p)$. Since $\mathfrak{L}_1$ is algebraically closed, q has a root $b \in L_1$. It follows that

$$\mathfrak{K}_0(a) \cong \mathfrak{K}_0[x]/(p) \cong \mathfrak{K}_1[x]/(q) \cong \mathfrak{K}_1(b),$$

and this isomorphism extends π.

It remains to consider the case that a is transcendental over K_0. Then the transcendence degree of L_0 over K_0 is at least 1 and we can find an element $b \in L_1$ that is transcendental over K_1. It follows that

$$\mathfrak{K}_0(a) \cong \mathrm{FF}(\mathfrak{K}_0[x]) \cong \mathrm{FF}(\mathfrak{K}_1[x]) \cong \mathfrak{K}_1(b). \qquad \square$$

Theorem 5.29. *Let $\mathfrak{K}$ be a field and $h_0 : \mathfrak{K} \to \mathfrak{L}_0$ and $h_1 : \mathfrak{K} \to \mathfrak{L}_1$ algebraically closed extensions of $\mathfrak{K}$. If $\mathfrak{L}_0$ and $\mathfrak{L}_1$ have the same transcendence degree over $\mathfrak{K}$ then there exists an isomorphism $\pi : \mathfrak{L}_0 \cong \mathfrak{L}_1$ with $\pi \circ h_0 = h_1$.*

Proof. Since $\mathfrak{L}_0$ and $\mathfrak{L}_1$ have the same transcendence degree λ of $\mathfrak{K}$ we have $|\mathfrak{L}_0| = |K| \oplus \lambda = |\mathfrak{L}_1|$. Fix enumerations $(a_i)_{i<\kappa}$ and $(b_i)_{i<\kappa}$ of, respectively, L_0 and L_1. By induction on α, we construct increasing sequences

$$\mathfrak{L}_d^0 \subseteq \mathfrak{L}_d^1 \subseteq \dots \subseteq \mathfrak{L}_d^\alpha \subseteq \dots \quad \text{and} \quad \pi_0 \subseteq \pi_1 \subseteq \dots \subseteq \pi_\alpha \subseteq \dots$$

of subfields $\mathfrak{L}_d^\alpha \subseteq \mathfrak{L}_d$ and isomorphisms $\pi_\alpha : \mathfrak{L}_0^\alpha \to \mathfrak{L}_1^\alpha$ such that

$$a_\alpha \in \operatorname{dom} \pi_{\alpha+1} \quad \text{and} \quad b_\alpha \in \operatorname{rng} \pi_{\alpha+1}.$$

Then $\pi := \bigcup_\alpha \pi_\alpha$ is an isomorphism with $\operatorname{dom} \pi = L_0$ and $\operatorname{rng} \pi = L_1$.

We start with $\mathfrak{L}_d^0 := \mathfrak{K}$ and $\pi_0 := \mathrm{id}_K$. For limit ordinals δ, we take unions $\mathfrak{L}_d^\delta := \bigcup_{\alpha<\delta} \mathfrak{L}_d^\alpha$ and $\pi_\delta := \bigcup_{\alpha<\delta} \pi_\alpha$. For the successor step, suppose that $\pi_\alpha : \mathfrak{L}_0^\alpha \to \mathfrak{L}_1^\alpha$ has already been defined. We apply the preceding lemma twice, first to construct an extension $\sigma \supseteq \pi_\alpha$ with $a_\alpha \in \operatorname{dom} \sigma$, and then to find an extension $\pi_{\alpha+1} \supseteq \sigma$ with $b_\alpha \in \operatorname{rng} \pi_{\alpha+1}$. $\square$

Corollary 5.30. *Two algebraically closed fields with the same characteristic and the same transcendence degree are isomorphic.*

Corollary 5.31. *Let $\mathfrak{L}$ be an algebraically closed field. For every isomorphism $\sigma : \mathfrak{K}_0 \to \mathfrak{K}_1$ between subfields $\mathfrak{K}_0, \mathfrak{K}_1 \subseteq \mathfrak{L}$, there exists an automorphism $\pi \in \operatorname{Aut}\mathfrak{L}$ such that $\pi \restriction K_0 = \sigma$.*

We can use automorphisms to study algebraic field extensions. This leads to what is called Galois theory. Here, we present only a simple lemma that is needed in the next section.

Definition 5.32. Let $h : \mathfrak{K} \to \mathfrak{L}$ be a field extension. We set

$$\operatorname{Aut}(\mathfrak{L}/\mathfrak{K}) := \{\, \pi \in \operatorname{Aut}\mathfrak{L} \mid \pi \restriction \operatorname{rng} h = \mathrm{id} \,\}.$$

Lemma 5.33. *Let $\mathfrak{K} \to \mathfrak{L}$ be a field extension where $\mathfrak{L}$ is algebraically closed.*

(a) *If $a \in L$ is an element such that $\pi(a) = a$, for all $\pi \in \operatorname{Aut}(\mathfrak{L}/\mathfrak{K})$, then $a \in K$.*

(b) *If $C \subseteq L$ is a finite set such that $\pi[C] \subseteq C$, for all $\pi \in \operatorname{Aut}(\mathfrak{L}/\mathfrak{K})$, then there exists a polynomial $p \in K[x]$ of degree $\deg p = |C|$ such that C is the set of roots of p.*

Proof. (a) For a contradiction, suppose that $a \notin K$. First, we consider the case that a is algebraic over K. Let p be its minimal polynomial and let $a_0, \dots, a_{n-1}$ be the roots of p. We have $n = \deg p$. Since

$$\mathfrak{K}(a_i) \cong \mathfrak{K}[x]/(p) \cong \mathfrak{K}(a),$$

we can use Corollary 5.31 to find automorphisms $\pi_i \in \operatorname{Aut}(\mathfrak{L}/\mathfrak{K})$ such that $\pi_i(a) = a_i$. By assumption, this implies $a_i = a$. Hence, we have

$$p = (x-a)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} x_i,$$

which implies that $a, a^2, \dots, a^n \in K$. Contradiction.

It remains to consider the case that a is transcendental over K. Then a^2 is also transcendental over K. Hence,

$$\mathfrak{K}(a) \cong \mathrm{FF}(\mathfrak{K}[x]) \cong \mathfrak{K}(a^2)$$

and we can use Corollary 5.31 to find an automorphism $\pi \in \mathrm{Aut}(\mathfrak{L}/\mathfrak{K})$ with $\pi(a) = a^2$. This implies $a^2 = a$, i.e., $a = 1 \in K$. Contradiction.

(b) Suppose that $C = \{c_0, \dots, c_{n-1}\}$ and set

$$p := (x - c_0)\cdots(x - c_{n-1}) .$$

Clearly, C is the set of roots of p. Hence, it remains to prove that $p \in K[x]$. For every $\pi \in \mathrm{Aut}(\mathfrak{L}/\mathfrak{K})$, we have

$$\pi(p) = (x - \pi(c_0))\cdots(x - \pi(c_{n-1})) = p .$$

Hence, every coefficient of p is fixed by every element of $\mathrm{Aut}(\mathfrak{L}/\mathfrak{K})$. By (a), it follows that all coefficients of p belong to K. □

We conclude this section with a result stating that every finite dimensional field extension is generated by a single element (at least in characteristic 0).

Theorem 5.34. *Let $\mathfrak{K} \to \mathfrak{L}$ be an extension of fields of characteristic* 0. *For all algebraic elements $a, b \in L$, there exists a finite subset $U \subseteq K$ such that*

$$\mathfrak{K}(a, b) = \mathfrak{K}(ac + b) , \quad \text{for all } c \in K \setminus U .$$

Proof. W.l.o.g. we may assume that L is algebraically closed. Let p and q be the minimal polynomials of a and b, respectively. Let $a'_0, \dots, a'_{m-1} \in L$ be the roots of p and $b'_0, \dots, b'_{n-1} \in L$ the roots of q where $a'_0 = a$ and $b'_0 = b$. We claim that the set

$$U := \{ (b'_j - b)(a - a'_i)^{-1} \mid 1 \leq i < m \text{ and } 0 \leq j < n \}$$

has the desired properties. Let $c \in K \setminus U$ and set $d := ac + b$. We have to show that

$$K(a, b) = K(d) .$$

Clearly, $K(d) \subseteq K(a,b)$. For the converse, let $r \in K(d)[x]$ be a polynomial such that

$$(r) = (p, q[d - cx]) .$$

Then $p[a] = 0$ and $q[d - ca] = q[b] = 0$ implies that $r[a] = 0$. Furthermore, if $r[z] = 0$, for some $z \in L$, then we have $p[z] = 0$ and $q[d - cz] = 0$. The former implies that $z = a'_i$, for some i, while the latter implies that $d - cz = b'_j$, for some j. Hence,

$$ac + b - cz = b'_j \quad \text{implies} \quad (a - z)c = b'_j - b .$$

Since $c \notin U$ it follows that $z = a$. Consequently, a is the only root of r and we have

$$r = (x - a)^k , \quad \text{for some } k < \omega .$$

Since r divides p it follows that $p = (x - a)^k p_0$, for some $p_0 \in K(a)[x]$. As p is irreducible, we can use Lemma 5.24 to conclude that $k = 1$. Hence, $r = x - a$. Since $r \in K(d)[x]$ it follows that $a \in K(d)$. This, in turn, implies that $b = d - ac \in K(d)$. Consequently, $K(a,b) \subseteq K(d)$. □

6. *Ordered fields*

The field $\mathbb{C}$ of complex numbers is the canonical example of an algebraically closed field of characteristic zero. We have studied such fields in the previous section. In this section we study fields like the field $\mathbb{R}$ of real numbers. It turns out that the theory of $\mathbb{R}$ is more complicated than that of $\mathbb{C}$. We start by looking at fields equipped with a partial order.

Definition 6.1. (a) A structure $\mathfrak{R} = \langle R, +, -, \cdot, 0, 1, < \rangle$ is a *partially ordered ring* if $\langle R, +, -, \cdot, 0, 1 \rangle$ is a ring and $<$ is a strict partial order on R satisfying the following conditions:

- $a < b$ implies $a + c < b + c$, for all $a, b, c \in R$.

- $a < b$ and $c > 0$ implies $a \cdot c < b \cdot c$.

If $<$ is a linear order then we call $\mathfrak{R}$ an *ordered ring*.

(b) A ring $\mathfrak{R}$ is *orderable* if there exists a linear order $<$ such that $\langle \mathfrak{R}, < \rangle$ is an ordered ring.

(c) For an element $a \in R$ of an ordered ring $\mathfrak{R}$, we define

$$|a| := \begin{cases} a & \text{if } a \geq 0\,, \\ -a & \text{if } a < 0\,. \end{cases}$$

(d) A field $\mathfrak{K}$ is *real* if -1 cannot be written as a sum of squares.

Exercise 6.1. Let $\mathfrak{K}$ be an ordered field. Prove that $-1 < 0$.

Lemma 6.2. *If $\mathfrak{K}$ is an ordered field then $a^2 \geq 0$, for all $a \in K$.*

Proof. If $a > 0$ then we have $a \cdot a > 0 \cdot a = 0$. Similarly, if $a = 0$ then $a^2 = 0^2 = 0 \geq 0$. Hence, suppose that $a < 0$. Then we have

$$0 = a + (-a) < 0 + (-a) = -a\,,$$

which implies that $-a^2 = a \cdot (-a) < 0 \cdot (-a) = 0$. Consequently, we have $0 = (-a^2) + a^2 < 0 + a^2 = a^2$. □

Lemma 6.3. *Every orderable field has characteristic* 0.

Proof. By the previous lemma, we have $1 = 1^2 > 0$. This implies that $0 + 1 < 1 + 1$ and, by induction it follows that

$$1 + 1 < 1 + 1 + 1\,, \quad 1 + 1 + 1 < 1 + 1 + 1 + 1\,, \quad \dots$$

If some sum $1 + \dots + 1$ equals 0 then we have

$$0 < 1 < 1 + 1 < \dots < 1 + \dots + 1 < 0\,.$$

A contradiction. □

Lemma 6.4. *Let $\mathfrak{K}$ be a real field. Then $\langle \mathfrak{K}, \leq \rangle$ is partially ordered where*

$$a \leq b \quad :\text{iff} \quad b - a \text{ is a sum of squares.}$$

Proof. We start by showing that $\leq$ is a partial order. It is clearly reflexive. For transitivity, suppose that $b - a = x$ and $c - b = y$ where x and y are sums of squares. Then $c - a = x + y$ is also a sum of squares. Finally, suppose that $a \leq b$ and $b \leq a$ for $a \neq b$. Then $x := b - a$ and $y := a - b$ are nonzero sums of squares with $x + y = 0$. Suppose that $x = x_0^2 + \cdots + x_m^2$ and $y = y_0^2 + \cdots + y_n^2$. Then

$$-x_0^2 = x_1^2 + \cdots + x_m^2 + y_0^2 + \cdots + y_n^2$$

implies

$$-1 = (x_1/x_0)^2 + \cdots + (x_m/x_0)^2 + (y_0/x_0)^2 + \cdots + (y_n/x_0)^2 .$$

Contradiction.

To show that $\mathfrak{K}$ is partially ordered by $\leq$ note that, if $b - a$ and $c = c - 0$ are sums of squares and d is an arbitrary element then

$$(b + d) - (a + d) = b - a \quad \text{and} \quad bc - ac = (b - a)c$$

are also sums of squares. □

We have seen that every real field can be equipped with a canonical partial order. We would like to extend this partial order to a linear one. To do so we consider field extensions such that, for every pair of elements a, b, one of $a - b$ and $b - a$ is a square. In the following we denote by $\sqrt{a}$ an arbitrary root of the polynomial $x^2 - a$, either in the given field $\mathfrak{K}$ itself or one of its extensions.

Lemma 6.5. *Let $\mathfrak{K}$ be a real field and $a \in K$ an element.*

(a) *If a is a sum of squares then $\mathfrak{K}(\sqrt{a})$ is a real field.*

(b) *If $-a$ cannot be written as a sum of squares then $\mathfrak{K}(\sqrt{a})$ is a real field.*

Proof. For a contradiction, suppose that $\mathfrak{K}(\sqrt{a})$ is not real. This implies that $\sqrt{a} \notin K$. Furthermore, there are numbers $b_i, c_i \in K$ such that

$$-1 = \sum_{i<n} (b_i + c_i\sqrt{a})^2 = \sum_{i<n} (b_i^2 + 2b_i c_i \sqrt{a} + ac_i^2)\,.$$

Since $\mathfrak{K}(\sqrt{a})$ is a $\mathfrak{K}$-vector space with basis $\{1, \sqrt{a}\}$ it follows that

$$-1 = \sum_{i<n} (b_i^2 + ac_i^2) \quad \text{and} \quad 0 = \sum_{i<n} 2b_i c_i \sqrt{a}\,.$$

Consequently, if a is a sum of squares then so is -1 and $\mathfrak{K}$ is not real. This contradiction proves (a).

For (b), note that setting $d := \sum_i c_i^2$ the above equation implies

$$\begin{aligned} -a = \frac{1 + \sum_i b_i^2}{\sum_i c_i^2} &= \frac{\sum_i c_i^2 + \sum_i b_i^2 \cdot \sum_i c_i^2}{(\sum_i c_i^2)^2} \\ &= \sum_i (c_i/d)^2 + \sum_i b_i^2 \cdot \sum_i (c_i/d)^2\,, \end{aligned}$$

and $-a$ is a sum of squares. Again a contradiction. □

Corollary 6.6. *If $\mathfrak{K}$ is real and $a \in K$ then at least one of $\mathfrak{K}(\sqrt{a})$ and $\mathfrak{K}(\sqrt{-a})$ is real.*

Lemma 6.7. *Let $\mathfrak{K}$ be a real field and $p \in K[x]$ an irreducible polynomial of odd degree. If a is a root of p (in some extension of $\mathfrak{K}$) then $\mathfrak{K}(a)$ is a real field.*

Proof. We prove the claim by induction on $n := \deg p$. Suppose that $\mathfrak{K}(a)$ is not real. Then there are elements $b_i \in K(a)$ with

$$-1 = b_0^2 + \dots + b_k^2\,.$$

Since $\mathfrak{K}(a) \cong \mathfrak{K}[x]/(p)$ we can find polynomials $q_i \in K[x]$ of degree less than n such that $b_i \equiv q_i \pmod{p}$. It follows that

$$-1 \equiv q_0^2 + \dots + q_k^2 \pmod{p}\,.$$

Hence, there is some polynomial $r \in K[x]$ such that

$$-1 = q_0^2 + \cdots + q_k^2 + rp\,.$$

Each square q_i^2 has an even degree. Let m be the degree of the sum $q_0^2 + \cdots + q_k^2$. If $m \leq 0$ then we would have $r = 0$ and -1 would be a sum of squares of elements in K. Hence, we have $0 < m \leq 2n - 2$. As $n = \deg p$ is odd, it follows that the degree of r is also odd and at most $n - 2$. Let r_0 be an irreducible factor of r of odd degree and let c be a root of r_0. Then

$$-1 = (q_0[c])^2 + \cdots + (q_k[c])^2$$

is a sum of squares in $\mathfrak{K}(c)$. Hence, $\mathfrak{K}(c)$ is not real. This contradicts the inductive hypothesis since the degree of r_0 is odd and less than n. □

Definition 6.8. (a) A field is *real closed* if it is real and it has no proper algebraic extension that is real.

(b) A *real closure* of a field $\mathfrak{K}$ is an algebraic extension $\mathfrak{K} \to \mathfrak{L}$ that is real closed.

Theorem 6.9. *Every real field has a real closure.*

Proof. Let $\mathfrak{K}$ be a real field and let $\mathcal{R}$ be the set of all real fields that are algebraic extensions of $\mathfrak{K}$. Then $\mathcal{R}$ is inductively ordered by inclusion. Hence, it has a maximal element $\mathfrak{L}$. This is the desired real closure of $\mathfrak{K}$. □

Lemma 6.10. *Let $\mathfrak{K}$ be a real closed field. There exists a unique linear order $<$ such that $\langle \mathfrak{K}, < \rangle$ is an ordered field.*

Proof. Let $\leq$ be the partial order of Lemma 6.4. We claim that $\leq$ is linear. Suppose that $a \nleq b$. Then $b - a$ is not a sum of squares. By Lemma 6.5 it follows that $\mathfrak{K}(\sqrt{a-b})$ is real. Since $\mathfrak{K}$ is real closed we have $\sqrt{a-b} \in K$. Hence, $a - b$ is a square and we have $b \leq a$, as desired.

Finally, note that, since every sum of squares must be non-negative $\leq$ is the only possible linear order on K. □

Theorem 6.11. *A field is orderable if and only if it is real.*

Proof. ($\Rightarrow$) If $\langle \mathfrak{K}, < \rangle$ is an ordered field then $a^2 \geq 0$, for all $a \in K$. Hence, every sum of squares is non-negative.

($\Leftarrow$) Let $\mathfrak{K}$ be a real field and let $\mathfrak{L}$ be a real closure of $\mathfrak{K}$. Then $\mathfrak{L}$ has a unique linear order $<$. The restriction of $<$ to $\mathfrak{K}$ yields the desired order of $\mathfrak{K}$. □

Lemma 6.12. *Let $\mathfrak{K}_0$ be an ordered field and $\mathfrak{K}_0 \to \mathfrak{K}_1$ an (unordered) field extension such that there are no elements $c_i \in K_1$ and $a_i \in K_0$ with $a_i > 0$ and*

$$-1 = a_0 c_0^2 + \dots + a_{n-1} c_{n-1}^2 .$$

Let $\mathfrak{A}$ be the algebraic closure of $\mathfrak{K}_1$ and $\mathfrak{L} \subseteq \mathfrak{A}$ the subfield generated by the set $K_1 \cup \{ \sqrt{c} \mid c \in K_0,\ c > 0 \}$. Then $\mathfrak{L}$ is a real field whose canonical partial order extends that of $\mathfrak{K}_0$.

Proof. Since every positive element of $\mathfrak{K}_0$ has a square root in $\mathfrak{L}$ it follows that the canonical order of $\mathfrak{L}$ extends the order of $\mathfrak{K}_0$. Hence, we only need to prove that $\mathfrak{L}$ is real.

If $\mathfrak{L}$ were not real then we would have

$$-1 = a_0 c_0^2 + \dots + a_{n-1} c_{n-1}^2 ,$$

where $a_i = 1$ and $c_i \in L$, for $i < n$. Furthermore, by definition of $\mathfrak{L}$, there would be elements $b_0, \dots, b_{k-1} \in K_0$ such that $c_0, \dots, c_{n-1} \in K_1(\sqrt{b_0}, \dots, \sqrt{b_{k-1}})$.

Consequently, it is sufficient to prove that we cannot find elements $a_0, \dots, a_{n-1}, b_0, \dots, b_{k-1} \in K_0$ and $c_0, \dots, c_{n-1} \in K_1(\sqrt{b_0}, \dots, \sqrt{b_1})$ such that $a_i, b_i > 0$ and

$$-1 = a_0 c_0^2 + \dots + a_{n-1} c_{n-1}^2 .$$

We proceed by induction on k. For $k = 0$ the claim follows by our assumption on $\mathfrak{K}_1$. Hence, let $k > 0$ and, for a contradiction, suppose that there are elements a_i, b_i, and c_i as above. Then

$$c_i = u_i + v_i\sqrt{b_{k-1}}\,, \quad \text{where} \quad u_i, v_i \in K_1(\sqrt{b_0}, \dots, \sqrt{b_{k-2}})\,.$$

Hence,

$$\begin{aligned} -1 &= \sum_{i<n} a_i(u_i + v_i\sqrt{b_{k-1}})^2 \\ &= \sum_{i<n}\Big(a_i u_i^2 + a_i b_{k-1} v_i^2 + 2a_i u_i v_i\sqrt{b_{k-1}}\Big)\,. \end{aligned}$$

If $b_{k-1} \in K_1(\sqrt{b_0}, \dots, \sqrt{b_{k-2}})$ then we obtain the desired contradiction by inductive hypothesis. Hence, assume that b_{k-1} is not contained in this field. Then 1 and $\sqrt{b_{k-1}}$ are linearly independent and it follows that

$$-1 = \sum_{i<n}(a_i u_i^2 + a_i b_{k-1} v_i^2) \quad \text{and} \quad 0 = \sum_{i<n} 2a_i u_i v_i\sqrt{b_{k-1}}\,.$$

But the first equation contradicts the inductive hypothesis. □

Theorem 6.13. *Every ordered field $\mathfrak{K}$ has a real closure $\mathfrak{R}$ such that the canonical ordering of $\mathfrak{R}$ extends the order of $\mathfrak{K}$.*

Proof. Applying Lemma 6.12 with $\mathfrak{K}_0 = \mathfrak{K}_1 = \mathfrak{K}$ we obtain a real field $\mathfrak{L}$ such that the canonical partial order of $\mathfrak{L}$ extends the order of $\mathfrak{K}$. The claim follows since the canonical order of every real closure of $\mathfrak{L}$ extends the canonical order of $\mathfrak{L}$. □

The next theorem gives a more concrete characterisation of when a field is real closed.

Theorem 6.14. *Let $\mathfrak{K}$ be a real field. The following statements are equivalent:*

(1) *$\mathfrak{K}$ is real closed.*

(2) *$\mathfrak{K}(\sqrt{-1})$ is algebraically closed.*

(3) *Every polynomial $p \in K[x]$ of odd degree has a root in $\mathfrak{K}$ and, for every $a \in K$, either a or $-a$ is a square.*

Proof. (1) ⇒ (3) follows from Lemmas 6.5 and 6.7.

(3) ⇒ (2) We start by showing that every element $a + b\sqrt{-1} \in K(\sqrt{-1})$ has a square root in $K(\sqrt{-1})$. Let $<$ be an ordering of $\mathfrak{K}$. Then $a^2 + b^2 > 0$ implies that $a^2 + b^2$ is a square. Since $-\sqrt{a^2 + b^2} \leq a \leq \sqrt{a^2 + b^2}$ we have

$$e := \frac{a + \sqrt{a^2 + b^2}}{2} > 0\,.$$

Hence, e is also a square. Set $c := \sqrt{e}$ and $d := \frac{b}{2c}$. It follows that

$$\begin{aligned}
(c + d\sqrt{-1})^2 &= e + b\sqrt{-1} - \frac{b^2}{4e} \\
&= \frac{a}{2} + \frac{\sqrt{a^2 + b^2}}{2} + b\sqrt{-1} - \frac{b^2}{2(a + \sqrt{a^2 + b^2})} \\
&= \frac{a}{2} + b\sqrt{-1} + \frac{\sqrt{a^2 + b^2}(a + \sqrt{a^2 + b^2}) - b^2}{2(a + \sqrt{a^2 + b^2})} \\
&= \frac{a}{2} + b\sqrt{-1} + \frac{a\sqrt{a^2 + b^2} + a^2}{2(a + \sqrt{a^2 + b^2})} \\
&= a + b\sqrt{-1}\,,
\end{aligned}$$

as desired.

To prove that $\mathfrak{K}(\sqrt{-1})$ is algebraically closed we have to show that every irreducible polynomial $p \in K[x]$ has a root in $\mathfrak{K}(\sqrt{-1})$. Suppose that the degree of p is $n = 2^m l$ where l is odd. We prove the claim by induction on m. If $m = 0$ then the claim holds by assumption on $\mathfrak{K}$. Suppose that $m > 0$. Let $\mathfrak{K} \to \mathfrak{L}$ be an algebraic field extension in which p has n roots $a_0, \dots, a_{n-1}$. By Theorem 5.34, there exist finite subsets $U_{ik} \subseteq K$ such that

$$\mathfrak{K}(a_i + a_k, a_i a_k) = \mathfrak{K}(a_i + a_k + c a_i a_k)\,, \quad \text{for all } c \in K \setminus U_{ik}\,.$$

Fix some element $c \in K \setminus \bigcup_{i,k} U_{ik}$. By Lemma 5.33, there is a polynomial $q \in K[x]$ of degree $n(n-1)/2$ whose roots are the elements $a_i + a_k + ca_i a_k$. By inductive hypothesis, one of them is in $\mathfrak{K}(\sqrt{-1})$. Suppose that $a_i + a_k + ca_i a_k \in K(\sqrt{-1})$.

First, we show that $b := a_i + a_k \in K(\sqrt{-1})$ and $b' := a_i a_k \in K(\sqrt{-1})$. For a contradiction, suppose otherwise. Note that, if one of b and b' is not in $K(\sqrt{-1})$ then $b + cb' \in K(\sqrt{-1})$ implies that the other one also does not belong to $K(\sqrt{-1})$. Hence, $K(b, b', \sqrt{-1})$ is a $K(\sqrt{-1})$-vector space with basis $\{1, b, b'\}$. But these vectors are not linearly independent since they satisfy the equation $\lambda 1 - b - b' = 0$ with $\lambda = b + cb' \in K(\sqrt{-1})$. Contradiction.

Consequently, a_i is the root of a quadratic polynomial in $K(\sqrt{-1})[x]$. Since every element of $K(\sqrt{-1})$ has a square root it follows that $a_i \in K(\sqrt{-1})$.

(2) ⇒ (1) By Lemma 6.4, there exists a partial order

$$a \leq b \quad :\text{iff} \quad b - a \text{ is a sum of squares}$$

on $\mathfrak{K}$. We claim that $\leq$ is linear. This implies that $\mathfrak{K}$ is real.

It is sufficient to show that every element $a \in K$ satisfies $a \geq 0$ or $-a \geq 0$. Suppose that $a \neq 0$ is not a sum of squares. Let b be a root of the polynomial $x^2 - a$. Since b is algebraic over K we have $\mathfrak{K}(b) \subseteq \mathfrak{K}(\sqrt{-1})$. Hence, there are elements $c, d \in K$ with $b = c + d\sqrt{-1}$. Consequently,

$$b^2 = c^2 + 2cd\sqrt{-1} - d^2 .$$

Since $\mathfrak{K}(\sqrt{-1})$ is a $\mathfrak{K}$-vector space with basis $\{1, \sqrt{-1}\}$ it follows that $cd = 0$ and $b^2 = c^2 - d^2$. Since $b \notin K$ we have $d \neq 0$. Hence, $c = 0$ and $-a = -b^2 = d^2$ is a square.

Finally, note that the real closure $\mathfrak{R}$ of $\mathfrak{K}$ is contained in $\mathfrak{K}(\sqrt{-1})$ since the latter is algebraically closed. To show that $\mathfrak{K}$ is real closed we have to prove that $\mathfrak{R} = \mathfrak{K}$. For a contradiction, suppose that there is some element $a \in R \setminus K$. Since $a \in K(\sqrt{-1})$ there are elements $b, c \in K$ with $a = b + c\sqrt{-1}$. Hence, $\sqrt{-1} = (a - b)/c \in R$ and -1 is a square in R. Contradiction. □

We continue our investigation of ordered fields by looking at the roots of polynomials.

Lemma 6.15. *If $\mathfrak{K}$ is real closed then every polynomial $p \in K[x]$ can be written as a product of polynomials of degree at most 2.*

Proof. Since $\mathfrak{K}(\sqrt{-1})$ is algebraically closed it follows that

$$p = u(x - a_0)\cdots(x - a_{n-1}),$$

for some $a_0, \dots, a_{n-1}, u \in K(\sqrt{-1})$. For $c = a + b\sqrt{-1} \in K(\sqrt{-1})$ we denote by $c^* := a - b\sqrt{-1}$ its complex conjugate. The mapping $c \mapsto c^*$ is a field homomorphism. Therefore, we have $p[c]^* = p[c^*]$. It follows that, for every $i < n$, there is some $l < n$ with $a_i^* = a_l$. If $i = l$ we have $a_i \in K$ and $x - a_i$ is a factor of p in $\mathfrak{K}[x]$. Otherwise, p has the factor

$$(x - a_i)(x - a_l) = x^2 - (a_i + a_i^*)x + a_i a_i^*$$

with $a_i + a_i^* \in K$ and $a_i a_i^* \in K$. ☐

Lemma 6.16. *Let $p = x^n + a_{n-1}x^{n-1} + \dots + a_1 x + a_0$ be a polynomial over an ordered field $\mathfrak{K}$ and suppose that $b \in K$ is some element with $b > 1 + |a_0| + \dots + |a_{2n}|$. Then*

$$p[b] > 0 \quad \textit{and} \quad (-1)^n p[-b] > 0 .$$

Proof. Note that $b > 1$ implies $b^{i+1} > b^i$, for all i. Hence,

$$p[b] > b^n - \sum_{i<n} |a_i| \cdot b^i \geq b^n - b^{n-1} \sum_{i<n} |a_i| > 0 .$$

Similarly,

$$p[-b] = (-1)^n b^n + \sum_{i<n} (-1)^i a_i b^i$$

implies

$$(-1)^n p[-b] > b^n - \sum_{i<n} |a_i| \cdot b^i > 0 .$$

☐

Proposition 6.17. *An ordered field $\mathfrak{K}$ is real closed if and only if, for every polynomial $p \in K[x]$ and all elements $a < b$ in K with $p[a] < 0 < p[b]$, there exists some $c \in (a, b)$ with $p[c] = 0$.*

Proof. ($\Leftarrow$) We use the characterisation of Theorem 6.14 (3).

For $a \in K$ set $p := x^2 - a$. If $a > 0$ then $p[0] = -a < 0 < a = p[2a]$. Hence, there is some element $c \in (0, 2a)$ with $p[c] = 0$. This implies that $a = c^2$ is a square.

Similarly, if $a < 0$ then $p[a] = 2a < 0 < -a = p[0]$. As above we find an element c with $p[c] = 0$. Hence, $-a = c^2$ is a square.

Finally, let $p = x^{2n+1} + a_{2n}x^{2n} + \cdots + a_1x + a_0$ be a polynomial of odd degree. Choose $b \in K$ such that $b > 1 + |a_0| + \cdots + |a_{2n}|$. By Lemma 6.16 we have $p[-b] < 0 < p[b]$. Therefore, p has a root $c \in (-b, b)$.

($\Rightarrow$) Let $p = p_0^{k_0} \cdots p_n^{k_n}$ where each p_i is irreducible. Choosing the interval (a, b) small enough we may assume that there is exactly one factor p_i with $p_i[a] < 0 < p_i[b]$ while all other factors have constant sign on the interval (a, b). If $p_i = x + c$ then $a + c < 0 < b + c$ implies $-c \in (a, b)$. Hence, $-c$ is the desired root of p.

Suppose that $p_i = x^2 + cx + d$. As p_i is irreducible we have $4d - c^2 > 0$. It follows that

$$p_i[z] = (z + c/2)^2 + (d - c^2/4) > 0, \quad \text{for all } z \in (a, b).$$

This contradicts our choice of p_i. ☐

Lemma 6.18. *Let $\mathfrak{K}$ be an ordered field and $p \in K[x]$ a polynomial. For every element $a \in K$ with $p[a] > 0$, there exists some $\varepsilon > 0$ such that*

$$p[z] > 0, \quad \textit{for all } a - \varepsilon \leq z \leq a + \varepsilon.$$

Proof. We consider the polynomial $q := p[a + x]$. Suppose that

$$q = c_n x^n + \cdots + c_1 x + c_0.$$

Set $k := \max_{1 \le i \le n} |c_i|$ and let ε be the minimum of 1 and $c_0/2kn$. For $|z| \le \varepsilon$ it follows that

$$\begin{aligned} q[z] &= c_0 + c_1 z + \dots + c_n z^n \\ &\ge c_0 - \varepsilon|c_1| - \dots - \varepsilon^n |c_n| \\ &\ge c_0 - \varepsilon k - \dots - \varepsilon k \\ &= c_0 - \varepsilon k n \\ &\ge \frac{c_0}{2} = \frac{p[a]}{2} > 0 . \end{aligned}$$

□

Lemma 6.19. *Let $\mathfrak{K}$ be an ordered field and $p \in K[x]$ a polynomial. If $p'[a] > 0$ then there exist some $\varepsilon > 0$ such that*

$$\begin{aligned} &p[z] > p[a] , \quad \textit{for } a < z < a + \varepsilon , \\ &p[z] < p[a] , \quad \textit{for } a - \varepsilon < z < a . \end{aligned}$$

Proof. Set $q := p[a + x] - p[a]$. Since $q[0] = 0$ we have $q = xq_0$, for some $q_0 \in K[x]$. Furthermore, we have

$$q_0[0] = q_0[0] + 0 \cdot q_0'[0] = q'[0] = p'[a] > 0 .$$

Hence, we can use Lemma 6.18 to find a number $\varepsilon > 0$ such that

$$q_0[z] > 0 , \quad \text{for all } -\varepsilon < z < \varepsilon .$$

This implies that

$$\begin{aligned} &q[z] > 0 , \quad \text{for } 0 < z < \varepsilon , \\ \text{and} \quad &q[z] < 0 , \quad \text{for } -\varepsilon < z < 0 . \end{aligned}$$

□

Lemma 6.20. *Let $\mathfrak{K}$ be a real closed field and $p \in K[x]$ a polynomial. If $a < b$ are elements such that*

$$p'[z] \ge 0 , \quad \textit{for all } a \le z \le b ,$$

then $p[a] < p[b]$.

Proof. First, suppose that $p'[z] > 0$, for all $a \leq z \leq b$. If $p[a] \geq p[b]$ then applying Lemma 6.19 to a and b, respectively, we obtain elements $a < c < d < b$ with $p[d] < p[b] \leq p[a] < p[c]$. Consequently, Proposition 6.17 implies that the polynomial $p - p[a]$ has a root b_1 with $c < b_1 < d$. Since $p[b_1] = p[a]$ we can repeat this argument to obtain a second root b_2 of $p - p[a]$ with $a < b_2 < b_1$. Continuing in this way we obtain an infinite descending sequence $b_1 > b_2 > \dots$ of roots of $p - p[a]$. But every nonzero polynomial has only finitely many roots. Contradiction.

For the general case, fix an enumeration $c_0 < \dots < c_{k-1}$ of all roots of p' in the interval (a, b), and let $d_0 < \dots < d_{2k+2}$ be the sequence defined by

$$a < \frac{a + c_0}{2} < c_0 < \frac{c_0 + c_1}{2} < c_2 < \dots$$
$$< \frac{c_{k-2} + c_{k-1}}{2} < c_{k-1} < \frac{c_{k-1} + b}{2} < b\,.$$

It is sufficient to prove that $p[d_i] < p[d_{i+1}]$, for all $i \leq 2k$. Therefore, we may assume that $p'[z] > 0$ for all z in the interval $[a, b]$ except possibly for one of the endpoints.

Suppose that $p'[a] = 0$ and $p'[b] > 0$. If $p[a] > p[b]$ then applying Lemma 6.18 to the polynomial $p - p[b]$ we obtain some element $a < c < b$ with $p[c] > p[b]$. Since $p'[z] > 0$, for all $z \in [c, b]$ this contradicts the first part of the proof. Consequently, we have $p[a] \leq p[b]$. By the same argument it follows that $p[a] \leq p[(a + b)/2]$. Hence, the first part of the proof implies that $p[a] \leq p[(a + b)/2] < p[b]$, as desired.

For $p'[a] > 0$ and $p'[b] = 0$ the claim follows in the same way by exchanging the roles of a and b. □

We conclude this section by proving that the real closure of an order field is unique.

Lemma 6.21. *Let $\mathfrak{L}_0$ and $\mathfrak{L}_1$ be real closures of an ordered field $\mathfrak{K}$ whose canonical orders extend the order of $\mathfrak{K}$. Suppose that $a \in L_0 \setminus K$ is an element whose minimal polynomial has minimal degree. Then there exists an order preserving embedding $\mathfrak{K}(a) \to \mathfrak{L}_1$.*

Proof. Let p be the minimal polynomial of a and set $n := \deg p$. We start by showing that p has a root in L_1. Note that, by Lemma 6.16, there are elements $b_-, b_+ \in K$ with $b_- < a < b_+$. Further, note that, if q is a polynomial of degree less than n then all roots of q are in $\mathfrak{K}$. Hence, when z varies over L_i then the sign of $q[z]$ changes only at points $z \in K$.

By choice of p we have $p'[a] \neq 0$ since, otherwise, we would have $p' = (x-a)q$, for some q. Hence, $p = (x-a)^2 r$, for some r, which contradicts Lemma 5.24. Therefore, replacing p by $-p$ if necessary, we may assume that $p'[a] > 0$.

We claim that there are elements $c, d \in K$ with $c < a < d$ such that p' is positive on the interval $[c, d]$. Let c' be the largest root of p' that is less than a. If such a root does not exist then we set $c' := b_-$. Similarly, let d' be the smallest root of p' that is greater than a, or set $d' := b_+$ if there is no such root. Since p' has degree $n-1$ it follows that $c', d' \in K$. Furthermore, Proposition 6.17 implies that p' has constant sign on the interval (c', d'). Setting $c := (c' + a)/2$ and $d := (d' + a)/2$ we obtain the desired elements.

By Lemma 6.20 it follows that $p[c] < 0 < p[d]$. Hence, we can use Proposition 6.17 to find a root $b \in L_1$ of p.

Let $a_0 < \cdots < a_{l-1}$ be an increasing enumeration of all roots of p in L_0 and let $b_0 < \cdots < b_{m-1}$ be an increasing enumeration of all roots of p in L_1. We claim that $l = m$ and that there exists an order preserving embedding $\sigma : \mathfrak{K}(\bar{a}) \to \mathfrak{K}(\bar{b})$ with $\sigma(a_i) = b_i$ and $\sigma \restriction K = \mathrm{id}$.

Fix elements $c_1, \dots, c_{n-1} \in L_0$ such that $c_i^2 = a_i - a_{i-1}$. There exists an embedding $\sigma' : \mathfrak{K}(\bar{a}\bar{c}) \to \mathfrak{L}_1$ of unordered fields with $\sigma' \restriction K = \mathrm{id}$. Since

$$\sigma'(a_i) - \sigma'(a_{i-1}) = \sigma'(c_i)^2$$

it follows that $\sigma'(a_{i-1}) < \sigma'(a_i)$. Furthermore, $\sigma'(a_i)$ is a root of p. Hence, $\sigma'(a_i) \in \bar{b}$. This implies that $l \leq m$. Similarly, we can show that $m \leq l$. Hence, there exists an embedding $\sigma : \mathfrak{K}(\bar{a}) \to \mathfrak{K}(\bar{b})$ with $\sigma(a_i) = b_i$ and $\sigma \restriction K = \mathrm{id}$. It remains to show that σ is order preserving.

Let $z \in K(\bar{a})$ be an element with $z > 0$. We fix some $u \in L_0$ such that $u^2 = z$. As above we can find an embedding of unordered fields $\sigma'' :$

$\mathfrak{K}(\bar{a}\bar{c}u) \to \mathfrak{L}$ with $\sigma''(a_i) = b_i$ and $\sigma'' \restriction K = \mathrm{id}$. Hence, $\sigma'' \restriction K(\bar{a}) = \sigma$. Furthermore, $\sigma(z) = \sigma''(z) = \sigma''(u)^2 > 0$. □

Theorem 6.22. *If $\mathfrak{L}_0$ and $\mathfrak{L}_1$ are ordered real closures of an ordered field $\mathfrak{K}$ then there exists a unique isomorphism $\pi : \mathfrak{L}_0 \to \mathfrak{L}_1$ with $\pi \restriction K = \mathrm{id}$.*

Proof. As in Theorem 5.29, we construct increasing sequences of isomorphisms

$$\pi_\alpha : \mathfrak{L}_0^\alpha \to \mathfrak{L}_1^\alpha$$

where $\mathfrak{L}_i^0 \subseteq \mathfrak{L}_i^1 \subseteq \dots \subseteq \mathfrak{L}_i$ are increasing chains of subfields with union $\bigcup_\alpha \mathfrak{L}_i^\alpha = \mathfrak{L}_i$. The limit $\pi := \bigcup_\alpha \pi_\alpha$ is the desired isomorphism.

We start with $\pi_0 := \mathrm{id}_K$. For limit steps, we take unions $\pi_\delta := \bigcup_{\alpha<\delta} \pi_\alpha$. For the inductive step, we apply Lemma 6.21 twice. First, we select some element $a \in L_0 \setminus L_0^\alpha$ such that its minimal polynomial over $\mathfrak{L}_0^\alpha$ has minimal degree and we extend π_α to an isomorphism $\mathfrak{L}_0^\alpha(a) \to \mathfrak{L}_1^\alpha(b)$, for some $b \in L_1$. Then we select some element $d \in L_1 \setminus L_1^\alpha(b)$ and extend the isomorphism to $\pi_{\alpha+1} : \mathfrak{L}_0^\alpha(a, c) \to \mathfrak{L}_1^\alpha(b, d)$, for some $c \in L_0$. □

Part C.

First-Order Logic and its Extensions

C1. First-order logic

1. Infinitary first-order logic

Logics are languages to talk about structures and their elements. They can be used to assert that a given structure has a certain property, to define classes of structures, or to define relations inside a given structure. Let us start with a simple, but typical example.

Example. Let $\mathfrak{K}$ be a field and X a set of variables. The *Zariski logic* over $\mathfrak{K}$ is the set $\mathrm{ZL}[\mathfrak{K}, X] := \mathfrak{K}[X]$ of all polynomials over $\mathfrak{K}$ with unknowns from X.

Let $\mathfrak{L} \supseteq \mathfrak{K}$ be a field extending $\mathfrak{K}$. For a polynomial $p \in \mathrm{ZL}[\mathfrak{K}, X]$ and a variable assignment $\beta : X \to L$, recall that $p^{\mathfrak{L}}[\beta]$ denotes the value of p when we assign to each variable $x \in X$ the value $\beta(x)$. A polynomial $p \in \mathrm{ZL}[\mathfrak{K}, X]$ *defines* in a given field $\mathfrak{L} \supseteq \mathfrak{K}$ the set

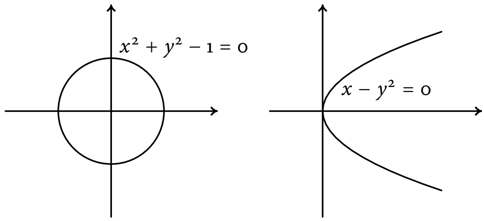
$$p^{\mathfrak{L}} := \{ \beta \in L^X \mid p^{\mathfrak{L}}[\beta] = 0 \}$$

of its roots. A set $A \subseteq L^n$ is *Zariski-definable over* $\mathfrak{K}$ if there exist finitely many polynomials $p_0, \dots, p_{k-1} \in \mathrm{ZL}[\mathfrak{K}, \{x_0, \dots, x_{n-1}\}]$ such that

$$A = \big\{ \langle \beta(x_0), \dots, \beta(x_{n-1}) \rangle \bigm| \beta \in p_0^{\mathfrak{L}} \cap \dots \cap p_{k-1}^{\mathfrak{L}} \big\} .$$

In case of algebraically closed fields the Zariski-definable relations are called *algebraic varieties.*

For instance, the polynomial $x^2 + y^2 - 1 = 0$ defines over $\mathbb{R}$ the unit circle S^1, while $x - y^2 = 0$ defines a rotated parabola.



Let us capture the above situation in a general definition.

Definition 1.1. A *logic* is a triple $\langle L, \mathcal{K}, \vDash \rangle$ consisting of a nonempty class L of *formulae,* a nonempty class $\mathcal{K}$ of *interpretations,* and a binary *satisfaction relation* $\vDash \subseteq \mathcal{K} \times L$.

Let $\mathfrak{J} \in \mathcal{K}$ be an interpretation and $\varphi \in L$ a formula. If $\mathfrak{J} \vDash \varphi$ then we say that φ *holds* in $\mathfrak{J}$, that $\mathfrak{J}$ *satisfies* φ, or that $\mathfrak{J}$ is a *model* of φ. For sets of formulae $\Phi \subseteq L$ we define

$$\mathfrak{J} \vDash \Phi \quad :\text{iff} \quad \mathfrak{J} \vDash \varphi \text{ for all } \varphi \in \Phi .$$

Example. (a) In the case of Zariski-logic $\mathrm{ZL}[\mathfrak{K}, X]$ the formulae are the polynomials $p \in \mathfrak{K}[X]$ and an interpretation is a pair $\langle \mathfrak{L}, \beta \rangle$ where $\mathfrak{L} \supseteq \mathfrak{K}$ is a field extension of $\mathfrak{K}$ and $\beta \in L^X$ is a variable assignment. We have

$$\langle \mathfrak{L}, \beta \rangle \vDash p \quad \text{iff} \quad p^{\mathfrak{L}}[\beta] = 0 .$$

(b) For a boolean algebra $\mathfrak{B}$, we define *boolean logic*

$$\mathrm{BL}(\mathfrak{B}) := \langle B, \ \mathrm{spec}(\mathfrak{B}), \ \vDash \rangle ,$$

where, for an element $b \in B$ and an ultrafilter $\mathfrak{u} \in \mathrm{spec}(\mathfrak{B})$,

$$\mathfrak{u} \vDash b \quad :\text{iff} \quad b \in \mathfrak{u} .$$

The main logic we will consider is *first-order logic,* also called *predicate logic.* We start by defining its *syntax,* that is, the set of first-order formulae. For convenience we simultaneously define two logics, basic first-order logic FO and a variant $\mathrm{FO}_{\kappa\aleph_0}$ where we allow infinite formulae.

Definition 1.2. Let Σ be a signature and κ an infinite cardinal. For each sort s of Σ, let X_s be a set of *variable symbols* of sort s, and set $X := \bigcup_s X_s$.

The set $\mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$ of *infinitary first-order formulae* is the smallest set of terms satisfying the following closure conditions:

- If $t_0, t_1 \in T[\Sigma, X]$ are of the same sort then $t_0 = t_1$ belongs to $\mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$.
- If $R \in \Sigma$ is of type $s_0 \dots s_{n-1}$ and $t_i \in T_{s_i}[\Sigma, X]$, for $i < n$, then $Rt_0 \dots t_{n-1}$ is in $\mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$.
- If $\varphi \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$ then $\neg\varphi \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$.
- If $\Phi \subseteq \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$ and $|\Phi| < \kappa$ then $\bigwedge \Phi, \bigvee \Phi \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$.
- If $\varphi \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X \cup \{x\}]$ then $\exists x\varphi, \forall x\varphi \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$.

For $\kappa = \aleph_0$, we obtain (*finitary*) *first-order logic*

$$\mathrm{FO}[\Sigma, X] := \mathrm{FO}_{\aleph_0\aleph_0}[\Sigma, X] .$$

If we omit the cardinality restriction, we get

$$\mathrm{FO}_{\infty\aleph_0}[\Sigma, X] := \bigcup_\kappa \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X] .$$

The operation $\neg$ is called *negation*, $\bigwedge$ and $\bigvee$ are *conjunction* and *disjunction*, and $\exists$ and $\forall$ are the *existential* and *universal quantifier*. An *atom* is a formula of the form

$$Rt_0 \dots t_{n-1} \quad \text{or} \quad t_0 = t_1 .$$

A formula that is either an atom or the negation of an atom is called a *literal*.

Remark. Every formula $\varphi \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$ is a term $\varphi : T \to \Lambda$ where $T \subseteq \kappa^{<\omega}$ and

$$\Lambda := \Sigma \cup X' \cup \{=, \neg, \bigwedge, \bigvee\} \cup \{\, \exists x, \forall x \mid x \in X' \,\} ,$$

for some $X' \supseteq X$. In particular, for $\kappa = \aleph_0$, we can regard $\mathrm{FO}[\Sigma, X]$ as a subset of $T[\Lambda, \varnothing]$.

It remains to define the meaning of these formulae, that is, the satisfaction relation. Before doing so, let us note that we can use induction on formulae.

Lemma 1.3. *If κ is a regular cardinal then we have* $\mathrm{frk}(\varphi) < \kappa$, *for all* $\varphi \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$.

Proof. Let Λ be the same set of symbols as in the preceding remark. The set Γ of all terms $t : T \to \Lambda$ such that $\mathrm{frk}(t) < \kappa$ is closed under all operations of Definition 1.2. Since $\mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$ is the smallest such set we have $\mathrm{FO}_{\kappa\aleph_0}[\Sigma, X] \subseteq \Gamma$, as desired. □

Corollary 1.4. $\mathrm{frk}(\varphi) < \infty$, *for all* $\varphi \in \mathrm{FO}_{\infty\aleph_0}[\Sigma, X]$.

This result implies that the reversed ordering on the domain of a formula is well-founded. Therefore, we can give proofs and definitions by induction on this order. A proof or a construction *by induction on* φ takes the following form. We have to distinguish several cases:

- φ is an atom.
- $\varphi = \neg\psi$ and the inductive hypothesis holds for ψ.
- $\varphi = \bigwedge \Phi$ or $\varphi = \bigvee \Phi$ and the inductive hypothesis holds for every element of Φ.
- $\varphi = \exists x\psi$ or $\varphi = \forall x\psi$ and the inductive hypothesis holds for ψ.

We use induction to define the *semantics* of first-order logic, that is, the satisfaction relation.

Definition 1.5. Let $\mathfrak{A}$ be a Σ-structure and $\beta : X \to A$ a variable assignment. The pair $\langle \mathfrak{A}, \beta \rangle$ is called a (*first-order*) *interpretation.* For $\varphi \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$ we define the *satisfaction relation* $\mathfrak{A} \vDash \varphi[\beta]$ by induction on φ.

$$
\begin{aligned}
&\mathfrak{A} \vDash t_0 = t_1[\beta] &&:\text{iff}\quad t_0^{\mathfrak{A}}[\beta] = t_1^{\mathfrak{A}}[\beta]\,,\\
&\mathfrak{A} \vDash Rt_0 \dots t_{n-1}[\beta] &&:\text{iff}\quad \big\langle t_0^{\mathfrak{A}}[\beta], \dots, t_{n-1}^{\mathfrak{A}}[\beta] \big\rangle \in R^{\mathfrak{A}}\,,\\
&\mathfrak{A} \vDash \neg\varphi[\beta] &&:\text{iff}\quad \mathfrak{A} \nvDash \varphi[\beta]\,,
\end{aligned}
$$

$$\begin{array}{lll}
\mathfrak{A} \vDash \bigvee \Phi[\beta] & :\text{iff} & \text{there is some } \varphi \in \Phi \text{ such that} \\
 & & \mathfrak{A} \vDash \varphi[\beta]\,, \\
\mathfrak{A} \vDash \bigwedge \Phi[\beta] & :\text{iff} & \mathfrak{A} \vDash \varphi[\beta] \text{ for all } \varphi \in \Phi\,, \\
\mathfrak{A} \vDash \exists x\varphi[\beta] & :\text{iff} & \text{there is some } a \in A \text{ such that} \\
 & & \mathfrak{A} \vDash \varphi[\beta[x/a]]\,, \\
\mathfrak{A} \vDash \forall x\varphi[\beta] & :\text{iff} & \mathfrak{A} \vDash \varphi[\beta[x/a]] \text{ for all } a \in A\,.
\end{array}$$

The set *defined* by a formula φ is $\varphi^{\mathfrak{A}} := \{\, \beta \in A^X \mid \mathfrak{A} \vDash \varphi[\beta] \,\}$.

Remark. For $X = \varnothing$, we simply write $\mathfrak{A} \vDash \varphi$ and we identify the pair $\langle \mathfrak{A}, \varnothing \rangle$ with the structure $\mathfrak{A}$. In this case $\varphi^{\mathfrak{A}}$ is either $\varnothing$ or $A^{\varnothing} = \{\varnothing\}$.

Exercise 1.1. Let $\mathfrak{N} := \langle \omega, +, 0, 1 \rangle$ be the natural numbers with addition and consider the formula

$$\varphi := \forall x \exists y [x = y + y \vee x = y + y + 1]\,.$$

Using the above definition, give a formal proof that $\mathfrak{N} \vDash \varphi$.

Definition 1.6. We will use the abbreviations

$$\begin{array}{ll}
\text{true} := \bigwedge \varnothing\,, & \text{false} := \bigvee \varnothing\,, \\
\varphi \vee \psi := \bigvee \{\varphi, \psi\}\,, & \varphi \rightarrow \psi := \neg\varphi \vee \psi\,, \\
\varphi \wedge \psi := \bigwedge \{\varphi, \psi\}\,, & \varphi \leftrightarrow \psi := (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)\,,
\end{array}$$

and $\quad t_0 \neq t_1 := \neg(t_0 = t_1)\,.$

The operation $\rightarrow$ is called *implication*. We abbreviate $\exists x_0 \cdots \exists x_{n-1}$ as $\exists \bar{x}$ and $\forall x_0 \cdots \forall x_{n-1}$ as $\forall \bar{x}$. Furthermore, we set

$$(\exists \bar{x}.\gamma)\varphi := \exists \bar{x}(\gamma \wedge \varphi) \quad \text{and} \quad (\forall \bar{x}.\gamma)\varphi := \forall \bar{x}(\gamma \rightarrow \varphi)\,.$$

Quantifiers of the form $(\exists \bar{x}.\gamma)$ and $(\forall \bar{x}.\gamma)$ are called *relativised quantifiers,* the formula γ is the *guard* of the quantifier.

Remark. To avoid unnecessary parenthesis we employ the following precedence rules.

- Unary operators like quantifiers, negation, and the large conjunction and disjunction signs bind strongest. For instance, the formula

$$\exists x \neg \bigwedge_{i<5} P_i x \wedge \exists y P_0 y \quad \text{is read as} \quad \Big(\exists x \neg \bigwedge_{i<5} P_i x\Big) \wedge (\exists y P_0 y)\,.$$

- $\wedge$ binds stronger than $\vee$, $\rightarrow$, and $\leftrightarrow$.
- The precedence between $\vee$, $\rightarrow$, and $\leftrightarrow$ is left unspecified.

Example. (a) Let $x_0, \dots, x_{n-1}$ be variables of sort s. The formula

$$\varphi_n := \exists x_0 \cdots \exists x_{n-1} \bigwedge_{i \neq k} x_i \neq x_k$$

expresses that the universe contains at least n different elements of sort s. Therefore, we can say that the domain of sort s is finite by the sentence

$$\varphi_{\text{fin}} := \bigvee \{\, \neg \varphi_n \mid n < \omega \,\}\,.$$

(b) Let $\Sigma = \{<\}$ be the signature of strict partial orders. We can express that an element y is the immediate successor of an element x by the formula

$$\varphi := x < y \wedge \neg \exists z (x < z \wedge z < y)\,.$$

(c) Let $\mathfrak{G} = \langle V, E \rangle$ be a graph. For every $n < \omega$, we can write down a first-order formula ψ_n saying that there exists a path of length at most n from the element x to y:

$$\psi_n := \exists z_0 \cdots \exists z_n \Big(z_0 = x \wedge z_n = y \wedge \bigwedge_{i<n} (z_i = z_{i+1} \vee E z_i z_{i+1}) \Big)\,.$$

The $\mathrm{FO}_{\aleph_1 \aleph_0}$-formula

$$\varphi_{\text{sc}} := \forall x \forall y \bigvee_{n<\omega} \psi_n$$

expresses that the graph is strongly connected.

(d) Let $\langle \mathbb{R}, +, -, 0, <, f \rangle$ be the additive ordered group of the real numbers with one unary function symbol f. We can say that $|x - y| < z$ by the formula

$$x - y < z \land y - x < z .$$

Making heavy use of relativised quantifiers, we can express that the function f is continuous at x by the formula

$$\begin{aligned}&(\forall \varepsilon.\varepsilon > 0)(\exists \delta.\delta > 0)\\&\qquad(\forall y.x - y < \delta \land y - x < \delta)\\&\qquad\qquad(fx - fy < \varepsilon \land fy - fx < \varepsilon) .\end{aligned}$$

Exercise 1.2. (a) Let $\langle A, \leq, P \rangle$ be a linear order with an additional unary predicate $P \subseteq A$. Write down a first-order formula $\varphi(x)$ which says that x is the supremum of P.

(b) Let $\langle V, E \rangle$ be a graph. Define a first-order formula φ which states that every vertex has exactly two outgoing edges.

Lemma 1.7. *For every ordinal $\alpha < \kappa$, there exists an* $\mathrm{FO}_{\kappa\aleph_0}$*-formula φ_α such that*

$$\mathfrak{A} \vDash \varphi_\alpha \quad \text{iff} \quad \mathfrak{A} \cong \langle \alpha, < \rangle .$$

Proof. We define a slightly more general formula $\psi_\alpha(x)$ such that

$$\mathfrak{A} \vDash \psi_\alpha(a) \quad \text{iff} \quad \langle \downarrow a, < \rangle \cong \langle \alpha, < \rangle .$$

The sentence

$$\vartheta := \forall x \neg(x < x) \land \forall x \forall y \forall z(x < y \land y < z \rightarrow x < z)$$

states that $<$ is a strict linear order. By induction on α, we set

$$\psi_0(x) := \neg\exists y(y < x) \land \vartheta ,$$

$$\text{and}\quad \psi_\alpha(x) := \bigwedge_{\beta<\alpha} (\exists y.y < x)\psi_\beta(y) \wedge (\forall y.y < x) \bigvee_{\beta<\alpha} \psi_\beta(y)\,.$$

Hence, we can define the desired formula φ_α by

$$\varphi_\alpha := \bigwedge_{\beta<\alpha} \exists y \psi_\beta(y) \wedge \forall y \bigvee_{\beta<\alpha} \psi_\beta(y)\,. \qquad \square$$

We can define the notions of a free variable, a subformula, substitution, etc. for formulae in the same way as for terms. But note that, unlike terms, formulae can contain variables that are not free.

Definition 1.8. Let $\varphi \in \mathrm{FO}_{\infty\aleph_0}[\Sigma, X]$.

(a) A subterm of φ is called a *subformula*.

(b) The set $\mathrm{free}(\varphi)$ of *free variables* of φ is the minimal set X_0 such that $\varphi \in \mathrm{FO}_{\infty\aleph_0}[\Sigma, X_0]$. A formula without free variables is called a *sentence*.

(c) An occurrence of a variable x in a formula φ is *bound* if it lies in a subformula of the form $\exists x\psi$ or $\forall x\psi$. Otherwise, the occurrence of x is *free*.

(d) For a sequence $\bar{s} \in S^I$ of sorts, let $X_{\bar{s}} := \{ x_i \mid i \in I \}$ be a standard set of variables where x_i is of sort s_i. We set

$$\mathrm{FO}^{\bar{s}}_{\kappa\aleph_0}[\Sigma] := \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X_{\bar{s}}]\,.$$

For ordinals α, we define

$$\mathrm{FO}^{\alpha}_{\kappa\aleph_0}[\Sigma] := \bigcup_{\bar{s}\in S^\alpha} \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X_{\bar{s}}]$$

$$\text{and}\quad \mathrm{FO}^{<\alpha}_{\kappa\aleph_0}[\Sigma] := \bigcup_{\beta<\alpha} \mathrm{FO}^{\beta}_{\kappa\aleph_0}[\Sigma]\,.$$

Remark. (a) Every $\mathrm{FO}_{\kappa\aleph_0}$-formula has less than κ free variables.

(b) Note that a variable x can occur both free and bound in the same formula φ.

Obviously, the truth value of a formula only depends on the symbols actually appearing in it. This triviality is recorded in the following lemma. Like the corresponding result for terms it can be proved by a straightforward induction on the structure of φ.

Lemma 1.9 (Coincidence Lemma). *Let $\varphi \in \mathrm{FO}_{\infty\aleph_0}[\Gamma, Y]$ be a formula and, for $i < 2$, let $\mathfrak{A}_i$ be a Σ_i-structure and $\beta_i : X_i \to A_i$ a variable assignment. If*

- $\Gamma \subseteq \Sigma_0 \cap \Sigma_1$ *and* $\mathrm{free}(\varphi) \subseteq X_0 \cap X_1$,
- $\mathfrak{A}_0|_\Gamma = \mathfrak{A}_1|_\Gamma$ *and* $\beta_0 \restriction \mathrm{free}(\varphi) = \beta_1 \restriction \mathrm{free}(\varphi)$

then we have $\mathfrak{A}_0 \vDash \varphi[\beta_0]$ *iff* $\mathfrak{A}_1 \vDash \varphi[\beta_1]$.

Remark. We will write $\varphi(x_0, \dots, x_{n-1})$ to indicate that

$$\mathrm{free}(\varphi) \subseteq \{x_0, \dots, x_{n-1}\}.$$

Furthermore, if $a_0, \dots, a_{n-1}$ are elements of the structure $\mathfrak{A}$, we write

$$\mathfrak{A} \vDash \varphi(a_0, \dots, a_{n-1})$$

instead of $\mathfrak{A} \vDash \varphi[\beta]$ for the assignment $\beta : x_i \mapsto a_i$. By the Coincidence Lemma, this notation is well-defined. Similarly, we write $\Phi(\bar{x})$ and $\mathfrak{A} \vDash \Phi(\bar{a})$, for sets $\Phi \subseteq \mathrm{FO}_{\infty\aleph_0}[\Sigma, X]$.

Let us compute the number of $\mathrm{FO}_{\kappa\aleph_0}$-formulae. Note that the number of finite formulae follows immediately from Lemma B3.1.5.

Lemma 1.10. *Let κ be a regular cardinal. Every formula $\varphi \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$ has less than κ subformulae.*

Proof. Using the same notation as in the remark after Definition 1.2, we see that φ is a Λ-term with $\mathrm{dom}\,\varphi \subseteq \kappa^{<\omega}$. If $\kappa = \aleph_0$ then φ is a finite term that has only finitely many subformulae. Suppose that $\kappa > \aleph_0$. Since κ is regular it follows by induction on φ that there exists a cardinal $\lambda < \kappa$ such that $\mathrm{dom}\,\varphi \subseteq \lambda^{<\omega}$. Hence, $|\mathrm{dom}\,\varphi| \leq \lambda^{<\omega} = \lambda \oplus \aleph_0 < \kappa$. ☐

Lemma 1.11. *Let Σ be a signature, X a set of variables, and κ a regular cardinal.*

$$|\mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]| \leq (|\Sigma| \oplus |X| \oplus \aleph_0)^{<\kappa}.$$

Proof. We have shown in the preceding lemma that every infinitary first-order formula $\varphi \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$ is a Λ-term with $|\mathrm{dom}\,\varphi| < \kappa$. Furthermore, we have

$$|\Lambda| \leq |\Sigma| \oplus |X'| \oplus \aleph_0 \leq |\Sigma| \oplus |X| \oplus |\mathrm{dom}\,\varphi| \oplus \aleph_0 .$$

Consequently, it follows that

$$\begin{aligned}|\mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]| &\leq \sup_{\lambda<\kappa} (|\Sigma| \oplus |X| \oplus \lambda \oplus \aleph_0)^\lambda \\ &= \sup_{\lambda<\kappa} (|\Sigma| \oplus |X| \oplus \aleph_0)^\lambda = (|\Sigma| \oplus |X| \oplus \aleph_0)^{<\kappa}.\end{aligned}$$ □

Remark. In the preceding lemma, we have tacitly identified formulae φ and ψ that differ only in the names of bound variables, i.e., variables in $X' \setminus X$. Hence, the above bound holds only up to this equivalence relation. Clearly, if we distinguish the formulae $\exists x Px$, $\exists y Py$, $\exists z Pz$,…then we can construct arbitrarily many formulae by using that many different variable names.

Exercise 1.3. Prove that every formula $\varphi \in \mathrm{FO}_{\infty\aleph_0}[\Sigma, X]$ can be rewritten to use only countably many different bound variables. That is, for every sort s, there exists a countable set Y_s such that φ can be written as Λ-term with

$$\Lambda := \Sigma \cup X \cup Y \cup \{=, \neg, \bigwedge, \bigvee\} \cup \{\, \exists x, \forall x \mid x \in X \cup Y \,\},$$

where $Y = \bigcup_s Y_s$. *Hint.* If ψ is a subformula of φ then $\mathrm{free}(\psi) \setminus X$ is finite.

We have seen that each $\mathrm{FO}_{\infty\aleph_0}$-formula has a foundation rank. Hence, we could measure the complexity of a formula by its foundation rank. But this measure is not very meaningful. There exists another rank for formulae that better reflects the semantics of first-order logic.

Definition 1.12. The *quantifier rank* $\mathrm{qr}(\varphi) \in \mathrm{On}$ of a formula $\varphi \in \mathrm{FO}_{\infty\aleph_0}$ is defined inductively by:

- $\mathrm{qr}(R\bar{t}) := 0$ and $\mathrm{qr}(t = t') := 0$.
- $\mathrm{qr}(\neg\varphi) := \mathrm{qr}(\varphi)$.
- $\mathrm{qr}(\exists x\varphi) := \mathrm{qr}(\forall x\varphi) := \mathrm{qr}(\varphi) + 1$.
- $\mathrm{qr}(\bigwedge \Phi) := \mathrm{qr}(\bigvee \Phi) := \sup \{ \mathrm{qr}(\varphi) \mid \varphi \in \Phi \}$.

A formula φ is *quantifier-free* if $\mathrm{qr}(\varphi) = 0$.

Example. For the formulae φ_{fin} and ψ_{sc} from the example on page 448, we have

$$\mathrm{qr}(\varphi_{\mathrm{fin}}) = \sup \{ \mathrm{qr}(\varphi_n) \mid n < \omega \} = \omega\,,$$

and $$\mathrm{qr}(\psi_{\mathrm{sc}}) = \sup \{ \mathrm{qr}(\psi_n) \mid n < \omega \} + 2 = \omega + 2\,.$$

Immediately from the respective definitions it follows that the foundation rank bounds the quantifier rank of a formula.

Lemma 1.13. $\mathrm{qr}(\varphi) \leq \mathrm{frk}(\varphi)$, *for all* $\varphi \in \mathrm{FO}_{\infty\aleph_0}[\Sigma, X]$.

Corollary 1.14. *If* κ *is a regular cardinal then we have* $\mathrm{qr}(\varphi) < \kappa$, *for all* $\varphi \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$.

If κ is singular then $\mathrm{FO}_{\kappa\aleph_0}$ can exhibit pathological behaviour. Fortunately, it is safe to ignore these logics and only consider $\mathrm{FO}_{\kappa\aleph_0}$ for regular cardinals κ.

Lemma 1.15. *For singular cardinals* κ, *the logics* $\mathrm{FO}_{\kappa\aleph_0}$ *and* $\mathrm{FO}_{\kappa^+\aleph_0}$ *have the same expressive power.*

Proof. Let κ be singular and fix a cofinal function $f : \mathrm{cf}\,\kappa \to \kappa$. Every conjunction of κ formulae can be written equivalently as nested conjunction of less then κ formulae:

$$\bigwedge_{i<\kappa} \varphi_i \quad \text{is equivalent to} \quad \bigwedge_{\alpha<\mathrm{cf}(\kappa)} \bigwedge_{i<f(\alpha)} \varphi_i\,.$$

Consequently, we can inductively transform every formula $\varphi \in \mathrm{FO}_{\kappa^+\aleph_0}$ into an equivalent $\mathrm{FO}_{\kappa\aleph_0}$-formula. □

2. Axiomatisations

Let us begin a more systematic investigation of what can be expressed in first-order logic. In this section we give examples of classes of structures that can be defined in $\mathrm{FO}_{\infty\aleph_0}$.

Definition 2.1. Let $\langle L, \mathcal{K}, \vDash \rangle$ be a logic.

(a) A set of formulae $\Phi \subseteq L$ *axiomatises* the class

$$\mathrm{Mod}_L(\Phi) := \{ \mathfrak{I} \in \mathcal{K} \mid \mathfrak{I} \vDash \Phi \} .$$

For a single formula we simply write $\mathrm{Mod}_L(\varphi) := \mathrm{Mod}_L(\{\varphi\})$.

(b) A class $\mathcal{C} \subseteq \mathcal{K}$ of interpretations is *L-axiomatisable* if

$$\mathcal{C} = \mathrm{Mod}_L(\Phi) , \quad \text{for some } \Phi \subseteq L .$$

If $\mathcal{C} = \mathrm{Mod}_L(\Phi)$, for a finite set $\Phi \subseteq L$, we say that $\mathcal{C}$ is *finitely L-axiomatisable.* If $\mathcal{C}$ is axiomatised by Φ, we call the set Φ an *axiom system* for $\mathcal{C}$ and every $\varphi \in \Phi$ is an *axiom.*

(c) A set of formulae $\Phi \subseteq L$ is *consistent,* or *satisfiable,* if $\mathrm{Mod}_L(\Phi) \neq \emptyset$. Otherwise, Φ is called *inconsistent,* or *unsatisfiable*. If $\mathrm{Mod}_L(\Phi) = \mathcal{K}$, then Φ is called *valid* or a *tautology.* We use the same terminology for single formulae φ.

Example (Partial orders). The class of all partial orders is finitely first-order axiomatised by

$$\forall x \forall y \forall z (x \leq y \land y \leq z \rightarrow x \leq z) ,$$
$$\forall x \forall y (x \leq y \land y \leq x \leftrightarrow x = y) .$$

We get an axiom system for the class of linear orders if we add the formula

$$\forall x \forall y (x \leq y \lor y \leq x) .$$

A linear order is *dense* if between any two elements there exists a third one. The corresponding first-order axiom is

$$\forall x \forall y (x < y \rightarrow \exists z (x < z \land z < y)) ,$$

where $x < y$ abbreviates $x \leq y \wedge x \neq y$. A dense linear order is *open* if it does not have a least and a greatest element.

$$\forall x \exists y \exists z (y < x \wedge x < z) .$$

A *discrete* linear order is an order where every element, except for the first one, has an immediate predecessor and every element, except for the last one, has an immediate successor.

$$\begin{aligned}
&\forall x \big[\exists y (y < x) \to \exists y (y < x \wedge \neg \exists z (y < z \wedge z < x)) \big] , \\
&\forall x \big[\exists y (x < y) \to \exists y (x < y \wedge \neg \exists z (x < z \wedge z < y)) \big] .
\end{aligned}$$

Example (Equivalence relations). The class of all structures $\mathfrak{A} = \langle A, \sim \rangle$ where $\sim$ is an equivalence relation can be axiomatised by the first-order formulae

$$\begin{aligned}
&\forall x (x \sim x) , \\
&\forall x \forall y (x \sim y \leftrightarrow y \sim x) , \\
&\forall x \forall y \forall z (x \sim y \wedge y \sim z \to x \sim z) .
\end{aligned}$$

Example (Lattices). An axiom system for the class of lattices was given in Lemma B2.2.4.

$$\begin{aligned}
&\forall x \forall y (x \sqsubseteq y \leftrightarrow x \sqcap y = x) \\
&\forall x (x \sqcap x = x \ \wedge \ x \sqcup x = x) \\
&\forall x \forall y (x \sqcap y = y \sqcap x \ \wedge \ x \sqcup y = y \sqcup x) \\
&\forall x \forall y (x \sqcap (x \sqcup y) = x \ \wedge \ x \sqcup (x \sqcap y) = x) \\
&\forall x \forall y \forall z (x \sqcap (y \sqcap z) = (x \sqcap y) \sqcap z) \\
&\forall x \forall y \forall z (x \sqcup (y \sqcup z) = (x \sqcup y) \sqcup z)
\end{aligned}$$

For boolean algebras we have to add the axioms

$$\begin{aligned}
&\bot \neq \top\,,\\
&\forall x(\bot \sqcap x = \bot \ \wedge\ \bot \sqcup x = x)\,,\\
&\forall x(\top \sqcap x = x \ \wedge\ \top \sqcup x = \top)\,,\\
&\forall x(x \sqcap x^* = \bot \wedge x \sqcup x^* = \top)\,,\\
&\forall x\forall y\forall z[x \sqcap (y \sqcup z) = (x \sqcap y) \sqcup (x \sqcap z)]\,,\\
&\forall x\forall y\forall z[x \sqcup (y \sqcap z) = (x \sqcup y) \sqcap (x \sqcup z)]\,.
\end{aligned}$$

Example (Groups). The class of all groups (in the signature $\{\cdot, {}^{-1}, e\}$) can be finitely axiomatised in first-order logic by the sentences

$$\begin{aligned}
&\forall x\forall y\forall z[x \cdot (y \cdot z) = (x \cdot y) \cdot z]\,,\\
&\forall x(x \cdot e = x)\,,\\
&\forall x(x \cdot x^{-1} = e)\,.
\end{aligned}$$

If we only allow multiplication then these axioms become

$$\begin{aligned}
&\forall x\forall y\forall z[x \cdot (y \cdot z) = (x \cdot y) \cdot z]\,,\\
&\exists e\forall x[x \cdot e = x \wedge \exists y(x \cdot y = e)]\,.
\end{aligned}$$

We can add the $\mathrm{FO}_{\aleph_1\aleph_0}$-sentence φ_{fin} from page 448 to obtain an axiom system for the class of all finite groups. But note that this is an infinitary formula. We will prove in Theorem c2.4.12 that this class cannot be axiomatised in first-order logic.

The class of all infinite groups on the other hand is first-order axiomatisable. To the group axioms we can add, for all $n < \omega$, the sentence

$$\exists x_0 \cdots \exists x_{n-1} \bigwedge_{i<k<n} x_i \neq x_k\,.$$

This axiom system is necessarily infinite. If the class of infinite groups where axiomatisable by a single first-order sentence, its negation could be used to construct an axiom system of the class of all finite groups.

Example (Rings). The class of all rings $\langle R, +, -, \cdot, 0, 1 \rangle$ is defined by

$$\begin{aligned}
&\forall x \forall y \forall z [x + (y + z) = (x + y) + z] , \\
&\forall x (x + 0 = x) , \\
&\forall x (x + (-x) = 0) , \\
&\forall x \forall y (x + y = y + x) , \\
&\forall x \forall y \forall z [x \cdot (y \cdot z) = (x \cdot y) \cdot z] , \\
&\forall x (x \cdot 1 = x \ \wedge \ 1 \cdot x = x) , \\
&\forall x \forall y \forall z [x \cdot (y + z) = x \cdot y + x \cdot z] , \\
&\forall x \forall y \forall z [(y + z) \cdot x = y \cdot x + z \cdot x] .
\end{aligned}$$

Example (Fields). We obtain an axiom system for the class of all fields if we add to the ring axioms the formulae

$$\begin{aligned}
&0 \neq 1 , \\
&\forall x \exists y (x \neq 0 \rightarrow x \cdot y = 1) , \\
&\forall x \forall y (x \cdot y = y \cdot x) .
\end{aligned}$$

To get axioms for the class of ordered fields, we further have to add the axioms for a linear order and the formulae

$$\begin{aligned}
&\forall x \forall y \forall z (x < y \rightarrow x + z < y + z) , \\
&\forall x \forall y \forall z (x < y \wedge 0 < z \rightarrow x \cdot z < y \cdot z) .
\end{aligned}$$

Example (Set theory). The axioms of set theory can be expressed in first-order logic. The signature consists just of one binary relation symbol $\in$.

First, let us collect some auxiliary formulae. The subset relation $x \subseteq y$ can be defined by the formula

$$\forall z (z \in x \rightarrow z \in y) .$$

There are formulae $\mathrm{Stage}(x)$ and $\mathrm{WellOrder}(x, y)$ that express, respectively, that the set x is a stage and that y is a well-order on the set x (exercise).

The Axiom of Extensionality reads

$$\forall a \forall b[a = b \leftrightarrow \forall x(x \in a \leftrightarrow x \in b)] .$$

To express the Axiom of Separation we need infinitely many formulae. For every first-order formula $\varphi(x, \bar{z}) \in \mathrm{FO}$, we have the formula

$$\forall \bar{z} \forall a \exists b \forall x[x \in b \leftrightarrow x \in a \wedge \varphi(x, \bar{z})] .$$

(Since the signature $\{\in\}$ of set theory does not contain constant symbols, we need parameters $\bar{z}$ for those sets that φ might refer to.)

The Axioms of Creation and Infinity are

$$\forall a(\exists s.\mathrm{Stage}(s))(a \in s)$$

and $$(\exists s.\mathrm{Stage}(s))\forall x[x \in s \rightarrow \wp(x) \in s] ,$$

where $\wp(x) \in s$ is an abbreviation for the formula

$$\exists z[z \in s \wedge \forall y(y \in z \leftrightarrow y \subseteq x)] .$$

For the Axiom of Choice we have the formula

$$\forall a \exists r \mathrm{WellOrder}(a, r) .$$

Finally, the Axiom of Replacement again consists of several formulae, one for every formula $\varphi(x, y, \bar{z}) \in \mathrm{FO}$.

$$\big(\forall \bar{z}.\mathrm{fun}_\varphi(\bar{z})\big)\big[\exists u \, \mathrm{dom}_\varphi(\bar{z}, u) \rightarrow \exists u \, \mathrm{rng}_\varphi(\bar{z}, u)\big] ,$$

where

$$\mathrm{fun}_\varphi(\bar{z}) := \forall x \forall y \forall y'[\varphi(x, y, \bar{z}) \wedge \varphi(x, y', \bar{z}) \rightarrow y = y']$$

says that φ defines a function and the formulae

$$\mathrm{dom}_\varphi(\bar{z}, u) := \forall x \forall y(\varphi(x, y, \bar{z}) \rightarrow x \in u)$$

and $$\mathrm{rng}_\varphi(\bar{z}, u) := \forall x \forall y(\varphi(x, y, \bar{z}) \rightarrow y \in u)$$

express that u contains, respectively, the domain and the range of the function defined by φ.

Exercise 2.1. Define the following formulae over the signature $\{\in\}$.

(a) $\mathrm{Stage}(x)$ states that the set x is a stage.

(b) $\mathrm{Pair}(x, y, z)$ expresses that $z = \langle x, y\rangle$.

(c) $\mathrm{WellOrder}(x, y)$ says that y is a well-order on the set x.

Lemma 2.2. *If $\mathfrak{A}$ is a finite Σ-structure then the class $\{\, \mathfrak{B} \mid \mathfrak{B} \cong \mathfrak{A} \,\}$ is first-order axiomatisable. If Σ is finite then it is finitely axiomatisable.*

Proof. First, we consider the case that Σ is finite. Let $a_0, \dots, a_{n-1}$ be an enumeration of A without repetitions. If $\mathfrak{A}$ has only one sort then we can axiomatise $\mathfrak{A}$ by the formula

$$\begin{aligned}\exists x_0 \cdots \exists x_{n-1}\Big(&\bigwedge_{0 \leq i < k < n} x_i \neq x_k \wedge \forall y \bigvee_{i<n} y = x_i \\ &\wedge \bigwedge \{\, R x_{i_0} \dots x_{i_k} \mid \langle a_{i_0}, \dots, a_{i_k}\rangle \in R^{\mathfrak{A}},\ R \in \Sigma \,\} \\ &\wedge \bigwedge \{\, \neg R x_{i_0} \dots x_{i_k} \mid \langle a_{i_0}, \dots, a_{i_k}\rangle \notin R^{\mathfrak{A}},\ R \in \Sigma \,\} \\ &\wedge \bigwedge \{\, f x_{i_0} \dots x_{i_k} = x_l \mid f^{\mathfrak{A}}(a_{i_0}, \dots, a_{i_k}) = a_l\,,\ f \in \Sigma \,\}\Big).\end{aligned}$$

The case of several sorts requires two modifications of this formula. We have to replace the subformula $\forall y \bigvee_i y = x_i$ by a conjunction of several such formulae where y is of the respective sort s and the disjunction ranges only over those i such that x_i has the same sort s. Furthermore, we have to remove from the conjunction $\bigwedge_{i<k} x_i \neq x_k$ all inequations $x_i \neq x_k$ where x_i and x_k have different sorts.

Suppose that Σ is infinite. For each finite subsignature $\Sigma_0 \subseteq \Sigma$, we can construct a formula φ_{Σ_0} axiomatising the Σ_0-reduct $\mathfrak{A}|_{\Sigma_0}$ of $\mathfrak{A}$. We claim that the set

$$\Phi := \{\, \varphi_{\Sigma_0} \mid \Sigma_0 \subseteq \Sigma \text{ is finite} \,\}$$

is the desired axiom system. Clearly, $\mathfrak{A} \vDash \Phi$. Conversely, suppose that $\mathfrak{B} \vDash \Phi$. Then B has exactly $n := |A|$ elements. For every finite signature

$\Sigma_0 \subseteq \Sigma$, there exists a sequence $\bar{b}^{\Sigma_0} \in B^n$ such that we can satisfy the formula φ_{Σ_0} if we assign to the variable x_i the element $b_i^{\Sigma_0}$. Define

$$S := \{ \Sigma_0 \subseteq \Sigma \mid \Sigma_0 \text{ finite} \}$$
$$\text{and} \quad S(\bar{b}) := \{ \Sigma_0 \in S \mid \bar{b}^{\Sigma_0} = \bar{b} \}, \quad \text{for } \bar{b} \in B^n.$$

Then $\langle S, \subseteq \rangle$ is a directed partial order with a finite partition $S = \bigcup_{\bar{b}} S(\bar{b})$. By Proposition B3.3.4, there exists some $\bar{b}$ such that $S(\bar{b})$ is a dense subset of S. It follows that the mapping $b_i \mapsto a_i$ is an isomorphism from $\mathfrak{B}$ to $\mathfrak{A}$. □

3. Theories

In the previous section we have studied sets of formulae and the classes they axiomatise. Now we turn to the dual question. Given a class of structures we try to determine which formulae hold.

Definition 3.1. Let $\langle L, \mathcal{K}, \vDash \rangle$ be a logic, $\mathfrak{J} \in \mathcal{K}$ an interpretation, $\varphi, \psi \in L$ formulae, and $\Phi \subseteq L$ a set of formulae.

(a) We write

$$\Phi \vDash \varphi \quad :\text{iff} \quad \mathrm{Mod}_L(\Phi) \subseteq \mathrm{Mod}_L(\varphi).$$

If $\Phi \vDash \varphi$ then φ is called a *consequence* of Φ. We also say that φ *follows* from Φ or that Φ *entails* φ.

If $\Phi = \{\psi\}$ we simply write $\psi \vDash \varphi$ and, for $\Phi = \varnothing$, we write $\vDash \varphi$. Note that we use the same symbol $\vDash$ both for the satisfaction relation and for the entailment relation. The object on the left-hand side can be used to resolve any ambiguities.

(b) If $\varphi \vDash \psi$ and $\psi \vDash \varphi$ then φ and ψ are called *equivalent* and we write $\varphi \equiv \psi$. Similarly, if $\Phi \cup \{\varphi\} \vDash \psi$ and $\Phi \cup \{\psi\} \vDash \varphi$, we say that φ and ψ are *equivalent modulo Φ*.

(c) The *closure of Φ under entailment* is the set

$$\Phi^{\vDash} := \{ \varphi \in L \mid \Phi \vDash \varphi \}.$$

Remark. Note that if L_0 and L_1 are logics with the same class of interpretations, we can generalise the above definitions of $\Phi \vDash \varphi$ and $\varphi \equiv \psi$ also to the case that $\Phi \subseteq L_0$, $\psi \in L_0$, and $\varphi \in L_1$.

Example. If $p, q \in \mathrm{ZL}[\mathfrak{K}, X]$ where $\mathfrak{K}$ is algebraically closed then we have

$$\begin{aligned} p \vDash q \quad &\text{iff} \quad \text{every zero of } p \text{ is a zero of } q \\ &\text{iff} \quad p \mid q^n, \quad \text{for some } n < \omega . \end{aligned}$$

Consequently, $p^{\vDash} = \{ q \in \mathfrak{K}[X] \mid p \mid q^n \text{ for some } n < \omega \} \trianglelefteq \mathfrak{K}[X]$ is the radical ideal generated by p and we have

$$p \equiv q \quad \text{iff} \quad p^m = aq^n \text{ for some } a \in \mathfrak{K} \text{ and } m, n < \omega .$$

Each non-constant polynomial is satisfiable. A constant polynomial p is satisfiable if and only if $p = 0$. The polynomial $p = 0$ is the only tautology.

The following properties of the entailment relation follow immediately from the definition.

Lemma 3.2. *Let $\langle L, \mathcal{K}, \vDash \rangle$ be a logic.*

(a) *$\vDash$ is a preorder on L.*

(b) *A set $\Phi \subseteq L$ is a final segment of $\langle L, \vDash \rangle$ if, and only if, $\Phi = \Phi^{\vDash}$.*

(c) *If $\Phi \subseteq L$ is inconsistent, then $\Phi^{\vDash} = L$.*

(d) *φ is a tautology if, and only if, $\varnothing \vDash \varphi$.*

Definition 3.3. Let $\langle L, \mathcal{K}, \vDash \rangle$ be a logic.

(a) An L-*theory* is a set of formulae $T \subseteq L$ with $T^{\vDash} = T$. The L-theory of a class $\mathcal{C} \in \mathcal{K}$ is the set

$$\mathrm{Th}_L(\mathcal{C}) := \{ \varphi \in L \mid \mathfrak{J} \vDash \varphi , \text{ for all } \mathfrak{J} \in \mathcal{C} \} .$$

The L-theory of a single interpretation $\mathfrak{J} \in \mathcal{K}$ is $\mathrm{Th}_L(\mathfrak{J}) := \mathrm{Th}_L(\{\mathfrak{J}\})$.

(b) An L-theory T is *complete* if it is of the form $T = \mathrm{Th}_L(\mathfrak{I})$, for some L-interpretation $\mathfrak{I}$.

(c) Two L-interpretations $\mathfrak{I}_0$ and $\mathfrak{I}_1$ are *L-equivalent* if

$$\mathrm{Th}_L(\mathfrak{I}_0) = \mathrm{Th}_L(\mathfrak{I}_1)\,.$$

We write $\mathfrak{I}_0 \equiv_L \mathfrak{I}_1$ to denote this fact. As usual we omit the index L if $L = \mathrm{FO}^\circ[\Sigma]$.

Example. Let $\mathfrak{B}$ be a boolean algebra, $a, b \in B$, and $\mathfrak{u} \in \mathrm{spec}(\mathfrak{B})$. For boolean logic $\mathrm{BL}(\mathfrak{B}) = \langle B,\ \mathrm{spec}(\mathfrak{B}),\ \vDash\rangle$, we have

$$\begin{aligned} a \vDash b \quad &\text{iff} \quad \text{every ultrafilter containing } a \text{ also contains } b \\ &\text{iff} \quad a \sqsubseteq b\,, \end{aligned}$$

and $\quad \mathrm{Th}_{\mathrm{BL}(\mathfrak{B})}(\mathfrak{u}) = \{\, b \in B \mid \mathfrak{u} \vDash b \,\} = \mathfrak{u}\,.$

Remark. (a) The function $\Phi \mapsto \Phi^{\vDash}$ is a closure operator on L whose closed sets are the theories. Consequently, the set of all L-theories forms a complete partial order where the least element is the set $\varnothing^{\vDash}$ of all tautologies and the greatest element is the set L of all formulae.

(b) For $\Phi \subseteq L$, we have

$$\Phi = \mathrm{Th}_L(\mathrm{Mod}_L(\Phi)) \quad \text{iff} \quad \Phi = \Phi^{\vDash} \quad \text{iff} \quad \Phi \text{ is a theory}\,.$$

Exercise 3.1. Let T be a satisfiable L-theory such that there is no satisfiable L-theory T' with $T \subset T'$. Prove that T is complete.

The following properties of the entailment relation follow immediately from the definition. We say that a logic L is *closed under negation* if, for every formula $\varphi \in L$, there is some formula $\neg\varphi \in L$ with

$$\mathfrak{I} \vDash \neg\varphi \quad \text{iff} \quad \mathfrak{I} \nvDash \varphi\,.$$

Similarly, L is *closed under implication* if there are formulae $\varphi \to \psi$ such that

$$\mathfrak{I} \vDash \varphi \to \psi \quad \text{iff} \quad \mathfrak{I} \nvDash \varphi \text{ or } \mathfrak{I} \vDash \psi \text{ or both.}$$

Lemma 3.4. *Let L be a logic, $\Phi \subseteq L$, and $\varphi, \psi \in L$.*

(a) *$\Phi \vDash \varphi$ implies $\Psi \vDash \varphi$, for every $\Psi \supseteq \Phi$.*

If L is closed under negation then we have

(b) *$\Phi \vDash \varphi$ if, and only if, $\Phi \cup \{\neg\varphi\}$ is inconsistent;*

(c) *φ is satisfiable if, and only if, $\neg\varphi$ is no tautology;*

(d) *Φ is a complete theory if, and only if, we have*

$$\Phi \nvDash \varphi \quad \text{iff} \quad \Phi \vDash \neg\varphi\,, \quad \textit{for all } \varphi \in L\,.$$

If L is closed under implication then we have

(f) *$\Phi \cup \{\varphi\} \vDash \psi$ if, and only if, $\Phi \vDash \varphi \to \psi$;*

(g) *$\varphi \equiv \psi$ modulo Φ if, and only if, $\Phi \vDash \varphi \to \psi$ and $\Phi \vDash \psi \to \varphi$.*

We conclude this section with a collection of equivalences that can be used to simplify first-order formulae. We start with the boolean operations which, of course, satisfy the laws of a boolean algebra.

Lemma 3.5. *The following equivalences hold for $\varphi, \psi, \vartheta \in \mathrm{FO}_{\infty\aleph_0}[\Sigma]$:*

(a) $\neg\neg\varphi \equiv \varphi$ (*elimination of double negation*)

(b) $\varphi \wedge \psi \equiv \psi \wedge \varphi$ (*commutativity*)
$\varphi \vee \psi \equiv \psi \vee \varphi$

(c) $(\varphi \wedge \psi) \wedge \vartheta \equiv \varphi \wedge (\psi \wedge \vartheta)$ (*associativity*)
$(\varphi \vee \psi) \vee \vartheta \equiv \varphi \vee (\psi \vee \vartheta)$

(d) $\varphi \wedge \varphi \equiv \varphi$ (*idempotence*)
$\varphi \vee \varphi \equiv \varphi$

(e) $\neg \bigwedge_{i<\alpha} \varphi_i \equiv \bigvee_{i<\alpha} \neg\varphi_i$ (*de Morgan's laws*)
$\neg \bigvee_{i<\alpha} \varphi_i \equiv \bigwedge_{i<\alpha} \neg\varphi_i$

(f) $\varphi \to \psi \equiv \neg\psi \to \neg\varphi$ (*contraposition*)

(g) $\varphi \wedge (\varphi \vee \psi) \equiv \varphi$ (*absorption*)
$\varphi \vee (\varphi \wedge \psi) \equiv \varphi$

(h) $\varphi \wedge (\psi \vee \vartheta) \equiv (\varphi \wedge \psi) \vee (\varphi \wedge \vartheta)$ (*distributivity*)
$\varphi \vee (\psi \wedge \vartheta) \equiv (\varphi \vee \psi) \wedge (\varphi \vee \vartheta)$

Lemma 3.6. *The following equivalences hold for* $\varphi, \psi \in \mathrm{FO}_{\infty\aleph_0}[\Sigma]$.

(a) $\exists x\varphi \vee \exists x\psi \equiv \exists x(\varphi \vee \psi)$
$\forall x\varphi \wedge \forall x\psi \equiv \forall x(\varphi \wedge \psi)$

(b) $\neg\exists x\varphi \equiv \forall x\neg\varphi$
$\neg\forall x\varphi \equiv \exists x\neg\varphi$

(c) $\exists x\exists y\varphi \equiv \exists y\exists x\varphi$
$\forall x\forall y\varphi \equiv \forall y\forall x\varphi$

Furthermore, if $x \notin \mathrm{free}(\varphi)$ *then we also have*

(d) $\varphi \wedge \exists x\psi \equiv \exists x(\varphi \wedge \psi)$
$\varphi \vee \forall x\psi \equiv \forall x(\varphi \vee \psi)$

(e) $\varphi \vee \exists x\psi \equiv \exists x(\varphi \vee \psi)$ *modulo* $\exists x(x = x)$
$\varphi \wedge \forall x\psi \equiv \forall x(\varphi \wedge \psi)$ *modulo* $\exists x(x = x)$

(f) $\varphi \equiv \exists x\varphi$ *modulo* $\exists x(x = x)$
$\varphi \equiv \forall x\varphi$ *modulo* $\exists x(x = x)$

Remark. Note that the equivalences (e) and (f) only hold in structures that contain at least one element of the corresponding sort.

Exercise 3.2. Prove some of the above equivalences.

Example. In general we have

$$\exists x(\varphi \wedge \psi) \not\equiv \exists x\varphi \wedge \exists x\psi\,,$$
$$\forall x(\varphi \vee \psi) \not\equiv \forall x\varphi \vee \forall x\psi\,,$$
$$\exists x \forall y\varphi \not\equiv \forall y \exists x\varphi\,.$$

For a counterexample, consider the structure $\mathfrak{A} = \langle A, P\rangle$ with $A = \{0,1\}$ and $P = \{1\}$. We have

$$\begin{aligned} &\mathfrak{A} \vDash \exists xPx \wedge \exists x\neg Px && \text{but} && \mathfrak{A} \nvDash \exists x(Px \wedge \neg Px)\,,\\ &\mathfrak{A} \vDash \forall x(Px \vee \neg Px) && \text{but} && \mathfrak{A} \nvDash \forall xPx \vee \forall x\neg Px\,,\\ &\mathfrak{A} \vDash \forall y\exists x(x = y) && \text{but} && \mathfrak{A} \nvDash \exists x\forall y(x = y)\,. \end{aligned}$$

4. *Normal forms*

In this section we study syntactic operations on first-order formulae. In particular, we will define several ways to simplify a given formula. We start by generalising the operation of substitution from terms to formulae.

Definition 4.1. Let $\varphi \in \mathrm{FO}_{\infty\aleph_0}[\Sigma, X]$ be a formula, $t \in T[\Sigma, X]$ a term, and $x \in X$ a variable. The *substitution* of t for x in φ is the formula $\varphi[x/t]$ obtained from φ by

- renaming the bound variables of φ such that no variable in $\mathrm{free}(t)$ is bound in φ, and
- replacing every free occurrence of x in φ by the term t.

Example. (a) When substituting terms in formulae, we have to take care to avoid clashes with bound variables in order not to change the meaning of the formula. For instance, consider the formula $\exists y(y + y = x)$ which expresses that x is divisible by 2. If we substitute y for x, we expect the formula to say that y is divisible by 2. If we rename the bound variable to z, we obtain the formula $\exists z(z + z = y)$ which has the expected semantics.

But if we forget the renaming, we get $\exists y(y + y = y)$ which has an altogether different meaning.

(b) Renaming bound variables does not change the meaning of a formula. But note that renaming of free variables does. For instance, we have $\exists zRxz \not\equiv \exists zRyz$ since the interpretation $\langle \mathfrak{A}, \beta \rangle$ with

$$\mathfrak{A} := \langle [2], \{\langle 0, 1 \rangle\} \rangle \quad \text{and} \quad \beta(x) := 0\,, \ \beta(y) := 1$$

satisfies the first formula but not the second one.

Remark. Note that, if $\varphi \equiv \psi$ are equivalent formulae, we have

$$\neg\varphi \equiv \neg\psi\,, \quad \exists x\varphi \equiv \exists x\psi\,, \quad \text{and} \quad \forall x\varphi \equiv \forall x\psi\,.$$

Similarly, $\varphi_i \equiv \psi_i$, for all i, implies that

$$\textstyle\bigwedge_i \varphi_i \equiv \bigwedge_i \psi_i \quad \text{and} \quad \bigvee_i \varphi_i \equiv \bigvee_i \psi_i\,.$$

By induction it follows that, if is φ a subformula of ϑ and $\varphi \equiv \psi$, then $\vartheta \equiv \vartheta[\varphi/\psi]$ where $\vartheta[\varphi/\psi]$ denotes the formula obtained from ϑ by replacing the subformula φ by ψ.

In the following we give a quick summary of various normal forms for first-order logic. That is, we present subsets $\Phi \subseteq \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$ defined by some syntactic criterion and we prove that every formula of $\mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$ is logically equivalent to an element of Φ. We start by simplifying the terms appearing in a formula.

Definition 4.2. A formula $\varphi \in \mathrm{FO}_{\infty\aleph_0}[\Sigma, X]$ is *term-reduced* if every atomic subformula of φ is of the form

$$R\bar{x}\,, \quad f\bar{x} = y\,, \quad \text{or} \quad y = z\,,$$

where $\bar{x}$, y, and z are variables.

Lemma 4.3. *For each formula $\varphi \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$, we can construct a term-reduced formula $\psi \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$ such that $\varphi \equiv \psi$.*

Proof. If φ is not term-reduced, it contains a subformula ϑ of the form $R\bar{t}$ or $f\bar{t} = s$ where not all elements of $\bar{t}$ and s are variables. Suppose that t_0 is not a variable. If z is a variable that does not appear in ϑ, we can replace $Rt_0 \dots t_{n-1}$ by the equivalent formula

$$\exists z(t_0 = z \land Rzt_1 \dots t_{n-1}) .$$

Similarly, we can replace $f\bar{t} = s$ by

$$\exists z(t_0 = z \land fzt_1 \dots t_{n-1} = s) .$$

By induction, it follows that, for every atomic subformula ϑ of φ, there exists a term-reduced formula $\chi_\vartheta \equiv \vartheta$. We obtain the desired formula ψ by replacing every atom ϑ in φ by the corresponding term-reduced formula χ_ϑ. □

Definition 4.4. (a) A formula is in *disjunctive normal form* if it is of the form

$$\bigvee \{ \bigwedge \Phi_i \mid i \in I \}$$

where each Φ_i is a set of literals.

(b) A formula is in *conjunctive normal form* if it is of the form

$$\bigwedge \{ \bigvee \Phi_i \mid i \in I \}$$

where each Φ_i is a set of literals.

Lemma 4.5. *For every quantifier-free formula $\varphi \in \mathrm{FO}[\Sigma, X]$, there exist equivalent $\mathrm{FO}[\Sigma, X]$-formulae $\mathrm{DNF}(\varphi)$ and $\mathrm{CNF}(\varphi)$ that are in, respectively, disjunctive normal form and conjunctive normal form.*

Proof. We construct $\mathrm{DNF}(\varphi)$ and $\mathrm{CNF}(\varphi)$ by induction on φ. If φ is a literal, we can set

$$\mathrm{DNF}(\varphi) := \varphi \quad \text{and} \quad \mathrm{CNF}(\varphi) := \varphi .$$

Suppose that, by inductive hypothesis, we have

$$\mathrm{DNF}(\psi) = \bigvee_i \bigwedge_k \alpha_{ik} \quad \text{and} \quad \mathrm{CNF}(\psi) = \bigwedge_i \bigvee_k \beta_{ik}$$

$$\mathrm{DNF}(\vartheta) = \bigvee_i \bigwedge_k \gamma_{ik} \quad \text{and} \quad \mathrm{CNF}(\vartheta) = \bigwedge_i \bigvee_k \delta_{ik}$$

Then we can set

$$\begin{aligned}
\mathrm{DNF}(\neg\psi) &:= \bigvee_i \bigwedge_k \neg\beta_{ik} \\
\mathrm{CNF}(\neg\psi) &:= \bigwedge_i \bigvee_k \neg\alpha_{ik} \\
\mathrm{DNF}(\psi \wedge \vartheta) &:= \bigvee_i \bigvee_j \Big(\bigwedge_k \alpha_{ik} \wedge \bigwedge_k \gamma_{jk}\Big) \\
\mathrm{CNF}(\psi \wedge \vartheta) &:= \mathrm{CNF}(\psi) \wedge \mathrm{CNF}(\vartheta) \\
\mathrm{DNF}(\psi \vee \vartheta) &:= \mathrm{DNF}(\psi) \vee \mathrm{DNF}(\vartheta) \\
\mathrm{CNF}(\psi \vee \vartheta) &:= \bigwedge_i \bigwedge_j \Big(\bigvee_k \beta_{ik} \wedge \bigvee_k \delta_{jk}\Big).
\end{aligned}$$

□

Exercise 4.1. Prove the corresponding statement for $\mathrm{FO}_{\infty\aleph_0}[\Sigma, X]$.

When doing inductions on the structure of a formula, it is sometimes useful not to have to treat the case of negations. In such cases we can use de Morgan's laws to move all negation signs directly in front of atoms.

Definition 4.6. Given a formula $\varphi \in \mathrm{FO}_{\infty\aleph_0}[\Sigma, X]$, we construct two formulae φ^+ and φ^- as follows. If φ is atomic, we set $\varphi^+ := \varphi$ and $\varphi^- := \neg\varphi$. For other formulae we define

$$\begin{aligned}
(\neg\psi)^+ &:= \psi^-, & (\neg\psi)^- &:= \psi^+, \\
(\textstyle\bigwedge \Phi)^+ &:= \textstyle\bigwedge \{\, \psi^+ \mid \psi \in \Phi \,\}, & (\textstyle\bigwedge \Phi)^- &:= \textstyle\bigvee \{\, \psi^- \mid \psi \in \Phi \,\}, \\
(\textstyle\bigvee \Phi)^+ &:= \textstyle\bigvee \{\, \psi^+ \mid \psi \in \Phi \,\}, & (\textstyle\bigvee \Phi)^- &:= \textstyle\bigwedge \{\, \psi^- \mid \psi \in \Phi \,\}, \\
(\exists x\psi)^+ &:= \exists x\psi^+, & (\exists x\psi)^- &:= \forall x\psi^-, \\
(\forall x\psi)^+ &:= \forall x\psi^+, & (\forall x\psi)^- &:= \exists x\psi^-.
\end{aligned}$$

The formula φ^+ is called *the negation normal form* of φ. It is denoted by $\text{NNF}(\varphi)$. We say that φ is *in negation normal form* if $\text{NNF}(\varphi) = \varphi$.

The following basic properties of the negation normal form of φ can be shown by a straightforward induction on the structure of φ.

Lemma 4.7. *Let* $\varphi \in \text{FO}_{\infty\aleph_0}[\Sigma, X]$.

(a) $\text{NNF}(\varphi) \equiv \varphi$ *and* $\varphi^- \equiv \neg\varphi$.

(b) $\text{NNF}(\varphi)$ *is in negation normal form.*

(c) φ *is in negation normal form if, and only if, the only subformulae of* φ *of the form* $\neg\psi$ *are literals.*

(d) $\text{qr}(\text{NNF}(\varphi)) = \text{qr}(\varphi)$.

Definition 4.8. A formula $\varphi \in \text{FO}_{\kappa\aleph_0}[\Sigma, X]$ is in *prenex normal form* if it is of the form

$$\varphi = Q_0 x_0 \cdots Q_{n-1} x_{n-1} \psi$$

where $Q_0, \dots, Q_{n-1} \in \{\exists, \forall\}$ and ψ is quantifier-free.

We can transform formulae into prenex normal form only for structures with nonempty universe.

Definition 4.9. Let T_{ne} be the theory consisting, for every sort s, of one formula $\exists x_s(x_s = x_s)$ where x_s is of sort s.

T_{ne} expresses that all domains of a structure are nonempty. For models of T_{ne} we can construct prenex normal forms.

Lemma 4.10. *For every formula* $\varphi \in \text{FO}[\Sigma, X]$*, there exists a formula* $\psi \in \text{FO}[\Sigma, X]$ *in prenex normal form such that* $\varphi \equiv \psi$ *modulo* T_{ne}.

Proof. By induction on φ, we can move the quantifiers to the front using the equivalences of Lemma 3.6. Suppose that the prenex normal forms of ψ and ϑ are, respectively,

$$Q_0 x_0 \cdots Q_{m-1} x_{m-1} \psi_0 \quad \text{and} \quad Q'_0 y_0 \cdots Q'_{n-1} y_{n-1} \vartheta_0 ,$$

where all variables x_i and y_k are distinct. For $Q \in \{\exists, \forall\}$, define $\overline{Q}$ by $\overline{\exists} := \forall$ and $\overline{\forall} := \exists$. The prenex normal form of φ is

$$\begin{array}{ll} \varphi & \text{if } \varphi \text{ is atomic}, \\ \overline{Q}_0 x_0 \cdots \overline{Q}_{m-1} x_{m-1} \neg \psi_0 & \text{for } \varphi = \neg\psi, \\ Q_0 x_0 \cdots Q_{m-1} x_{m-1} Q'_0 y_0 \cdots Q'_{n-1} y_{n-1} (\psi_0 \wedge \vartheta_0) & \text{for } \varphi = \psi \wedge \vartheta, \\ Q_0 x_0 \cdots Q_{m-1} x_{m-1} Q'_0 y_0 \cdots Q'_{n-1} y_{n-1} (\psi_0 \vee \vartheta_0) & \text{for } \varphi = \psi \vee \vartheta, \\ \exists z Q_0 x_0 \cdots Q_{m-1} x_{m-1} \psi_0 & \text{for } \varphi = \exists z \psi, \\ \forall z Q_0 x_0 \cdots Q_{m-1} x_{m-1} \psi_0 & \text{for } \varphi = \forall z \psi. \end{array}$$ □

In some cases we can get a prenex normal form that is fully equivalent instead of being only equivalent modulo T_{ne}.

Corollary 4.11. *Let Σ be an S-sorted signature satisfying either of the following conditions:*

- *For every $s \in S$, there is a constant symbol of sort s.*
- *$|S| = 1$ and S does not contain relations of arity 0.*

For every formula $\varphi \in \text{FO}[\Sigma, X]$, there exists a formula $\psi \in \text{FO}[\Sigma, X]$ in prenex normal form such that $\varphi \equiv \psi$.

Proof. In the first case, every Σ-structure is a model of T_{ne}. Hence, logical equivalence and equivalence modulo T_{ne} coincide.

In the second case, we can obtain ψ as follows. There exists a formula ψ' in prenex normal form such that $\varphi \equiv \psi'$ modulo T_{ne}. Note that, up to isomorphism, there exists exactly one Σ-structure $\mathfrak{A}_0$ with empty universe since we have no relations of arity 0. Let $x \notin \text{free}(\varphi)$ be a new variable. Note that $\mathfrak{A}_0 \models \forall x \psi'$ and $\mathfrak{A}_0 \not\models \exists x \psi'$ regardless of what the formula ψ' looks like. Hence, we can set

$$\psi := \begin{cases} \forall x \psi' & \text{if } \mathfrak{A}_0 \models \varphi, \\ \exists x \psi' & \text{otherwise}. \end{cases}$$

For every nonempty structure $\mathfrak{A}$, we have $\mathfrak{A} \models \psi'$ iff $\mathfrak{A} \models \psi$. Consequently, $\varphi \equiv \psi$. □

Remark. Infinitary formulae usually have no prenex normal form. For example, consider the sentence

$$\varphi := \bigwedge_{n<\omega} \exists x_0 \cdots \exists x_{n-1} \bigwedge_{i \neq k} x_i \neq x_k \,.$$

If we move all quantifiers to the front, we obtain a formula starting with an infinite string of quantifiers. This is forbidden by the definition of $\mathrm{FO}_{\infty\aleph_0}$.

When we are interested in whether some theory is satisfiable, we can also perform translations that, while preserving satisfiability, do not respect logical equivalence. For infinitary formulae the following reduction to first-order logic is useful. Another example is Skolemisation which transforms an arbitrary theory into a universal one (see Section C2.3).

Lemma 4.12 (Chang's Reduction). *For every $\varphi \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$, there exists a signature $\Sigma_\varphi \supseteq \Sigma$ and a set $\Phi_\varphi \subseteq \mathrm{FO}_{\kappa\aleph_0}[\Sigma_\varphi, X]$ with the following properties:*

- *Every model of φ can be expanded in exactly one way to a model of Φ_φ.*
- *Every model of Φ_φ is a model of φ.*
- *Every subformula of φ is equivalent modulo Φ_φ to an atomic formula.*
- *Every formula in Φ_φ is either a first-order formula or a sentence of the form $\forall \bar{x} \bigvee_i \psi_i(\bar{x})$ where each ψ_i is atomic.*

Proof. For every subformula $\psi(\bar{x})$ of φ with n free variables, choose two new n-ary relation symbols $R_\psi, R_{\neg\psi} \notin \Sigma$. Let Σ_φ be the signature consisting of Σ and all the new symbols $R_\psi, R_{\neg\psi}$. The set Φ_φ consists of the following formulae.

$$\forall \bar{x}(R_\psi \bar{x} \leftrightarrow \psi(\bar{x}))\,, \qquad \text{if } \psi \text{ is atomic.}$$
$$\forall \bar{x}(R_{\neg\psi} \bar{x} \leftrightarrow \neg R_\psi \bar{x})\,,$$

$$\begin{aligned}
&\forall \bar{x}(R_{\exists y\psi}\bar{x} \leftrightarrow \exists y R_\psi \bar{x}y)\,, \\
&\forall \bar{x}(R_{\forall y\psi}\bar{x} \leftrightarrow \forall y R_\psi \bar{x}y)\,, \\
&\forall \bar{x}(R_{\bigwedge_{i<\lambda}\psi_i}\bar{x} \rightarrow R_{\psi_i}\bar{x})\,, && \text{for all } i < \lambda\,, \\
&\forall \bar{x}(R_{\psi_i}\bar{x} \rightarrow R_{\bigvee_{i<\lambda}\psi_i}\bar{x})\,, && \text{for all } i < \lambda\,, \\
&\forall \bar{x}\Big[R_{\bigwedge_{i<\lambda}\psi_i}\bar{x} \vee \textstyle\bigvee_{i<\lambda} R_{\neg\psi_i}\bar{x}\Big]\,, \\
&\forall \bar{x}\Big[R_{\neg\bigvee_{i<\lambda}\psi_i}\bar{x} \vee \textstyle\bigvee_{i<\lambda} R_{\psi_i}\bar{x}\Big]\,.
\end{aligned}$$ □

5. Translations

In the last section we have considered transformations of formulae respecting logical equivalence. Now we turn to operations on structures and we investigate how to compute the theory of the resulting structure from the theories of the original ones. We start with a trivial example that illustrates the general situation.

Lemma 5.1. *Let $\Sigma \subseteq \Gamma$ be signatures. For every formula $\varphi(\bar{x}) \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma]$, there exists a formula $\psi(\bar{x}) \in \mathrm{FO}_{\kappa\aleph_0}[\Gamma]$ such that*

$$\mathfrak{A}|_\Sigma \vDash \varphi(\bar{a}) \quad \text{iff} \quad \mathfrak{A} \vDash \psi(\bar{a})\,,$$

for every Γ-structure $\mathfrak{A}$ and all $\bar{a} \subseteq A$.

Proof. We can set $\psi := \varphi$. □

Corollary 5.2. *Let $\mathfrak{A}$ and $\mathfrak{B}$ Γ-structures.*

$$\mathfrak{A} \equiv_{\mathrm{FO}_{\kappa\aleph_0}[\Gamma]} \mathfrak{B} \quad \textit{implies} \quad \mathfrak{A}|_\Sigma \equiv_{\mathrm{FO}_{\kappa\aleph_0}[\Sigma]} \mathfrak{B}|_\Sigma\,, \quad \textit{for all } \Sigma \subseteq \Gamma\,.$$

The other results of this section are all of the above form. We consider an operation F on structures and logics L and L', and we prove that, for every formula $\varphi \in L$, one can construct a formula $\varphi' \in L'$ such that

$$F(\mathfrak{A}) \vDash \varphi \quad \text{iff} \quad \mathfrak{A} \vDash \varphi'\,, \quad \text{for every } \mathfrak{A}\,.$$

As a consequence we obtain the result that

$$\mathfrak{A} \equiv_{L'} \mathfrak{B} \quad \text{implies} \quad F(\mathfrak{A}) \equiv_L F(\mathfrak{B}) .$$

In the case that $L' = L$ we call such operations *compatible* with L.

As a converse to the introductory example we consider expansions of a structure. Of course, there is no hope to reduce the theory of an arbitrary expansion to the original structure. But if we expand a structure by definable relations, such a reduction is possible.

Definition 5.3. Let $\mathfrak{A}$ be a Σ-structure and $\Gamma \supseteq \Sigma$. A Γ-structure $\mathfrak{B}$ is an *L-definable expansion* of $\mathfrak{A}$ if $\mathfrak{B}|_\Sigma = \mathfrak{A}$ and, for every symbol $\xi \in \Gamma \smallsetminus \Sigma$, there is some L-formula φ_ξ such that

$$\begin{aligned} \bar a \in R^{\mathfrak{B}} \quad &\text{iff} \quad \mathfrak{A} \vDash \varphi_R(\bar a) , \qquad \text{for all relations } R \in \Gamma \smallsetminus \Sigma , \\ f^{\mathfrak{B}}(\bar a) = b \quad &\text{iff} \quad \mathfrak{A} \vDash \varphi_f(\bar a, b) , \quad \text{for all functions } f \in \Gamma \smallsetminus \Sigma . \end{aligned}$$

In this case we also say that $(\varphi_\xi)_{\xi \in \Gamma \smallsetminus \Sigma}$ *defines the expansion* $\mathfrak{B}$ of $\mathfrak{A}$.

Lemma 5.4. *Let $\Sigma \subseteq \Gamma$ be signatures and let $\varphi_\xi(\bar x) \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma]$, for $\xi \in \Gamma \smallsetminus \Sigma$, be formulae. For every formula $\psi(\bar x) \in \mathrm{FO}_{\kappa\aleph_0}[\Gamma]$, there exists a formula $\psi^+(\bar x) \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma]$ such that*

$$\mathfrak{A}_+ \vDash \psi(\bar a) \quad \text{iff} \quad \mathfrak{A} \vDash \psi^+(\bar a) ,$$

whenever $\bar a \subseteq A$ and $\mathfrak{A}_+$ is the expansion of $\mathfrak{A}$ defined by $(\varphi_\xi)_\xi$.

Proof. Let ψ' be a term-reduced formula equivalent to ψ. We can obtain ψ^+ by replacing in ψ'

- every atom $R\bar t$ with $R \in \Gamma \smallsetminus \Sigma$ by the formula $\varphi_R(\bar t)$ and
- every atom $f\bar t = s$ with $f \in \Gamma \smallsetminus \Sigma$ by $\varphi_f(\bar t, s)$. □

Next we consider substructures. Again we have to restrict ourselves to those where the universe is definable.

Definition 5.5. Let $\mathfrak{A}$ be an S-sorted Σ-structure and $\delta_s(x) \in \mathrm{FO}^s_{\kappa\aleph_0}[\Sigma]$, for $s \in S$.

(a) If $\bigcup_{s\in S} \delta_s^{\mathfrak{A}}$ induces a substructure $\mathfrak{A}_0$ of $\mathfrak{A}$, we call $\mathfrak{A}_0$ the substructure *defined* by $(\delta_s)_{s\in S}$.

(b) The *relativisation* of a formula $\varphi \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma]$ to $(\delta_s)_{s\in S}$ is the formula $\varphi^{(\bar{\delta})} \in \mathrm{FO}_{\kappa\aleph_0}$ obtained from φ by replacing every subformula of the form $\exists y\psi$ and $\forall y\psi$ by, respectively,

$$(\exists y.\delta_s(y))\psi \quad \text{and} \quad (\forall y.\delta_s(y))\psi\,,$$

where s is the sort of y.

Lemma 5.6. *If a sequence $(\delta_s)_{s\in S}$ of $\mathrm{FO}_{\kappa\aleph_0}$-formulae defines a substructure $\mathfrak{A}_0$ of $\mathfrak{A}$, we have*

$$\mathfrak{A}_0 \vDash \varphi(\bar{a}) \quad \text{iff} \quad \mathfrak{A} \vDash \varphi^{(\bar{\delta})}(\bar{a})\,,$$

for every $\varphi \in \mathrm{FO}_{\kappa\aleph_0}$ and all $\bar{a} \subseteq \bigcup_{s\in S} \delta_s^{\mathfrak{A}}$.

Exercise 5.1. Prove Lemma 5.6.

Factorisation by definable congruences is also compatible with first-order logic.

Lemma 5.7. *Let Σ be an S-sorted signature and $\varepsilon_s(x,y) \in \mathrm{FO}^{ss}_{\kappa\aleph_0}[\Sigma]$, for $s \in S$. For every formula $\varphi(\bar{x}) \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$, there exists a formula $\varphi'(\bar{x}) \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$ such that, if $\approx\, := \bigcup_{s\in S} \varepsilon_s^{\mathfrak{A}}$ is a congruence relation on $\mathfrak{A}$, then*

$$\mathfrak{A}/{\approx} \vDash \varphi([\bar{a}]_{\approx}) \quad \text{iff} \quad \mathfrak{A} \vDash \varphi'(\bar{a})\,, \quad \textit{for all } \bar{a} \subseteq A\,.$$

Proof. We can obtain φ' from φ by replacing every atom of the form $t = u$ by the formula $\varepsilon_s(t,u)$, where s is the sort of t and u. □

If we combine all of the above operations, we obtain the notion of a first-order interpretation.

Definition 5.8. Let Σ be an S-sorted signatures and Γ a T-sorted one.

(a) An $\mathrm{FO}_{\kappa\aleph_0}$-*interpretation* from Σ to Γ is a sequence

$$\mathcal{I} = \langle \alpha, (\delta_t)_{t\in T}, (\varepsilon_t)_{t\in T}, (\varphi_\xi)_{\xi\in\Gamma} \rangle$$

of formulae where, for some function $\sigma : T \to S^{<\omega}$,

$$\alpha \in \mathrm{FO}^{0}_{\kappa\aleph_0}[\Sigma], \quad \delta_t \in \mathrm{FO}^{\sigma(t)}_{\kappa\aleph_0}[\Sigma], \quad \varepsilon_t \in \mathrm{FO}^{\sigma(t)\sigma(t)}_{\kappa\aleph_0}[\Sigma],$$

for every relation symbol $R \in \Gamma$ of type $t_0 \dots t_{n-1}$,

$$\varphi_R \in \mathrm{FO}^{\sigma(t_0)\dots\sigma(t_{n-1})}_{\kappa\aleph_0}[\Sigma],$$

and for every function symbol $f \in \Gamma$ of type $t_0 \dots t_{n-1} \to t'$,

$$\varphi_f \in \mathrm{FO}^{\sigma(t_0)\dots\sigma(t_{n-1})\sigma(t')}_{\kappa\aleph_0}[\Sigma].$$

(b) Each $\mathrm{FO}_{\kappa\aleph_0}$-interpretation $\mathcal{I}$ defines an operation on structures as follows. Intuitively, given a Σ-structure $\mathfrak{A}$ the interpretation $\mathcal{I}$ constructs a Γ-structure $\mathcal{I}(\mathfrak{A})$ every element of which is a tuple of elements of $\mathfrak{A}$ and where the relations and functions are defined by the formulae φ_ξ. The formulae δ_t define those tuples that encode elements of sort t and the formula ε_t is used to check whether two such tuples encode the same element. Finally, the *admissibility condition* α says when $\mathcal{I}(\mathfrak{A})$ is defined.

Formally, if $\mathfrak{A}$ is a Σ-structure with $\mathfrak{A} \vDash \alpha$, we define the Γ-structure

$$\mathcal{I}(\mathfrak{A}) := \big\langle (\delta^{\mathfrak{A}}_t)_{t\in T},\ (\varphi^{\mathfrak{A}}_\xi)_{\xi\in\Gamma} \big\rangle / {\approx},$$

which is obtained from the structure $\big\langle (\delta^{\mathfrak{A}}_t)_{t\in T},\ (\varphi^{\mathfrak{A}}_\xi)_{\xi\in\Gamma} \big\rangle$, where the domain of sort t is $\delta^{\mathfrak{A}}_t \subseteq A^{\sigma(t)}$ and every symbol $\xi \in \Gamma$ is interpreted as the relation or function $\varphi^{\mathfrak{A}}_\xi$, by factorising by the congruence relation $\approx$ defined by the $\varepsilon^{\mathfrak{A}}_t$. We regard $\mathcal{I}(\mathfrak{A})$ as undefined if

- $\mathfrak{A} \nvDash \alpha$, or

- $\varepsilon_t^{\mathfrak{A}}$ is not a congruence relation of $\langle (\delta_t^{\mathfrak{A}})_t,\ (\varphi_\xi^{\mathfrak{A}})_\xi \rangle$, or
- there is some function symbol $f \in \Gamma$ such that $\varphi_f^{\mathfrak{A}}$ is not a function.

Example. We construct an interpretation

$$\mathcal{I} = \big\langle \delta(\bar{x}), \varepsilon(\bar{x}, \bar{y}), \varphi_+(\bar{x}, \bar{y}, \bar{z}), \varphi_\cdot(\bar{x}, \bar{y}, \bar{z}) \big\rangle$$

such that

$$\mathcal{I}\langle \mathbb{Z}, +, \cdot, 0, < \rangle \cong \langle \mathbb{Q}, +, \cdot \rangle .$$

We encode a rational number p/q by the pair $\langle p, q \rangle$.

$$\begin{aligned}
\delta(x, x') &:= x' > 0 , \\
\varepsilon(x, x', y, y') &:= x \cdot y' = y \cdot x' , \\
\varphi_+(x, x', y, y', z, z') &:= \varepsilon(z, z',\ x \cdot y' + y \cdot x',\ x' \cdot y') , \\
\varphi_\cdot(x, x', y, y', z, z') &:= \varepsilon(z, z',\ x \cdot y,\ x' \cdot y') .
\end{aligned}$$

Exercise 5.2. Consider the structures $\mathfrak{N} := \langle \mathbb{N}, +, \cdot \rangle$ of arithmetic, $\mathfrak{S} := \langle \mathrm{HF}, \in \rangle$ of hereditary finite sets, and $\mathfrak{M} := \langle 2^{<\omega}, \cdot \rangle$ of finite sequences over $[2]$ with concatenation. Define interpretations $\mathcal{I}_0$, $\mathcal{I}_1$, and $\mathcal{I}_2$ such that

$$\mathfrak{N} = \mathcal{I}_0(\mathfrak{S}) , \quad \mathfrak{S} = \mathcal{I}_1(\mathfrak{M}) , \quad \mathfrak{M} = \mathcal{I}_2(\mathfrak{N}) .$$

For the next lemma, we denote by $\iota_s : \delta_s^{\mathfrak{A}} \to \mathcal{I}(\mathfrak{A})$ the canonical function mapping a tuple to the element it encodes.

Lemma 5.9 (Interpretation Lemma). *Let $\mathcal{I} = \langle \alpha, (\delta_s)_s, (\varepsilon_s)_s, (\varphi_\xi)_\xi \rangle$ be an $\mathrm{FO}_{\kappa\aleph_0}$-interpretation from Σ to Γ.*

(a) *For every formula $\psi(x_0, \dots, x_{m-1}) \in \mathrm{FO}^{\bar{s}}_{\kappa\aleph_0}[\Gamma]$, we can construct an formula $\psi^{\mathcal{I}}(\bar{x}_0, \dots, \bar{x}_{m-1}) \in \mathrm{FO}^{<\omega}_{\kappa\aleph_0}[\Sigma]$ such that*

$$\mathcal{I}(\mathfrak{A}) \vDash \psi(\iota_{s_0}\bar{a}_0, \dots, \iota_{s_{m-1}}\bar{a}_{m-1}) \quad \text{iff} \quad \mathfrak{A} \vDash \psi^{\mathcal{I}}(\bar{a}_0, \dots, \bar{a}_{m-1}) ,$$

for all structures $\mathfrak{A}$ *such that* $\mathcal{I}(\mathfrak{A})$ *is defined and all* $\bar{a}_i \subseteq \delta_{s_i}^{\mathfrak{A}}$.

(b) *There exists a formula* $\chi \in \mathrm{FO}^{\circ}_{\kappa\aleph_0}[\Sigma]$ *such that, for every* Σ*-structure* $\mathfrak{A}$,

$$\mathfrak{A} \vDash \chi \quad \text{iff} \quad \mathcal{I}(\mathfrak{A}) \text{ is defined.}$$

Proof. (a) W.l.o.g. we may assume that ψ is term-reduced. We define $\psi^{\mathcal{I}}$ by induction on ψ. For atomic formulae, we have

$$\begin{aligned} (f x_0 \dots x_{m-1} = y)^{\mathcal{I}} &:= \varphi_f(\bar{x}_0, \dots, \bar{x}_{m-1}, \bar{y}) , \\ (R x_0 \dots x_{m-1})^{\mathcal{I}} &:= \varphi_R(\bar{x}_0, \dots, \bar{x}_{m-1}) , \end{aligned}$$

and, if x and y are of sort s then

$$(x = y)^{\mathcal{I}} := \varepsilon_s(\bar{x}, \bar{y}) .$$

(Note that we assume that every tuple satisfying φ_ξ also satisfies the corresponding δ_s. Otherwise, we have to add the conjunction of all $\delta_{s_i}(\bar{x}_i)$ to the above formulae.) Boolean combinations are left unchanged.

$$\begin{aligned} (\neg\vartheta)^{\mathcal{I}} &:= \neg\vartheta^{\mathcal{I}} , \\ (\bigwedge \Phi)^{\mathcal{I}} &:= \bigwedge \{ \vartheta^{\mathcal{I}} \mid \vartheta \in \Phi \} , \\ (\bigvee \Phi)^{\mathcal{I}} &:= \bigvee \{ \vartheta^{\mathcal{I}} \mid \vartheta \in \Phi \} . \end{aligned}$$

And if y is a variable of sort s, we have to restrict quantifiers over y to δ_s.

$$\begin{aligned} (\exists y \vartheta)^{\mathcal{I}} &:= (\exists \bar{y}.\delta_s(\bar{y}))\vartheta^{\mathcal{I}} , \\ (\forall y \vartheta)^{\mathcal{I}} &:= (\forall \bar{y}.\delta_s(\bar{y}))\vartheta^{\mathcal{I}} . \end{aligned}$$

(b) We can set

$$\chi := \alpha \wedge \bigwedge_{\xi \in \Gamma} \vartheta_\xi$$

where, for each relation symbol $R \in \Gamma$ of type $s_0 \dots s_{n-1}$, the formula

$$\begin{aligned}\vartheta_R := \forall \bar{x}_0 \cdots \bar{x}_{n-1} \bar{y}_0 \cdots \bar{y}_{n-1} \\ \Big(\bigwedge_{i<n} \varepsilon_{s_i}(\bar{x}_i, \bar{y}_i) \to \\ (\varphi_R(\bar{x}_0, \dots, \bar{x}_{n-1}) \leftrightarrow \varphi_R(\bar{y}_0, \dots, \bar{y}_{n-1})\Big)\end{aligned}$$

expresses that the ε_s define a congruence with respect to the relation defined by φ_R and, for each function symbol $f \in \Gamma$ of type $s_0 \dots s_{n-1} \to t$, the formula

$$\begin{aligned}\vartheta_f := {} & \forall \bar{x}_0 \cdots \bar{x}_{n-1} \exists \bar{y} \varphi_f(\bar{x}_0, \dots, \bar{x}_{n-1}, \bar{y}) \\ & \wedge \forall \bar{x}_0 \cdots \bar{x}_{n-1} \bar{y}_0 \cdots \bar{y}_{n-1} \bar{u} \bar{v} \\ & \Big(\Big(\bigwedge_{i<n} \varepsilon_{s_i}(\bar{x}_i, \bar{y}_i) \wedge \varphi_f(\bar{x}_0, \dots, \bar{x}_{n-1}, \bar{u}) \\ & \wedge \varphi_f(\bar{y}_0, \dots, \bar{y}_{n-1}, \bar{v}) \Big) \to \varepsilon_t(\bar{u}, \bar{v}) \Big)\end{aligned}$$

says that φ_f defines a function and the ε_s define a congruence with respect to this function. □

The general scheme of these constructions is summarised in the following definition.

Definition 5.10. Let $\langle L_0, \mathcal{K}_0, \vDash \rangle$ and $\langle L_1, \mathcal{K}_1, \vDash \rangle$ be logics.

(a) A *morphism* from L_0 to L_1 is a pair $\langle \alpha, \beta \rangle$ of functions $\alpha : L_0 \to L_1$ and $\beta : \mathcal{K}_1 \to \mathcal{K}_0$ such that

$$\mathfrak{J} \vDash \alpha(\varphi) \quad \text{iff} \quad \beta(\mathfrak{J}) \vDash \varphi \,, \qquad \text{for all } \varphi \in L_0 \text{ and } \mathfrak{J} \in \mathcal{K}_1 \,.$$

The category consisting of all logics and these morphisms is called $\mathfrak{Logic}$.

(b) An *embedding* is a morphism $\langle \alpha, \beta \rangle : L_0 \to L_1$ where β is surjective.

(c) A *comorphism* from L_0 to L_1 is a morphism $\langle \alpha, \beta \rangle : L_1 \to L_0$.

(d) By abuse of terminology we call a function $\alpha : L_0 \to L_1$ a morphism if there exists a function $\beta : \mathcal{K}_1 \to \mathcal{K}_0$ such that the pair $\langle \alpha, \beta \rangle$ forms a morphism $L_0 \to L_1$. Similarly, we call $\beta : \mathcal{K}_0 \to \mathcal{K}_1$ a comorphism if there is some $\alpha : L_1 \to L_0$ such that $\langle \alpha, \beta \rangle$ is a comorphism $L_0 \to L_1$.

Remark. The only difference between a morphism and a comorphism is the direction of the arrow. We will use the former term if we want to stress the translation of formulae, while the latter term is used when we are mainly interested in the operation on structures.

Example. Each of the operations introduced in this section induces a comorphism. For instance, we have seen in Lemma 5.1 that the reduct operation $r : \mathfrak{A} \mapsto \mathfrak{A}|_\Sigma$ induces the comorphism

$$\langle i, r \rangle : \mathrm{FO}_{\kappa\aleph_0}[\Gamma, X] \to \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X],$$

where $i : \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X] \to \mathrm{FO}_{\kappa\aleph_0}[\Gamma, X]$ is the inclusion map.

In the case of interpretations we face a minor technical difficulty since these are partial operations. An $\mathrm{FO}_{\kappa\aleph_0}$-interpretation $\mathcal{I}$ from Σ to Γ induces a comorphism $L \to \mathrm{FO}_{\kappa\aleph_0}[\Gamma, X]$ where L is not $\mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$ but the sublogic $\langle \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X], \mathcal{C}, \vDash \rangle$, where the class $\mathcal{C} \subseteq \mathrm{Str}[\Sigma]$ of interpretations consists of those Σ-structures $\mathfrak{A}$ such that $\mathcal{I}(\mathfrak{A})$ is defined.

Exercise 5.3. Prove that a morphism $\langle \alpha, \beta \rangle : L_0 \to L_1$ is a monomorphism if, and only if, α is injective and β is surjective. Show that it is an epimorphism if, and only if, α is surjective and β is injective.

Lemma 5.11. *Let $\langle \alpha, \beta \rangle : L_0 \to L_1$ be a morphism of logics.*

(a) *$\langle \alpha, \beta \rangle$ is a monomorphism if, and only if, it has a left inverse.*

(b) *$\langle \alpha, \beta \rangle$ is an epimorphism if, and only if, it has a right inverse.*

Lemma 5.12. *Let $\langle \alpha, \beta \rangle : L_0 \to L_1$ be a morphism of logics, $\Phi \subseteq L_0$, $\varphi, \psi \in L_0$, and $\mathfrak{J}$ an L_1-interpretation.*

(a) *$\varphi \vDash \psi$ implies $\alpha(\varphi) \vDash \alpha(\psi)$.*

(b) *If Φ is inconsistent then so is $\alpha[\Phi]$.*

(c) $\mathrm{Th}_{L_0}(\beta(\mathfrak{J})) = \alpha^{-1}[\mathrm{Th}_{L_1}(\mathfrak{J})]$.
(d) $\mathrm{Mod}_{L_1}(\alpha[\Phi]) = \beta^{-1}[\mathrm{Mod}_{L_0}(\Phi)]$.

Proof. (a) For every L_1-interpretation $\mathfrak{J}$, we have the following chain of implications:

$$\begin{aligned} \mathfrak{J} \vDash \alpha(\varphi) \quad &\Rightarrow \quad \beta(\mathfrak{J}) \vDash \varphi \\ &\Rightarrow \quad \beta(\mathfrak{J}) \vDash \psi \quad \Rightarrow \quad \mathfrak{J} \vDash \alpha(\psi) . \end{aligned}$$

It follows that $\alpha(\varphi) \vDash \alpha(\psi)$.

(b) Suppose that $\alpha[\Phi]$ has a model $\mathfrak{J}$. Then $\mathfrak{J} \vDash \alpha[\Phi]$ implies that $\beta(\mathfrak{J}) \vDash \Phi$. Hence, Φ is satisfiable.

(c) For a formula $\varphi \in L_0$ and an L_1-interpretation $\mathfrak{J}$, we have

$$\beta(\mathfrak{J}) \vDash \varphi \quad \text{iff} \quad \mathfrak{J} \vDash \alpha(\varphi) \quad \text{iff} \quad \varphi \in \alpha^{-1}[\mathrm{Th}_{L_1}(\mathfrak{J})] .$$

(d) By definition of a morphism, we have

$$\mathfrak{J} \vDash \alpha[\Phi] \quad \text{iff} \quad \beta(\mathfrak{J}) \vDash \Phi \quad \text{iff} \quad \mathfrak{J} \in \beta^{-1}[\mathrm{Mod}_{L_1}(\Phi)] . \qquad \square$$

Corollary 5.13. *Let $\langle \alpha, \beta \rangle : L_0 \to L_1$ be a comorphism of logics and suppose that $\mathfrak{J}_0, \mathfrak{J}_1$ are L_0-interpretations.*

$$\mathfrak{J}_0 \equiv_{L_0} \mathfrak{J}_1 \quad \textit{implies} \quad \beta(\mathfrak{J}_0) \equiv_{L_1} \beta(\mathfrak{J}_1) .$$

Proof. The claim follows immediately from Lemma 5.12 (c). $\square$

Every monomorphism of logics is an embedding. Statement (a) of the following lemma states that, conversely, every embedding is a monomorphism 'up to logical equivalence'.

Lemma 5.14. *Let $\langle \alpha, \beta \rangle : L_0 \to L_1$ be an embedding of logics, $\Phi \subseteq L_0$, and $\varphi, \psi \in L_0$ formulae.*

(a) $\varphi \vDash \psi$ iff $\alpha(\varphi) \vDash \alpha(\psi)$.
(b) $\mathrm{Mod}_{L_0}(\Phi) = \beta[\mathrm{Mod}_{L_1}(\alpha[\Phi])]$.

Proof. (a) We have already seen in Lemma 5.12 (a) that $\varphi \vDash \psi$ implies $\alpha(\varphi) \vDash \alpha(\psi)$. Conversely, suppose that $\alpha(\varphi) \vDash \alpha(\psi)$ and let $\mathfrak{J}_0$ be an L_0-interpretation. By assumption, there is some L_1-interpretation $\mathfrak{J}_1$ with $\beta(\mathfrak{J}_1) = \mathfrak{J}_0$. Hence, we have

$$\mathfrak{J}_0 \vDash \varphi \quad \Rightarrow \quad \mathfrak{J}_1 \vDash \alpha(\varphi) \quad \Rightarrow \quad \mathfrak{J}_1 \vDash \alpha(\psi) \quad \Rightarrow \quad \mathfrak{J}_0 \vDash \psi\,.$$

It follows that $\varphi \vDash \psi$.

(b) By Lemmas A2.1.10 and 5.12 (d), it follows that

$$\beta[\mathrm{Mod}_{L_1}(\alpha[\Phi])] = \beta[\beta^{-1}[\mathrm{Mod}_{L_0}(\Phi)]] = \mathrm{Mod}_{L_0}(\Phi)\,. \qquad \square$$

6. Extensions of first-order logic

Lindström quantifiers

First-order logic seems to be ill-suited to talk about cardinalities. To express that there are infinitely many elements we had to use an infinite set of formulae, and we will see in Lemma C2.4.9 that, even if we allow infinitely many formulae, we cannot express that something is finite.

To obtain a logic where these things can be expressed, we add to ordinary first-order logic a *cardinality quantifier* $\exists^\lambda$ with the meaning of 'there are at least λ many'.

Definition 6.1. By $\mathrm{FO}_{\kappa\aleph_0}(\exists^\lambda)[\Sigma, X]$ we denote the logic obtained from $\mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$ by adding the syntax rule:

- if $\varphi \in \mathrm{FO}_{\kappa\aleph_0}(\exists^\lambda)[\Sigma, X \cup \{x\}]$ then $\exists^\lambda x\varphi \in \mathrm{FO}_{\kappa\aleph_0}(\exists^\lambda)[\Sigma, X]$.

We define the semantics of this quantifier by

$$\mathfrak{A} \vDash \exists^\lambda x\varphi[\beta] \quad :\text{iff} \quad \big|\big\{\, a \in A \;\big|\; \mathfrak{A} \vDash \varphi[\beta[x/a]] \,\big\}\big| \geq \lambda\,.$$

Example. We can axiomatise the order $\langle \omega, < \rangle$ up to isomorphism by the formula

$$\begin{aligned} \forall x \neg(x < x) \wedge \forall x \forall y \forall z(x < y \wedge y < z \rightarrow x < z) \\ \wedge \forall x \forall y(x < y \vee x = y \vee y < x) \\ \wedge \forall x \exists y(x < y) \\ \wedge \forall x \neg \exists^{\aleph_0} y(y < x) . \end{aligned}$$

Another property that infinitary first-order logic is unable to express is well-foundedness. As above, we can introduce a new quantifier expressing that a definable relation is a well-order. This logic will play an important role in Section C5.6.

Definition 6.2. Let $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{wo})$ be the extension of $\mathrm{FO}_{\kappa\aleph_0}$ by the *well-ordering quantifier* W whose semantics is given by

$$\mathfrak{A} \models \mathsf{W}\bar{x}\bar{y}\varphi(\bar{x}, \bar{y}, \bar{c}) \quad \text{iff} \quad \text{the relation } \varphi^{\mathfrak{A}}(\bar{x}, \bar{y}, \bar{c}) \text{ is a well-ordering of its field.}$$

Note that the quantifier W cannot be used to express 'there exists a well-order'. We can only check whether some definable relation is a well-order.

Generalising the above examples we can define extensions of (infinitary) first-order logic by quantifiers for any given property.

Definition 6.3. Let $\Gamma = \{R_0, \ldots, R_n\}$ be a finite relational signature and $\mathcal{K}$ a class of Γ-structures. The *Lindström quantifier* for $\mathcal{K}$ is of the form $\mathsf{Q}_{\mathcal{K}}\bar{x}_0 \ldots \bar{x}_n \varphi_0(\bar{x}_0, \bar{z}) \ldots \varphi_n(\bar{x}_n, \bar{z})$. The semantics of such a formula is defined by

$$\begin{aligned} &\mathfrak{A} \models \mathsf{Q}_{\mathcal{K}}\bar{x}_0 \ldots \bar{x}_n \varphi_0(\bar{x}_0, \bar{c}) \ldots \varphi_n(\bar{x}_n, \bar{c}) \\ &:\text{iff } \langle A, \varphi_0(\bar{x}_0, \bar{c})^{\mathfrak{A}}, \ldots, \varphi_n(\bar{x}_n, \bar{c})^{\mathfrak{A}} \rangle \in \mathcal{K} . \end{aligned}$$

Example. (a) The cardinality quantifier $\exists^\lambda$ is the quantifier $\mathsf{Q}_{\mathcal{K}}$ where

$$\mathcal{K} := \{ \langle A, P \rangle \mid A \text{ a set}, P \subseteq A, |P| \geq \lambda \} .$$

(b) The *cardinality comparison quantifier* is defined by the class

$$\mathcal{K} := \{ \langle A, P, Q \rangle \mid |P| = |Q| \} .$$

(c) The well-ordering quantifier W is defined by the class

$$\mathcal{K} := \{ \langle A, R \rangle \mid R \text{ is a well-order on its field} \} .$$

Second-order logic

In second-order logic we extend first-order logic by variables for relations and functions and we allow quantification over such variables. When we equip each variable with a type, the set of variables becomes a signature where the constant symbols play the role of the first-order variables. This particular point of view makes the definition of syntax and semantics much more streamlined. We could also have adopted this convention for the definition of first-order logic. But for expositionary reasons we have refrained from doing so.

Definition 6.4. Let Σ and Ξ be S-sorted signatures. The set $\mathrm{SO}_{\kappa\aleph_0}[\Sigma, \Xi]$ of *infinitary second-order formulae* is the smallest set of terms satisfying the following closure conditions:

- If $t_0, t_1 \in T[\Sigma \cup \Xi, \varnothing]$ are terms of the same sort, we have $t_0 = t_1 \in \mathrm{SO}_{\kappa\aleph_0}[\Sigma, \Xi]$.
- If $R \in \Sigma \cup \Xi$ is of type $s_0 \dots s_{n-1}$ and $t_i \in T_{s_i}[\Sigma \cup \Xi, \varnothing]$, for $i < n$, then $Rt_0 \dots t_{n-1} \in \mathrm{SO}_{\kappa\aleph_0}[\Sigma, \Xi]$.
- If $\varphi \in \mathrm{SO}_{\kappa\aleph_0}[\Sigma, \Xi]$, then $\neg\varphi \in \mathrm{SO}_{\kappa\aleph_0}[\Sigma, \Xi]$.
- If $\Phi \subseteq \mathrm{SO}_{\kappa\aleph_0}[\Sigma, \Xi]$ and $|\Phi| < \kappa$, then $\bigwedge \Phi$, $\bigvee \Phi \in \mathrm{SO}_{\kappa\aleph_0}[\Sigma, \Xi]$.
- If $\varphi \in \mathrm{SO}_{\kappa\aleph_0}[\Sigma, \Xi \cup \{\xi\}]$, then $\exists\xi\varphi$, $\forall\xi\varphi \in \mathrm{SO}_{\kappa\aleph_0}[\Sigma, \Xi]$.

We define *monadic second-order logic* $\mathrm{MSO}_{\kappa\aleph_0}[\Sigma, \Xi]$ as the restriction of $\mathrm{SO}_{\kappa\aleph_0}[\Sigma, \Xi]$ were we allow only constant symbols and *unary* relation symbols in the variable signature Ξ.

For a $(\Sigma \cup \Xi)$-structure $\mathfrak{A}$ and an $\mathrm{SO}_{\kappa\aleph_0}[\Sigma, \Xi]$-formula φ, we define the satisfaction relation $\mathfrak{A} \vDash \varphi$ by induction on φ.

$$
\begin{aligned}
&\mathfrak{A} \vDash t_0 = t_1 &&:\text{iff} \quad t_0^{\mathfrak{A}} = t_1^{\mathfrak{A}} ,\\
&\mathfrak{A} \vDash Rt_0 \dots t_{n-1} &&:\text{iff} \quad \langle t_0^{\mathfrak{A}}, \dots, t_{n-1}^{\mathfrak{A}} \rangle \in R^{\mathfrak{A}} ,\\
&\mathfrak{A} \vDash \neg\varphi &&:\text{iff} \quad \mathfrak{A} \nvDash \varphi ,\\
&\mathfrak{A} \vDash \bigvee \Phi &&:\text{iff} \quad \text{there is some } \varphi \in \Phi \text{ such that } \mathfrak{A} \vDash \varphi ,\\
&\mathfrak{A} \vDash \bigwedge \Phi &&:\text{iff} \quad \mathfrak{A} \vDash \varphi \text{ for all } \varphi \in \Phi ,\\
&\mathfrak{A} \vDash \exists\xi\varphi &&:\text{iff} \quad \text{there is some relation or function } \xi^{\mathfrak{A}}\\
& && \qquad\quad \text{such that } \langle \mathfrak{A}, \xi^{\mathfrak{A}} \rangle \vDash \varphi ,\\
&\mathfrak{A} \vDash \forall\xi\varphi &&:\text{iff} \quad \langle \mathfrak{A}, \xi^{\mathfrak{A}} \rangle \vDash \varphi \text{ for all suitable relations or}\\
& && \qquad\quad \text{functions } \xi^{\mathfrak{A}}.
\end{aligned}
$$

Example (Peano Axioms). The structure $\langle \omega, s, 0 \rangle$, where $s : n \mapsto n + 1$ is the successor function, can be axiomatised in monadic second-order logic up to isomorphism by the *Peano Axioms*.

$$
\begin{aligned}
&\forall x(sx \neq 0) ,\\
&\forall x \forall y(sx = sy \rightarrow x = y) ,\\
&\forall Z[Z0 \wedge \forall x(Zx \rightarrow Zsx) \rightarrow \forall x Zx] .
\end{aligned}
$$

The third axiom which states the induction principle is not first-order.

Example. (a) The class of all well-orders can be axiomatised by the MSO-formulae

$$
\begin{aligned}
&\forall x \forall y(x \leq y \wedge y \leq x \leftrightarrow x = y) ,\\
&\forall x \forall y \forall z(x \leq y \wedge y \leq z \rightarrow x \leq z) ,\\
&\forall x \forall y(x \leq y \vee y \leq x) ,\\
&\forall Z[\exists x Zx \rightarrow (\exists x.Zx)(\forall y.Zy)(x \leq y)] ,
\end{aligned}
$$

which express that $\leq$ is a linear order such that every nonempty set Z has a minimal element.

(b) Let $\langle V, E \rangle$ be a graph. The transitive closure of the relation E can be defined by the monadic second-order formula

$$\varphi(x, y) := \forall Z[Zx \wedge \forall u \forall v(Zu \wedge Euv \rightarrow Zv) \rightarrow Zy] .$$

Consequently, we can express that a graph is strongly connected by

$$\psi := \forall x \forall y \varphi(x, y) .$$

(c) Let $\varphi(x)$ and $\psi(x)$ be second-order formulae. We can express that the sets defined by φ and ψ have the same cardinality by the second-order formula

$$\begin{aligned} \exists f\big[&(\forall x.\varphi(x))(\forall y.\varphi(y))(x \neq y \rightarrow fx \neq fy) \\ &\wedge (\forall x.\varphi(x))(\exists y.\psi(y))(fx = y) \\ &\wedge (\forall x.\psi(x))(\exists y.\varphi(y))(fy = x)\big] \end{aligned}$$

which states that there exists a bijection between these sets.

Logical systems

We have already introduced several logics and we will define some more below. To facilitate a uniform treatment let us define a general framework for the kind of logic we are interested in. We have two basic requirements. Firstly, the logic should talk about structures and, secondly, it should be well-behaved with respect to reducts and expansions of signatures. Like in the first-order case we will therefore consider not a single logic but a family of them, one logic for each signature. We start by giving a general definition of a family of logics.

Definition 6.5. Let $\mathfrak{S}$ be a category. A *logical system* parametrised by $\mathfrak{S}$ is a functor $\mathcal{L} : \mathfrak{S} \rightarrow \mathfrak{Logic}$. To each logical system $\mathcal{L}$ we associate a covariant functor L and a contravariant functor $\mathcal{C}$ such that

$$\begin{aligned} \mathcal{L}[s] &= \langle L[s], \mathcal{C}[s], \vDash_s \rangle , &&\text{for } s \in \mathfrak{S} , \\ \mathcal{L}[f] &= \langle L[f], \mathcal{C}[f] \rangle , &&\text{for } f \in \mathfrak{S}(s, s') . \end{aligned}$$

L is called the *syntax functor* of $\mathcal{L}$ and $\mathcal{C}$ is the *semantics functor.*

Remark. (a) An alternative, more concrete definition of a logical system would be as follows. A logical system consists of a covariant functor $L : \mathfrak{S} \to \mathfrak{L}$, a contravariant functor $\mathcal{C} : \mathfrak{S} \to \mathfrak{Int}$, and a family $(\vDash_s)_{s \in \mathfrak{S}}$ of binary relations $\vDash_s \subseteq \mathcal{C}[s] \times L[s]$ that satisfy the following conditions:

- $\langle L[s], \mathcal{C}[s], \vDash_s \rangle$ is a logic, for all $s \in \mathfrak{S}$.
- For every morphism $f : s \to t$ of $\mathfrak{S}$, all formulae $\varphi \in L[s]$, and each interpretation $\mathfrak{J} \in \mathcal{C}[t]$, we have

$$\mathcal{C}[f](\mathfrak{J}) \vDash_s \varphi \quad \text{iff} \quad \mathfrak{J} \vDash_t L[f](\varphi) .$$

Note that the second condition is a generalisation of the property of terms stated in Lemma B3.1.16.

(b) Usually the category $\mathfrak{S}$ specifies a signature Σ and a set of variables X, and $\mathcal{C}[\Sigma, X]$ is the class of all pairs $\langle \mathfrak{A}, \beta \rangle$ where $\mathfrak{A}$ is a Σ-structure and β a variable assignment for the variables in X. In fact, we will mostly deal with logics without free variables where the interpretations consists of only a structure (see Definition 6.7 below).

Example. We define a logical system based on Zariski logic. The category $\mathfrak{S}$ of parameters consists of all pairs $\langle \mathfrak{K}, X \rangle$ where $\mathfrak{K}$ is a field and X a set of variables. If $\mathfrak{L}$ is an extension of $\mathfrak{K}$ then $\mathfrak{S}(\langle \mathfrak{K}, X \rangle, \langle \mathfrak{L}, Y \rangle)$ consists of all functions $f : X \to Y$. If $\mathfrak{L}$ is not an extension of $\mathfrak{K}$ then there are no morphisms $\langle \mathfrak{K}, X \rangle \to \langle \mathfrak{L}, Y \rangle$.

The logical system maps a parameter $\langle \mathfrak{K}, X \rangle \in \mathfrak{S}$ to the Zariski logic $\mathrm{ZL}[\mathfrak{K}, X]$. Each morphism $f : \langle \mathfrak{K}, X \rangle \to \langle \mathfrak{L}, Y \rangle$ of $\mathfrak{S}$ is mapped to the morphism $\langle \alpha, \beta \rangle : \mathrm{ZL}[\mathfrak{K}, X] \to \mathrm{ZL}[\mathfrak{L}, Y]$ where

- α maps a polynomial $p(\bar{x}) \in \mathfrak{K}[X]$ to $p(f(\bar{x})) \in \mathfrak{L}[Y]$ and
- β maps a variable assignment $\gamma \in \mathfrak{M}^Y$ to $\gamma \circ f \in \mathfrak{M}^X$.

Note that $\langle \alpha, \beta \rangle$ is indeed a morphism since

$$\gamma \circ f \vDash p(x_0, \dots, x_{n-1}) \quad \text{iff} \quad \gamma \vDash p(f(x_0), \dots, f(x_{n-1})) .$$

Recall the categories $\mathfrak{Sig}$, $\mathfrak{SigVar}$, and $\mathfrak{Str}$ introduced in Section B3.1.

Definition 6.6. By $\mathrm{FO}_{\kappa\aleph_0}$ we denote the logical system $\mathfrak{SigVar} \to \mathfrak{Logic}$ with

$$\langle \Sigma, X \rangle \mapsto \langle \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X], \mathrm{Str}[\Sigma, X], \models \rangle ,$$

and $\mathrm{FO}^{\bar{s}}_{\kappa\aleph_0} : \mathfrak{Sig} \to \mathfrak{Logic}$ is the subsystem with

$$\Sigma \mapsto \langle \mathrm{FO}^{\bar{s}}_{\kappa\aleph_0}[\Sigma], \mathrm{Str}, \models \rangle .$$

Exercise 6.1. Prove that $\mathrm{FO}_{\kappa\aleph_0}$ and $\mathrm{FO}^{\bar{s}}_{\kappa\aleph_0}$ are indeed logical systems.

Exercise 6.2. Let $\mathcal{L} : \mathfrak{S} \to \mathfrak{Logic}$ be a logical system with

$$\begin{aligned} \mathcal{L}(s) &= \langle L[s], \mathcal{C}[s], \models_s \rangle , && \text{for } s \in \mathfrak{S} , \\ \mathcal{L}(f) &= \langle \alpha_f, \beta_f \rangle , && \text{for } f \in \mathfrak{S}(s,t) . \end{aligned}$$

Show that the function $\mathcal{L}^{\mathrm{op}} : \mathfrak{S}^{\mathrm{op}} \to \mathfrak{Logic}$ defined by

$$\begin{aligned} \mathcal{L}^{\mathrm{op}}(s) &:= \langle \mathcal{C}[s], L[s], (\models_s)^{-1} \rangle , && \text{for } s \in \mathfrak{S} , \\ \mathcal{L}^{\mathrm{op}}(f) &:= \langle \beta_f, \alpha_f \rangle , && \text{for } f \in \mathfrak{S}(s,t) \end{aligned}$$

is a logical system.

We are mainly interested in logical systems that, like first-order logic, talk about structures.

Definition 6.7. An *algebraic logic* is a logical system $\mathcal{L} : \mathfrak{Sig} \to \mathfrak{Logic}$ parametrised by $\mathfrak{Sig}$ such that

- the semantics functor is the canonical functor $\mathrm{Str} : \mathfrak{Sig} \to \mathfrak{Str}$ and
- every logic $L[\Sigma]$ is invariant under isomorphisms, that is,

$$\mathfrak{A} \cong \mathfrak{B} \quad \text{implies} \quad \mathfrak{A} \equiv_{L[\Sigma]} \mathfrak{B} , \quad \text{for all } \mathfrak{A}, \mathfrak{B} \in \mathrm{Str}[\Sigma] .$$

Example. We will prove in Lemma C2.1.3 (c) that first-order logic is invariant under isomorphisms. Consequently, $\mathrm{FO}^{0}_{\kappa\aleph_0}$ is algebraic. Clearly, $\mathrm{FO}^{\alpha}_{\kappa\aleph_0}$ is not, for $\alpha > 0$, since the interpretations are not structures.

Remark. Note that it follows immediately from the definition of an algebraic logic that the reduct operation $\mathfrak{A} \mapsto \mathfrak{A}|_\Sigma$ is a comorphism $L[\Gamma] \to L[\Sigma]$, for every algebraic logic L.

When defining the semantics of second-order logic we have treated the variables as symbols of a signature. This trick can be used to simulate free variables in every algebraic logic.

Definition 6.8. Let L be an algebraic logic, Σ a signature, and X a set of variables disjoint from Σ. We set

$$L[\Sigma, X] := L[\Sigma \cup X],$$

where we regard the elements of X as constant symbols. If $\mathfrak{A}$ is a Σ-structure and $\beta : X \to A$ a variable assignment, we define

$$\mathfrak{A} \vDash \varphi[\beta] \quad :\text{iff} \quad \mathfrak{A}_\beta \vDash \varphi,$$

where $\mathfrak{A}_\beta$ is the $(\Sigma \cup X)$-expansion of $\mathfrak{A}$ where we assign to the additional constants $x \in X$ the value $x^{\mathfrak{A}_\beta} := \beta(x)$.

We define $\varphi^{\mathfrak{A}}$, $\text{free}(\varphi)$, $\mathfrak{A} \vDash \varphi(\bar{a})$, and $L^{\bar{s}}[\Sigma]$ in the same way as for first-order logic.

Lindenbaum algebras

Usually we are only interested in the expressive power of a logic and, hence, we will not distinguish between equivalent formulae. To this end we associate with every logic L a partial order that reflects the structural properties of L while abstracting away from the concrete syntax. We have seen in Lemma 3.2 that the entailment relation $\vDash$ is a preorder. If we identify equivalent formulae, we obtain the partial order $\langle L, \vDash \rangle / \equiv$. In this way we can define a functor $\mathfrak{Logic} \to \mathfrak{PO}$ where $\mathfrak{PO}$ is the category of all partial orders with homomorphisms.

Definition 6.9. The *Lindenbaum functor* $\mathfrak{Lb} : \mathfrak{Logic} \to \mathfrak{PO}$ is defined by

$$\mathfrak{Lb}(L) := \langle L, \vDash \rangle / \equiv, \qquad \text{for } L \in \mathfrak{Logic},$$
$$\mathfrak{Lb}(\mu)([\varphi]_\equiv) := [\alpha(\varphi)]_\equiv, \quad \text{for } \mu = \langle \alpha, \beta \rangle \in \mathfrak{Logic}(L_0, L_1).$$

The partial order $\mathfrak{Lb}(L)$ is called the *Lindenbaum algebra* of L.

Remark. Note that it follows by Lemma 5.12 (a) that the image $\mathfrak{Lb}(\mu)$ of a morphism $\mu : L_0 \to L_1$ is well-defined and that it is indeed a homomorphism of partial orders.

Example. (a) Let $\mathfrak{K}$ be an algebraically closed field. For Zariski logic $\mathrm{ZL}[\mathfrak{K}, X]$, we have shown that

$$p \equiv q \quad \text{iff} \quad p^m = aq^n \quad \text{for some } a \in K \text{ and } m, n < \omega\,.$$

The Lindenbaum algebra $\mathfrak{Lb}(\mathrm{ZL}[\mathfrak{K}, X])$ is an upper semilattice where

$$\top = [0]_\equiv\,, \quad \bot = [1]_\equiv\,, \quad \text{and} \quad [p]_\equiv \sqcup [q]_\equiv = [pq]_\equiv\,.$$

(b) Let $\mathfrak{B}$ be a boolean algebra. The Lindenbaum algebra $\mathfrak{Lb}(\mathrm{BL}(\mathfrak{B}))$ is isomorphic to $\mathfrak{B}$ since, for $a, b \in B$,

$$a \equiv b \quad \text{implies} \quad a = b\,.$$

Lemma 6.10. *Let $\mu : L_0 \to L_1$ be a morphism of logics.*

(a) *If μ is an epimorphism then so is $\mathfrak{Lb}(\mu)$.*

(b) *If μ is an embedding then so is $\mathfrak{Lb}(\mu)$.*

Proof. Suppose that $\mu = \langle \alpha, \beta \rangle$.

(a) Let $[\varphi]_\equiv \in \mathfrak{Lb}(L_1)$. The map α is surjective since μ is an epimorphism. Consequently, there is some $\psi \in L_0$ with $\alpha(\psi) = \varphi$. Hence, $\mathfrak{Lb}(\mu)([\psi]_\equiv) = [\varphi]_\equiv$, as desired.

(b) follows immediately from Lemma 5.14 (a). □

Definition 6.11. Let L be a logic and $\varphi, \psi \in L$ formulae.

(a) A *negation* of φ is a formula $\vartheta \in L$ such that, for all L-interpretations $\mathfrak{I}$, we have

$$\mathfrak{I} \vDash \vartheta \quad \text{iff} \quad \mathfrak{I} \nvDash \varphi\,.$$

If φ has negations, we fix one and denote it by $\neg\varphi$.

(b) A *disjunction* of φ and ψ is a formula $\vartheta \in L$ such that, for all L-interpretations $\mathfrak{J}$, we have

$$\mathfrak{J} \vDash \vartheta \quad \text{iff} \quad \mathfrak{J} \vDash \varphi \text{ or } \mathfrak{J} \vDash \psi \text{ or both.}$$

If disjunctions of φ and ψ exist, we fix one and denote it by $\varphi \vee \psi$.

(c) A *conjunction* of φ and ψ is a formula $\vartheta \in L$ such that, for all L-interpretations $\mathfrak{J}$, we have

$$\mathfrak{J} \vDash \vartheta \quad \text{iff} \quad \mathfrak{J} \vDash \varphi \text{ and } \mathfrak{J} \vDash \psi\,.$$

If conjunctions of φ and ψ exist, we fix one and denote it by $\varphi \wedge \psi$.

(d) We say that L is *closed under negation, disjunction,* or *conjunction* if all L-formulae have, respectively, negations, disjunctions, or conjunctions. We call L *boolean closed* if L is closed under all three operations.

Remark. (a) Note that $\neg\varphi$, $\varphi \vee \psi$, and $\varphi \wedge \psi$ are only determined up to logical equivalence, but they are unique when regarded as elements of $\mathfrak{Lb}(L)$.

(b) If L is closed under conjunction and disjunction, the Lindenbaum algebra $\mathfrak{Lb}(L)$ is a lattice where

$$[\varphi]_\equiv \sqcap [\psi]_\equiv = [\varphi \wedge \psi]_\equiv \quad \text{and} \quad [\varphi]_\equiv \sqcup [\psi]_\equiv = [\varphi \vee \psi]_\equiv\,.$$

Exercise 6.3. Define a logic L such that $\mathfrak{Lb}(L)$ is a boolean algebra but L is closed under neither negation, nor disjunction, nor conjunction.

Lemma 6.12. *Let $\langle L, \mathcal{K}, \vDash\rangle$ be a logic.*

(a) *If L is closed under conjunction and disjunction then $\mathfrak{Lb}(L)$ is a distributive lattice.*

(b) *If L is boolean closed then $\mathfrak{Lb}(L)$ is a boolean algebra.*

Proof. (a) $\mathfrak{Lb}(L)$ is clearly a lattice if it has the above closure properties. To show that it is distributive note that the function

$$f : \mathfrak{Lb}(L) \to \wp(\mathcal{K}) : [\varphi]_\equiv \mapsto \mathrm{Mod}_L(\varphi)$$

is an embedding of $\mathfrak{Lb}(L)$ into a power-set lattice and such lattices are always distributive.

(b) If L is boolean closed, it contains tautologies $\varphi \vee \neg\varphi$ and unsatisfiable formulae $\varphi \wedge \neg\varphi$. Hence, $\mathfrak{Lb}(L)$ forms a boolean algebra. □

When investigating a logical theory T we usually are only interested in the class of models of T. In these cases we can restrict the logic by removing all interpretations that do not satisfy T.

Definition 6.13. Let $\langle L, \mathcal{K}, \vDash\rangle$ be a logic, $\Phi \subseteq L$ a set of formulae, and let $i : \Phi \to L$ and $j : \mathrm{Mod}_L(\Phi) \to \mathcal{K}$ be the corresponding inclusion maps.

(a) The *restriction* of L to Φ is the logic

$$L|_\Phi := \langle \Phi, \mathcal{K}, \vDash\rangle\,,$$

where the set of formulae is restricted to Φ. The morphism

$$\langle i, \mathrm{id}_\mathcal{K}\rangle : L|_\Phi \to L$$

is the *inclusion morphism* associated with Φ and L.

(b) The *localisation* of L to Φ is the logic

$$L/\Phi := \langle L,\ \mathrm{Mod}_L(\Phi),\ \vDash\rangle\,,$$

where the class of interpretations is restricted to those satisfying Φ. The morphism

$$\langle \mathrm{id}_L, j\rangle : L \to L/\Phi$$

is the *localisation morphism* associated with Φ and L. We define the relations

$$\begin{aligned} \varphi \vDash_\Phi \psi \quad &:\text{iff} \quad \Phi \cup \{\varphi\} \vDash \psi\,, \\ \varphi \equiv_\Phi \psi \quad &:\text{iff} \quad \varphi \equiv \psi \text{ modulo } \Phi\,. \end{aligned}$$

(c) If L is an algebraic logic and $\Phi \subseteq L^\circ[\Sigma]$ then we set

$$L^{\bar{s}}/\Phi := L^{\bar{s}}[\Sigma]/\Phi\,.$$

The next lemma and its corollary state that the restriction and the localisation of a logic yield something like 'short exact sequences' of logics and Lindenbaum algebras

$$L|_\Phi \to L \to L/\Phi \quad \text{and} \quad \mathfrak{Lb}(L|_\Phi) \to \mathfrak{Lb}(L) \to \mathfrak{Lb}(L/\Phi)\,.$$

Lemma 6.14. *Let $\langle L, \mathcal{K}, \vDash \rangle$ be a logic and $\Phi \subseteq L$ a set of formulae.*

(a) *The inclusion morphism $i : L|_\Phi \to L$ is a monomorphism of logics.*

(b) *The localisation morphism $\lambda : L \to L/\Phi$ is an epimorphism of logics.*

Corollary 6.15. *Let L be a logic and $\Phi \subseteq L$.*

(a) *There exists an embedding $\mathfrak{Lb}(L|_\Phi) \to \mathfrak{Lb}(L)$.*

(b) *There exists a surjective homomorphism $\mathfrak{Lb}(L) \to \mathfrak{Lb}(L/\Phi)$.*

Proof. The claims follow from Lemmas 6.14 and 6.10. □

We can describe the entailment relation of a localisation as follows.

Lemma 6.16. *Let L be a logic and $T \subseteq L$.*

(a) $\varphi \vDash \psi$ *in* L/T iff $\varphi \vDash_T \psi$ *in* L.

(b) $\mathfrak{Lb}(L/T) = \langle L, \vDash_T \rangle / \equiv_T$.

Proof. (a) We have $\varphi \vDash \psi$ in L/T if, and only if, every model of T that satisfies φ also satisfies ψ. This is equivalent to $T \cup \{\varphi\} \vDash \psi$.

(b) follows immediately from (a). □

C2. *Elementary substructures and embeddings*

1. *Homomorphisms and embeddings*

We can compare structures by looking at the functions between them. In this section we investigate how such maps are related to the theories of the structures in question.

Definition 1.1. Let L be an algebraic logic and $f : \mathfrak{A} \to \mathfrak{B}$ a partial function between Σ-structures.

(a) We say that f *preserves* a formula $\varphi(\bar{x}) \in L[\Sigma, X]$ if

$$\mathfrak{A} \vDash \varphi(\bar{a}) \quad \text{implies} \quad \mathfrak{B} \vDash \varphi(f\bar{a}), \qquad \text{for all } \bar{a} \subseteq \operatorname{dom} f.$$

(b) Let $\Delta \subseteq L[\Sigma, X]$ be a set of formulae. We call f a *Δ-map* if it preserves every formula in Δ. A *Δ-embedding* is a Δ-map that is an embedding. We say that f is *strict* if we have

$$\mathfrak{A} \vDash \varphi(\bar{a}) \quad \text{iff} \quad \mathfrak{B} \vDash \varphi(f\bar{a}),$$

for all formulae $\varphi(\bar{x}) \in \Delta$ and every $\bar{a} \subseteq \operatorname{dom} f$.

If $C \subseteq A \subseteq B$ then we say that $f : A \to B$ is a Δ-map or a Δ-embedding *over* C if f additionally satisfies $f \restriction C = \mathrm{id}_C$. For historical reasons FO-maps and FO-embeddings are usually called *elementary.*

(c) We denote by $\mathrm{Emb}_L(\mathfrak{A}, \mathfrak{B})$ the set of all $L^{<\omega}[\Sigma]$-embeddings $h : \mathfrak{A} \to \mathfrak{B}$. We write $\mathfrak{Emb}_L(\Sigma)$ for the category of all $L^{<\omega}[\Sigma]$-embeddings between Σ-structures.

Remark. If Δ is closed under negation then every Δ-map is strict.

Example. Let $f : \mathfrak{A} \to \mathfrak{B}$. Immediately from the definition it follows that

(a) f is injective if and only if it preserves the formula $x \neq y$;

(b) f is a homomorphism if and only if it preserves every atomic formula;

(c) f is an embedding if and only if it preserves every literal.

Definition 1.2. (a) We write $\mathrm{QF}_{\kappa\aleph_0}[\Sigma, X]$ for the set of all quantifier-free $\mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$-formulae.

(b) For $\Delta \subseteq \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$ we denote by $\exists\Delta$ the closure of Δ under existential quantifiers and conjunctions and disjunctions of less than κ formulae. Similarly, $\forall\Delta$ denotes the closure of Δ under conjunctions, disjunctions, and universal quantifiers. The intended value of κ should always be clear from the context.

(c) The set of *existential formulae* is $\exists_{\kappa\aleph_0}[\Sigma, X] := \exists\mathrm{QF}_{\kappa\aleph_0}[\Sigma, X]$ and the set of *universal formulae* is $\forall_{\kappa\aleph_0}[\Sigma, X] := \forall\mathrm{QF}_{\kappa\aleph_0}[\Sigma, X]$. For $\kappa = \aleph_0$, we simply write $\exists[\Sigma, X]$ and $\forall[\Sigma, X]$.

(d) The set $\exists^+_{\kappa\aleph_0}[\Sigma, X]$ of *positive existential formulae* consists of all $\mathrm{FO}_{\kappa\aleph_0}$-formulae containing neither negations nor universal quantifiers.

Lemma 1.3. *Let* $f : \mathfrak{A} \to \mathfrak{B}$.

(a) *f is a homomorphism if, and only if, it preserves all $\exists^+_{\infty\aleph_0}$-formulae.*

(b) *f is an embedding if, and only if, it preserves all $\exists_{\infty\aleph_0}$-formulae.*

(c) *If f is an isomorphism, it preserves all $\mathrm{FO}_{\infty\aleph_0}$-formulae.*

Proof. One direction follows immediately from the definition (see the above example) since every function preserving all atomic formulae is a homomorphism and every function preserving all literals is an embedding.

For the other direction, we prove all three claims simultaneously by induction on the structure of φ. For claims (b) and (c), we may assume that φ is in negation normal form.

If $\varphi = Rt_0 \dots t_n$ then we have

$$\begin{aligned}
\mathfrak{A} \vDash (Rt_0 \dots t_{n-1})(\bar{a}) &\Rightarrow \left\langle t_0^{\mathfrak{A}}(\bar{a}), \dots, t_{n-1}^{\mathfrak{A}}(\bar{a}) \right\rangle \in R^{\mathfrak{A}} \\
&\Rightarrow \left\langle f(t_0^{\mathfrak{A}}(\bar{a})), \dots, f(t_{n-1}^{\mathfrak{A}}(\bar{a})) \right\rangle \in R^{\mathfrak{B}} \\
&\Rightarrow \left\langle t_0^{\mathfrak{B}}(f\bar{a}), \dots, t_{n-1}^{\mathfrak{B}}(f\bar{a}) \right\rangle \in R^{\mathfrak{B}} \\
&\Rightarrow \mathfrak{B} \vDash (Rt_0, \dots, t_{n-1})(f\bar{a}) .
\end{aligned}$$

The proof for $\varphi = t_0 = t_1$ is similar.

For (b) and (c), we also have to consider the case that $\varphi = \neg Rt_0 \dots t_n$. Since in these cases f is a strict homomorphism we have

$$\begin{aligned}
\mathfrak{A} \vDash \neg(Rt_0, \dots, t_{n-1})(\bar{a}) &\Rightarrow \left\langle t_0^{\mathfrak{A}}(\bar{a}), \dots, t_{n-1}^{\mathfrak{A}}(\bar{a}) \right\rangle \notin R^{\mathfrak{A}} \\
&\Rightarrow \left\langle f(t_0^{\mathfrak{A}}(\bar{a})), \dots, f(t_{n-1}^{\mathfrak{A}}(\bar{a})) \right\rangle \notin R^{\mathfrak{B}} \\
&\Rightarrow \left\langle t_0^{\mathfrak{B}}(f\bar{a}), \dots, t_{n-1}^{\mathfrak{B}}(f\bar{a}) \right\rangle \notin R^{\mathfrak{B}} \\
&\Rightarrow \mathfrak{B} \vDash \neg(Rt_0, \dots, t_{n-1})(f\bar{a}) .
\end{aligned}$$

The proof for $\varphi = t_0 \neq t_1$ is similar.

The cases that $\varphi = \bigwedge \Phi$ or $\varphi = \bigvee \Phi$ follow immediately from the inductive hypothesis. Therefore, it remains to consider quantifiers. Suppose that $\varphi = \exists y \psi(\bar{x}, y)$. We have

$$\begin{aligned}
\mathfrak{A} \vDash \exists y \psi(\bar{a}, y) &\Rightarrow \mathfrak{A} \vDash \psi(\bar{a}, b) \text{ for some } b \in A \\
&\Rightarrow \mathfrak{B} \vDash \psi(f\bar{a}, fb) \text{ for some } b \in A \\
&\Rightarrow \mathfrak{B} \vDash \exists y \psi(f\bar{a}, y) .
\end{aligned}$$

Finally, for claim (c) there is the case that $\varphi = \forall y \psi(\bar{x}, y)$. Then we have

$$\begin{aligned}
\mathfrak{A} \vDash \forall y \psi(\bar{a}, y) &\Rightarrow \mathfrak{A} \vDash \psi(\bar{a}, b) \text{ for all } b \in A \\
&\Rightarrow \mathfrak{B} \vDash \psi(f\bar{a}, fb) \text{ for all } b \in A \\
&\Rightarrow \mathfrak{B} \vDash \forall y \psi(f\bar{a}, y) ,
\end{aligned}$$

since f is surjective. □

Corollary 1.4. *Let $\mathfrak{A}$ be a Σ-structure. For every relation $R := \varphi^{\mathfrak{A}}$ defined by some formula $\varphi(\bar{x}) \in \mathrm{FO}^{<\omega}_{\infty\aleph_0}[\Sigma]$, we have*

$$\bar{a} \in R \quad \text{iff} \quad \pi\bar{a} \in R\,, \quad \textit{for each automorphism } \pi : \mathfrak{A} \to \mathfrak{A}\,.$$

Example. We can use the above characterisation to prove that certain relations are not definable. Let $\mathfrak{A}$ be a structure and R a relation. If we can find an automorphism of $\mathfrak{A}$ that is not an automorphism of the expansion $\langle \mathfrak{A}, R\rangle$ then we know that R is not definable in $\mathfrak{A}$.

(a) Addition is not definable in the structure $\langle \mathbb{N}, \cdot\rangle$. Define the function $\pi : \mathbb{N} \to \mathbb{N}$ that maps a number of the form $2^m 3^n k$, where k is not divisible by 2 or 3, to the number $2^n 3^m k$. Then π is an automorphism of $\langle \mathbb{N}, \cdot\rangle$, but it is not an automorphism of $\langle \mathbb{N}, \cdot, +\rangle$ since we have

$$4 + 3 = 7 \quad \text{and} \quad \pi(4) + \pi(3) = 9 + 2 \neq 7 = \pi(7)\,.$$

(b) Similarly, we can show that multiplication is not definable in the structure $\langle \mathbb{Z}, +\rangle$ since the mapping $\pi : x \mapsto -x$ is an automorphism of $\langle \mathbb{Z}, +\rangle$ but not of $\langle \mathbb{Z}, +, \cdot\rangle$.

Definition 1.5. A formula $\varphi(\bar{x})$ is *preserved in substructures* if

$$\mathfrak{A} \vDash \varphi(\bar{a}) \quad \text{implies} \quad \mathfrak{A}_0 \vDash \varphi(\bar{a})\,,$$

whenever $\mathfrak{A}_0 \subseteq \mathfrak{A}$ is a substructure containing $\bar{a}$.

Lemma 1.6. *$\forall_{\infty\aleph_0}$-formulae are preserved in substructures.*

Proof. This is just the dual statement of Lemma 1.3 (b). Let $\varphi \in \forall_{\infty\aleph_0}$ and suppose there exist structures $\mathfrak{A}_0 \subseteq \mathfrak{A}$ and elements $\bar{a} \subseteq A_0$ such that

$$\mathfrak{A} \vDash \varphi(\bar{a}) \quad \text{but} \quad \mathfrak{A}_0 \nvDash \varphi(\bar{a})\,.$$

Let $\mathrm{id} : \mathfrak{A}_0 \to \mathfrak{A}$ be the embedding of $\mathfrak{A}_0$ into $\mathfrak{A}$. Since $\neg\varphi$ is equivalent to some existential formula $\psi \in \exists_{\infty\aleph_0}$ it follows from Lemma 1.3 (b) that

$$\mathfrak{A}_0 \vDash \neg\varphi(\bar{a}) \quad \text{implies} \quad \mathfrak{A} \vDash \neg\varphi(\bar{a})\,.$$

Contradiction. □

Example. Groups can be axiomatised by universal sentences:

$$\forall x \forall y \forall z (x \cdot (y \cdot z) = (x \cdot y) \cdot z)$$
$$\forall x (x \cdot e = x)$$
$$\forall x (x \cdot x^{-1} = e)$$

It follows that every substructure of a group $\langle G, \cdot, ^{-1}, e \rangle$ is itself a group.

Note that, if we use the smaller signature consisting only of group multiplication $\cdot$, this property fails since the axioms are no longer universal:

$$\forall x \forall y \forall z (x \cdot (y \cdot z) = (x \cdot y) \cdot z)$$
$$\exists e \forall x [x \cdot e = x \wedge \exists y (x \cdot y = e)]$$

For instance, the group $\langle \mathbb{Z}, + \rangle$ has the substructure $\langle \mathbb{N}, + \rangle$ which is not a group.

Definition 1.7. A formula $\varphi(\bar{x})$ is *preserved in unions of chains* if, for all chains $(\mathfrak{A}_i)_{i<\alpha}$ and every tuple $\bar{a} \subseteq A_0$,

$$\mathfrak{A}_i \vDash \varphi(\bar{a}) \,, \text{ for all } i < \alpha \,, \quad \text{implies} \quad \bigcup_{i<\alpha} \mathfrak{A}_i \vDash \varphi(\bar{a}) \,.$$

Lemma 1.8. *Every* $\forall\exists_{\infty\aleph_0}$*-formula* φ *is preserved in unions of chains.*

Proof. Let $(\mathfrak{A}_i)_{i<\alpha}$ be a chain with union $\mathfrak{B} := \bigcup_{i<\alpha} \mathfrak{A}_i$. Suppose that $\varphi \in \forall\exists_{\infty\aleph_0}$ is a formula such that $\mathfrak{A}_i \vDash \varphi(\bar{a})$, for all $i < \alpha$, where $\bar{a} \subseteq A_0$. We prove by induction on φ that $\mathfrak{B} \vDash \varphi(\bar{a})$.

If $\varphi \in \exists_{\infty\aleph_0}$ then $\mathfrak{A}_0 \vDash \varphi(\bar{a})$ and $\mathfrak{A}_0 \subseteq \mathfrak{B}$ implies that $\mathfrak{B} \vDash \varphi(\bar{a})$, by Lemma1.3 (b). If $\varphi = \bigwedge \Phi$ or $\varphi = \bigvee \Phi$, for $\Phi \subseteq \forall\exists_{\infty\aleph_0}$ then the claim follows immediately from the inductive hypothesis.

Hence, it remains to consider the case that $\varphi = \forall y \psi(\bar{x}, y)$, for some $\psi \in \forall\exists_{\infty\aleph_0}$. For every $b \in B$, there is some index k such that $b \in A_k$. By assumption, we have $\mathfrak{A}_i \vDash \psi(\bar{a}, b)$, for every $i \geq k$. By inductive hypothesis, it follows that $\bigcup_{i \geq k} \mathfrak{A}_i \vDash \psi(\bar{a}, b)$. Since $\bigcup_{i \geq k} \mathfrak{A}_i = \mathfrak{B}$ we have shown that $\mathfrak{B} \vDash \psi(\bar{a}, b)$, for all $b \in B$. This implies that $\mathfrak{B} \vDash \forall y \psi(\bar{a}, y)$. □

Remark. Similarly to Lemma **??**, we can show that $\forall\exists_{\infty\aleph_0}$-formulae are preserved in direct limits of diagrams of embeddings. Analogously it follows that $\forall\exists^+_{\infty\aleph_0}$-formulae are preserved in arbitrary direct limits.

Example. The class of all fields is $\forall\exists$-axiomatisable. It follows that the union of a chain of fields is again a field.

Exercise 1.1. Prove that every $\forall\exists^+_{\infty\aleph_0}$-formula is preserved in direct limits.

2. Elementary embeddings

Definition 2.1. Let L be an algebraic logic, $\Delta \subseteq L[\Sigma, X]$ a set of formulae, and $\mathfrak{A}$ and $\mathfrak{B}$ Σ-structures.

We say that $\mathfrak{B}$ is a *Δ-extension* of $\mathfrak{A}$, or that $\mathfrak{A}$ is a *Δ-substructure* of $\mathfrak{B}$, if $\mathfrak{A} \subseteq \mathfrak{B}$ and the inclusion map $A \to B$ is a Δ-embedding. We write $\mathfrak{A} \preceq_\Delta \mathfrak{B}$ to indicate that $\mathfrak{A}$ is a Δ-substructure of $\mathfrak{B}$. In the case $\Delta = \mathrm{FO}[\Sigma]$ we also speak of *elementary* embeddings and extensions, and we write $\mathfrak{A} \preceq \mathfrak{B}$ instead of $\mathfrak{A} \preceq_{\mathrm{FO}} \mathfrak{B}$.

Example. (a) $\langle \mathbb{N}, \leq \rangle \subseteq \langle \mathbb{Q}, \leq \rangle$ is not elementary since

$$\langle \mathbb{N}, \leq \rangle \vDash \exists x \forall y (x \leq y) \quad \text{but} \quad \langle \mathbb{Q}, \leq \rangle \nvDash \exists x \forall y (x \leq y)\,.$$

(b) There are structures $\mathfrak{A} \subseteq \mathfrak{B}$ such that $\mathfrak{A} \equiv \mathfrak{B}$ but $\mathfrak{A} \npreceq \mathfrak{B}$. For instance, let $\mathfrak{A} := \langle 2\mathbb{Z}, \leq \rangle$ and $\mathfrak{B} := \langle \mathbb{Z}, \leq \rangle$. Then we even have $\mathfrak{A} \cong \mathfrak{B}$ but $\mathfrak{A} \npreceq \mathfrak{B}$ since

$$\langle 2\mathbb{Z}, \leq \rangle \nvDash \exists x (0 < x \wedge x < 2) \quad \text{but} \quad \langle \mathbb{Z}, \leq \rangle \vDash \exists x (0 < x \wedge x < 2)\,.$$

(c) $\langle \mathbb{Q}, \leq \rangle \preceq_{\mathrm{FO}} \langle \mathbb{R}, \leq \rangle$. (The easiest proof of this statement is based on so-called 'back-and-forth' arguments which will be introduced in Chapter c4. See Lemma c4.1.4).

Exercise 2.1. Find an elementary extension of $\langle \mathbb{Z}, s \rangle$ where $s : x \mapsto x + 1$ is the successor function.

Remark. If L is closed under negation then $\mathfrak{A} \leq_L \mathfrak{B}$ implies $\mathfrak{A} \equiv_L \mathfrak{B}$.

Definition 2.2. Let L be an algebraic logic and $\mathfrak{A}$ a Σ-structure.

(a) For a set $U \subseteq A$, we denote by $\mathfrak{A}_U$ the expansion of $\mathfrak{A}$ by one constant c_a, for each element $a \in U$, with value $c_a^{\mathfrak{A}} := a$. By Σ_U we denote the corresponding expansion of the signature. In the following we will not distinguish between the element a and the symbol c_a denoting it, and we simply write a in both cases.

(b) If T is a complete theory and $\mathfrak{A}$ a model of T with $U \subseteq A$ then we define $T(U) := \mathrm{Th}_L(\mathfrak{A}_U)$. For $U = A$, we call $T(A)$ the *L-diagram* of $\mathfrak{A}$.

Let $\Delta_0 \subseteq \mathrm{FO}[\Sigma]$ be the set of all atomic first-order formulae and $\Delta_1 \subseteq \mathrm{FO}[\Sigma]$ the set of all literals. The Δ_0-diagram of $\mathfrak{A}$ is called the *atomic diagram,* and the Δ_1-diagram is the *algebraic diagram.* As usual, the FO-diagram is called *elementary.*

The next lemma states that in order to construct an L-extension of a structure $\mathfrak{A}$ we can take any model of its L-diagram.

Lemma 2.3 (Diagram Lemma). *Let L be an algebraic logic and $\mathfrak{A}$ and $\mathfrak{B}$ Σ-structures. There exists an L-map $g : \mathfrak{A} \to \mathfrak{B}$ if and only if we have*

$$\mathfrak{B}^+ \vDash \mathrm{Th}_L(\mathfrak{A}_A)\,, \quad \textit{for some } \Sigma_A\textit{-expansion } \mathfrak{B}^+ \textit{ of } \mathfrak{B}\,.$$

Proof. ($\Rightarrow$) By definition, $\mathfrak{B} \vDash \varphi(g\bar{a})$, for all $\varphi(\bar{a}) \in \mathrm{Th}(\mathfrak{A}_A)$. Hence, if $\bar{a}$ is an enumeration of A then we can define the desired expansion of $\mathfrak{B}$ by $\mathfrak{B}^+ := \langle \mathfrak{B}, g(\bar{a})\rangle$.

($\Leftarrow$) We claim that the function $g : A \to B : a \mapsto c_a^{\mathfrak{B}^+}$ is the desired L-embedding. Since $\mathrm{Th}_L(\mathfrak{B}^+) = \mathrm{Th}_L(\mathfrak{A}_A)$ we have

$$\begin{aligned}
\mathfrak{A} \vDash \varphi(a_0, \dots, a_{n-1}) \quad &\text{iff} \quad \varphi(c_{a_0}, \dots, c_{a_{n-1}}) \in \mathrm{Th}_L(\mathfrak{A}_A) \\
&\Rightarrow \quad \mathfrak{B}^+ \vDash \varphi(c_{a_0}, \dots, c_{a_{n-1}}) \\
&\text{iff} \quad \mathfrak{B} \vDash \varphi(g(a_0), \dots, g(a_{n-1}))\,. \qquad \square
\end{aligned}$$

Corollary 2.4. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be structures. Let $\Delta_0(\mathfrak{A})$ be the atomic diagram of $\mathfrak{A}$ and $\Delta_1(\mathfrak{A})$ the algebraic diagram.*

(a) *There exists a homomorphism* $\mathfrak{A} \to \mathfrak{B}$ *if and only if*

$$\mathfrak{B}_A \vDash \Delta_0(\mathfrak{A}), \quad \textit{for some expansion } \mathfrak{B}_A \textit{ of } \mathfrak{B}.$$

(b) *There exists an embedding* $\mathfrak{A} \to \mathfrak{B}$ *if and only if*

$$\mathfrak{B}_A \vDash \Delta_1(\mathfrak{A}), \quad \textit{for some expansion } \mathfrak{B}_A \textit{ of } \mathfrak{B}.$$

For first-order logic there is a simple test to check whether some extension is elementary.

Theorem 2.5 (Tarski-Vaught Test). *Let* $\mathfrak{A} \subseteq \mathfrak{B}$ *be* Σ*-structures and suppose that* $\Delta \subseteq \mathrm{FO}_{\infty\aleph_0}[\Sigma]$ *is closed under negation, subformulae, and negation normal forms.*

We have $\mathfrak{A} \preceq_\Delta \mathfrak{B}$ *if and only if, for every formula* $\exists y\varphi(\bar{x}, y) \in \Delta$ *and all tuples* $\bar{a} \subseteq A$,

$$\mathfrak{B} \vDash \exists y\varphi(\bar{a}, y) \quad \textit{implies} \quad \mathfrak{B} \vDash \varphi(\bar{a}, b), \textit{ for some } b \in A.$$

Proof. ($\Rightarrow$) Since $\mathfrak{A} \preceq_\Delta \mathfrak{B}$ and Δ is closed under negation we have

$$\begin{aligned} \mathfrak{B} \vDash \exists y\varphi(\bar{a}, y) \quad &\text{iff} \quad \mathfrak{A} \vDash \exists y\varphi(\bar{a}, y) \\ &\text{iff} \quad \mathfrak{A} \vDash \varphi(\bar{a}, b) \qquad \text{for some } b \in A \\ &\text{iff} \quad \mathfrak{B} \vDash \varphi(\bar{a}, b) \qquad \text{for some } b \in A. \end{aligned}$$

($\Leftarrow$) Since Δ is closed under subformulae we can prove by induction on φ that

$$\mathfrak{A} \vDash \varphi(\bar{a}) \quad \text{implies} \quad \mathfrak{B} \vDash \varphi(\bar{a}), \quad \text{for all } \varphi \in \Delta.$$

Moreover, it is sufficient to consider only formulae φ in negation normal form.

We will only give the inductive step for the universal quantifier. The other cases are handled in the same way as in the proof of Lemma 1.3. Suppose that

$$\mathfrak{A} \vDash \forall y\psi(\bar{a}, y) \quad \text{but} \quad \mathfrak{B} \nvDash \forall y\psi(\bar{a}, y).$$

Since $\forall y\psi \in \Delta$ we have $\textsc{nnf}(\neg\forall y\psi) = \exists y(\textsc{nnf}(\neg\psi)) \in \Delta$. Therefore, $\mathfrak{B} \vDash \exists y\neg\psi(\bar{a}, y)$ implies that $\mathfrak{B} \vDash \neg\psi(\bar{a}, b)$, for some $b \in A$. On the other hand, $\mathfrak{A} \vDash \forall y\psi(\bar{a}, y)$ implies that $\mathfrak{A} \vDash \psi(\bar{a}, b)$ and, by inductive hypothesis, it follows that $\mathfrak{B} \vDash \psi(\bar{a}, b)$. Contradiction. □

Proposition 2.6. *Let $\mathcal{D} : \mathfrak{I} \to \mathrm{Hom}_s(\Sigma)$ be a directed diagram of strict homomorphisms with cone $h_i : \mathcal{D}(i) \to \varinjlim \mathcal{D}$, $i \in I$, and suppose that $\Delta \subseteq \mathrm{FO}_{\infty\aleph_0}[\Sigma, X]$ is closed under subformulae and negation. If each map $\mathcal{D}(i, j)$ is a Δ-map then so is every h_i.*

Proof. By induction on $\varphi \in \Delta$ we prove that

$$\mathcal{D}(i) \vDash \varphi(\bar{a}) \quad \text{iff} \quad \varinjlim \mathcal{D} \vDash \varphi(h_i(\bar{a}))\,.$$

Since $\forall y\psi(\bar{x}, y) \equiv \neg\exists y\neg\psi(\bar{x}, y)$ and Δ is closed under negation we may w.l.o.g. assume that φ does not contain universal quantifiers.

If φ is atomic then the claim follows from the fact that h_i is a strict homomorphism. The cases that $\varphi = \neg\psi$, $\varphi = \bigwedge \Phi$, or $\varphi = \bigvee \Phi$ follow immediately from inductive hypothesis.

Suppose that $\varphi = \exists y\psi(\bar{x}, y)$. If $\mathcal{D}(i) \vDash \exists y\psi(\bar{a}, y)$ then there is some $b \in \mathcal{D}(i)$ such that $\mathcal{D}(i) \vDash \psi(\bar{a}, b)$. By inductive hypothesis, it follows that $\varinjlim \mathcal{D} \vDash \psi(h_i(\bar{a}b))$. Hence, $\varinjlim \mathcal{D} \vDash \varphi(h_i(\bar{a}))$. Conversely, suppose that $\varinjlim \mathcal{D} \vDash \exists y\psi(h_i(\bar{a}), y)$. Then there is some element b such that $\varinjlim \mathcal{D} \vDash \psi(h_i(\bar{a}), b)$. By definition of a direct limit there is some index k with $b \in \operatorname{rng} h_k$. Let $l \in I$ be an index with $i, k \leq l$ and let $c \in \mathcal{D}(l)$ be an element with $h_l(c) = b$. By inductive hypothesis, it follows that $\mathcal{D}(l) \vDash \psi(\mathcal{D}(i, l)(\bar{a}), c)$. Hence, $\mathcal{D}(l) \vDash \varphi(\mathcal{D}(i, l)(\bar{a}))$. Since $\mathcal{D}(i, l)$ is a Δ-map and Δ is closed under negation we have $\mathcal{D}(i) \vDash \varphi(\bar{a})$, as desired. □

Definition 2.7. A chain $(\mathfrak{A}_i)_{i<\alpha}$ is an *L-chain* if $\mathfrak{A}_i \preceq_L \mathfrak{A}_k$, for all $i < k$. As usual, FO-chains are also called *elementary.*

Corollary 2.8. *If $(\mathfrak{A}_i)_{i<\alpha}$ is an $\mathrm{FO}_{\kappa\aleph_0}$-chain then $\mathfrak{A}_k \preceq_{\mathrm{FO}_{\kappa\aleph_0}} \bigcup_{i<\alpha} \mathfrak{A}_i$, for all $k < \alpha$.*

If $\Delta \subseteq \mathrm{FO}_{\kappa\aleph_0}$ is not closed under negation then obtain a similar result if we require the diagram to be κ-directed and Δ to not contain universal quantifiers.

Proposition 2.9. *Let $\mathcal{D} : \mathfrak{I} \to \mathrm{Hom}(\Sigma)$ be a κ-directed diagram with cone $h_i : \mathcal{D}(i) \to \varinjlim \mathcal{D}$, $i \in I$, and suppose that $\Delta \subseteq \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$ is closed under subformulae and no formula in Δ contains universal quantifiers. If each map $\mathcal{D}(i,j)$ is a Δ-map then so is every h_i.*

Proof. By induction on $\varphi \in \Delta$ we prove that

$$\varinjlim \mathcal{D} \vDash \varphi(\bar{a}) \quad \text{iff} \quad \text{there is some } i \in I \text{ and a tuple } \bar{b} \text{ with } h_i(\bar{b}) = \bar{a} \text{ such that } \mathcal{D}(i) \vDash \varphi(\bar{b})\,.$$

(φ atomic) follows from the definition of $\varinjlim \mathcal{D}$.

($\varphi = \bigvee \Psi$) If $\mathcal{D}(i) \vDash \varphi(\bar{b})$ then there is a formula $\psi \in \Psi$ with $\mathcal{D}(i) \vDash \psi(\bar{b})$. By inductive hypothesis it follows that $\varinjlim \mathcal{D} \vDash \varphi(\bar{a})$. Conversely, if $\varinjlim \mathcal{D} \vDash \psi(\bar{a})$, for some $\psi \in \Psi$, then we have $\mathcal{D}(i) \vDash \psi(\bar{b})$ and $h_i(\bar{b}) = \bar{a}$, for suitable i and $\bar{b}$.

($\varphi = \bigwedge \Psi$) If $\mathcal{D}(i) \vDash \varphi(\bar{b})$ then the inductive hypothesis implies that $\varinjlim \mathcal{D} \vDash \psi(\bar{a})$, for each $\psi \in \Psi$. Conversely, if $\varinjlim \mathcal{D} \vDash \varphi(\bar{a})$ then we can find, for every $\psi \in \Psi$, an index $i_\psi \in I$ and a tuple $\bar{b}_\psi$ with $h_{i_\psi}(\bar{b}_\psi) = \bar{a}$ and $\mathcal{D}(i_\psi) \vDash \psi(\bar{b}_\psi)$. Since $h_{i_\psi}(\bar{b}_\psi) = h_{i_\vartheta}(\bar{b}_\vartheta)$, for $\psi, \theta \in \Psi$, there exists, by definition of $\varinjlim \mathcal{D}$, an index $l_{\psi\vartheta} \geq i_\psi, i_\theta$ with

$$\mathcal{D}(i_\psi, l_{\psi\vartheta})(\bar{b}_\psi) = \mathcal{D}(i_\vartheta, l_{\psi\vartheta})(\bar{b}_\vartheta)\,.$$

Since $\mathfrak{I}$ is κ-directed we can find index $k \in I$ with $l_{\psi\vartheta} \leq k$, for all ψ, ϑ. Let $\bar{c} := \mathcal{D}(i_\psi, k)(\bar{b}_\psi)$, for some/all ψ. It follows that $h_k(\bar{c}) = \bar{a}$ and $\mathcal{D}(k) \vDash \psi(\bar{c})$, for every $\psi \in \Psi$.

($\varphi = \neg\psi$) Since all homomorphisms $\mathcal{D}(i,k)$ are Δ-maps and $\neg\psi \in \Delta$ we have

$$\mathcal{D}(i) \vDash \psi(\bar{b}) \quad \text{iff} \quad \mathcal{D}(k) \vDash \psi(\mathcal{D}(i,k)(\bar{b}))\,, \quad \text{for all } i \leq k\,.$$

Consequently, $h_i(\bar{b}) = h_j(\bar{c})$, for arbitrary $i, j \in I$, implies

$$\mathcal{D}(i) \vDash \psi(\bar{b}) \quad \text{iff} \quad \mathcal{D}(j) \vDash \psi(\bar{c}) .$$

Therefore, we have

$$\begin{aligned} & \varinjlim \mathcal{D} \nvDash \psi(\bar{a}) \\ \text{iff} \quad & \mathcal{D}(i) \nvDash \psi(\bar{b}) && \text{for all } i \text{ and } \bar{b} \in h_i^{-1}(\bar{a}) , \\ \text{iff} \quad & \mathcal{D}(i) \vDash \neg\psi(\bar{b}) && \text{for all } i \text{ and } \bar{b} \in h_i^{-1}(\bar{a}) , \\ \text{iff} \quad & \mathcal{D}(i) \vDash \neg\psi(\bar{b}) && \text{for some } i \text{ and } \bar{b} \in h_i^{-1}(\bar{a}) . \end{aligned}$$

($\varphi = \exists y \psi(\bar{x}, y)$) If $\mathcal{D}(i) \vDash \exists y \psi(\bar{b}, y)$ then there is some $c \in \mathcal{D}(i)$ such that $\mathcal{D}(i) \vDash \psi(\bar{b}, c)$. By inductive hypothesis, it follows that

$$\varinjlim \mathcal{D} \vDash \psi(h_i(\bar{b}c)) .$$

Hence, $\varinjlim \mathcal{D} \vDash \varphi(h_i(\bar{b}))$. Conversely, suppose that $\varinjlim \mathcal{D} \vDash \exists y \psi(\bar{a}, y)$. Then there is some element c such that $\varinjlim \mathcal{D} \vDash \psi(\bar{a}, c)$. By inductive hypothesis, we can find an index i and elements $\bar{b}d \in h_i^{-1}(\bar{a}c)$ such that $\mathcal{D}(i) \vDash \psi(\bar{b}, d)$. Hence, $\mathcal{D}(i) \vDash \varphi(\bar{b})$. ☐

Exercise 2.2. Find an example showing that the above Proposition does not hold if Δ contains a formula with a universal quantifier.

We conclude this section with the observation that interpretations preserve elementary embeddings.

Lemma 2.10. *Let Σ and Γ be signatures. Every first-order interpretation $\mathcal{I}$ from Σ to Γ induces a functor $\mathcal{I} : \mathfrak{Emb}_{\mathcal{I}} \to \mathfrak{Emb}_{\mathrm{FO}}(\Gamma)$, where $\mathfrak{Emb}_{\mathcal{I}}$ denotes the subcategory of $\mathfrak{Emb}_{\mathrm{FO}}(\Sigma)$ consisting of all structures $\mathfrak{A}$ such that $\mathcal{I}(\mathfrak{A})$ is defined.*

Proof. Suppose that $\mathcal{I} = \langle \alpha, (\delta_s)_s, (\varepsilon_s)_s, (\varphi_\xi)_\xi \rangle$, let $h : \mathfrak{A} \to \mathfrak{B}$ be an elementary embedding such that $\mathcal{I}(\mathfrak{A})$ and $\mathcal{I}(\mathfrak{B})$ are defined, and let

$$\iota_s : \delta_s^{\mathfrak{A}} \to \mathcal{I}(\mathfrak{A}) \quad \text{and} \quad \kappa_s : \delta_s^{\mathfrak{B}} \to \mathcal{I}(\mathfrak{B})$$

be the canonical functions mapping a tuple to the element it encodes. We define $\mathcal{I}(h) : \mathcal{I}(\mathfrak{A}) \to \mathcal{I}(\mathfrak{B})$ as follows. For every element c of $\mathcal{I}(\mathfrak{A})$ of sort s, we set

$$\mathcal{I}(h)(c) := \kappa_s(h(\bar{a}))\,, \quad \text{for any } \bar{a} \in \iota_s^{-1}(c)\,.$$

We claim that $\mathcal{I}(h)$ is a well-defined elementary embedding.

To show that it is well-defined, suppose that $\bar{a}, \bar{a}' \in \iota_s^{-1}(c)$. Then

$$\mathfrak{A} \vDash \varepsilon_s(\bar{a}, \bar{a}') \quad \text{implies} \quad \mathfrak{B} \vDash \varepsilon_s(h(\bar{a}), h(\bar{a}'))\,.$$

Consequently,

$$\kappa_s(h(\bar{a})) = \kappa_s(h(\bar{a}'))\,,$$

as desired.

Hence, it remains to show that $\mathcal{I}(h)$ is an elementary embedding. Let $\bar{c}$ be an n-tuple in $\mathcal{I}(\mathfrak{A})$ with sorts $\bar{s}$ and let $\varphi(\bar{x})$ be a first-order formula. Choosing tuples $\bar{a}_i \in \iota_{s_i}^{-1}(c_i)$, it follows by Lemma c1.5.9 that

$$\begin{aligned}
& \mathcal{I}(\mathfrak{A}) \vDash \varphi(\bar{c}) \\
\text{iff}\quad & \mathfrak{A} \vDash \varphi^{\mathcal{I}}(\bar{a}_0, \dots, \bar{a}_{n-1}) \\
\text{iff}\quad & \mathfrak{B} \vDash \varphi^{\mathcal{I}}(h(\bar{a}_0), \dots, h(\bar{a}_{n-1})) \\
\text{iff}\quad & \mathcal{I}(\mathfrak{B}) \vDash \varphi(\kappa_{s_0}(h(\bar{a}_0)), \dots, \kappa_{s_{n-1}}(h(\bar{a}_{n-1}))) \\
\text{iff}\quad & \mathcal{I}(\mathfrak{B}) \vDash \varphi(\mathcal{I}(h)(\bar{c}))\,.
\end{aligned}$$

□

3. *The Theorem of Löwenheim and Skolem*

A general method to eliminate existential quantifiers consists in replacing them by functions. Consider a formula $\psi = \exists y \varphi(\bar{x}, y)$ which states that, for a given value of $\bar{x}$, there exists some element y satisfying φ. If we define a function f that maps all suitable values of $\bar{x}$ to such an element y then we can write ψ equivalently as $\varphi(\bar{x}, f\bar{x})$. Informally we say that the function f we constructed yields a 'witness' that asserts the truth of $\exists y \varphi$.

Definition 3.1. Let $\Phi \subseteq \mathrm{FO}^{0}_{\infty\aleph_0}[\Sigma]$ and $\Delta \subseteq \mathrm{FO}^{<\omega}_{\infty\aleph_0}[\Sigma]$.

(a) A Σ-term $t(\bar{x})$ defines a *Skolem function* for a formula $\exists y\varphi(\bar{x}, y)$ (w.r.t. Φ) if

$$\Phi \vDash \forall\bar{x}\big[\exists y\varphi(\bar{x}, y) \rightarrow \varphi(\bar{x}, t(\bar{x}))\big].$$

A formula of the form $\forall\bar{x}[\exists y\varphi(\bar{x}, y) \rightarrow \varphi(\bar{x}, t)]$ is called a *Skolem axiom* for $\exists y\varphi$.

(b) A Δ*-Skolemisation* of Φ is a set $\Phi^+ \subseteq \mathrm{FO}^{0}_{\infty\aleph_0}[\Sigma^+]$, for some signature $\Sigma^+ \supseteq \Sigma$, such that

- $\Phi \subseteq \Phi^+$,
- every model $\mathfrak{M} \vDash \Phi$ has an Σ^+-expansion $\mathfrak{M}^+ \vDash \Phi^+$ and,
- for every formula $\exists y\varphi \in \Delta$, there exists a Σ^+-term defining a Skolem function for $\exists y\varphi$.

(c) We say that a theory $T \subseteq \mathrm{FO}^{0}_{\infty\aleph_0}[\Sigma]$ is a Δ*-Skolem theory* if T is a Δ-Skolemisation of itself. If $\Delta = \mathrm{FO}^{<\omega}_{\kappa\aleph_0}[\Sigma]$ we simply speak of a *Skolemisation* and a *Skolem theory.* The intended value of κ and Σ should always be clear from the context.

Example. Consider the ordered additive group of the real numbers $\mathfrak{R} = \langle\mathbb{R}, +, <, f\rangle$ expanded by the (definable) function $f(x) := x/2$. The term $f(x_0 + x_1)$ defines a Skolem function for the formula

$$\varphi(x_0, x_1) := \exists y(x_0 < y < x_1).$$

The main reason why Skolem theories are interesting is the property of their models that *all* substructures are elementary.

Lemma 3.2. *Let* $T \subseteq \mathrm{FO}^{0}_{\kappa\aleph_0}[\Sigma]$ *be a* Δ*-Skolem theory where the set* $\Delta \subseteq \mathrm{FO}^{<\omega}_{\kappa\aleph_0}[\Sigma]$ *is closed under negation, subformulae, and negation normal forms. If* $\mathfrak{A} \vDash T$ *and* $\mathfrak{B} \subseteq \mathfrak{A}$ *then* $\mathfrak{B} \preceq_\Delta \mathfrak{A}$.

Proof. We apply the Tarski-Vaught Test. Suppose that $\exists y\varphi(\bar{x}, y) \in \Delta$ is a formula and $\bar{a} \subseteq B$ a tuple such that

$$\mathfrak{A} \vDash \exists y\varphi(\bar{a}, y).$$

Let t be a term defining a Skolem function for $\exists y\varphi$. Then

$$\mathfrak{A} \vDash \varphi(\bar{a}, t(\bar{a}))\,.$$

Since $\bar{a} \subseteq B$ and B is closed under all functions of $\mathfrak{A}$ it follows that $t^{\mathfrak{A}}(\bar{a}) \in B$, as desired. ☐

Syntactically we can use Skolemisation to eliminate existential quantifiers.

Lemma 3.3. *Suppose that $T \subseteq \mathrm{FO}^{0}_{\kappa\aleph_0}[\Sigma]$ is a Skolem theory. For every formula $\varphi \in \mathrm{FO}^{<\omega}_{\kappa\aleph_0}[\Sigma]$, we can construct a formula $\varphi^* \in \forall^{<\omega}_{\kappa\aleph_0}[\Sigma]$ such that*

$$\varphi^* \vDash \varphi \quad \textit{and} \quad T \vDash \varphi \to \varphi^*.$$

In particular, $\varphi \equiv \varphi^*$ modulo T.

Proof. We define φ^* by induction on φ. W.l.o.g. we may assume that φ is in negation normal form. For $\varphi \in \forall_{\kappa\aleph_0}$ we set $\varphi^* := \varphi$. For conjunctions, disjunctions, and universal quantifiers, we set

$$\begin{aligned}(\bigwedge \Psi)^* &:= \bigwedge \{\, \psi^* \mid \psi \in \Psi \,\}\,,\\ (\bigvee \Psi)^* &:= \bigvee \{\, \psi^* \mid \psi \in \Psi \,\}\,,\\ (\forall y\psi)^* &:= \forall y\psi^*.\end{aligned}$$

Finally, for $\varphi = \exists y\psi(\bar{x}, y)$ we set $\varphi^* := \psi^*(\bar{x}, t_\varphi)$ where the term t_φ defines a Skolem function for φ. ☐

Corollary 3.4. *For every Skolem theory $T \subseteq \mathrm{FO}^{0}_{\kappa\aleph_0}[\Sigma]$ there exists a set $\Phi \subseteq \forall_{\kappa\aleph_0}[\Sigma]$ such that $T \equiv \Phi$.*

Proof. Let $\Phi := \{\, \varphi \in \forall^{0}_{\kappa\aleph_0}[\Sigma] \mid T \vDash \varphi \,\}$. Then we have $T \vDash \Phi$. Conversely, we can use the preceding lemma to assign to every formula $\varphi \in T$ a formula $\varphi^* \in \Phi$ with $\varphi^* \vDash \varphi$. This implies that $\Phi \vDash T$. ☐

Constructing Δ-Skolemisations is easy. We just have to add Skolem axioms for all formulae in Δ.

Lemma 3.5. *For all* $\Phi \subseteq \mathrm{FO}^{0}_{\kappa\aleph_0}[\Sigma]$ *and* $\Delta \subseteq \mathrm{FO}^{<\omega}_{\kappa\aleph_0}[\Sigma]$, *there exists a* Δ*-Skolemisation* $\Phi^+ \subseteq \mathrm{FO}^{0}_{\kappa\aleph_0}[\Sigma^+]$ *of* Φ *with* $|\Phi^+| \leq |\Phi| \oplus |\Delta|$ *and* $|\Sigma^+| \leq |\Sigma| \oplus |\Delta|$.

Proof. Let Σ^+ be the signature obtained from Σ by adding new function symbols $f_{\exists y\varphi}$, for every formula $\exists y\varphi \in \Delta$. We construct Φ^+ by adding to Φ all Skolem axioms

$$\chi_{\exists y\varphi} := \forall \bar{x}\big[\exists y\varphi(\bar{x}, y) \rightarrow \varphi(\bar{x}, f_{\exists y\varphi}\bar{x})\big]$$

with $\exists y\varphi \in \Delta$. Clearly, $|\Phi^+| \leq |\Phi| \oplus |\Delta|$ and $|\Sigma^+| \leq |\Sigma| \oplus |\Delta|$.

We claim that Φ^+ is a Δ-Skolemisation of Φ. By construction, we have $\Phi \subseteq \Phi^+$ and every formula $\exists y\varphi \in \Delta$ has the Skolem function $f_{\exists y\varphi}$. Hence, it remains to prove that every model of Φ can be expanded to one of Φ^+.

Suppose that $\mathfrak{A} \models \Phi$. We construct an expansion $\mathfrak{A}^+ \models \Phi^+$ as follows. Let $\exists y\varphi \in \Delta$ and $\bar{a} \subseteq A$. If $\mathfrak{A} \models \exists y\varphi(\bar{a}, y)$ then we select some $b \in A$ such that $\mathfrak{A} \models \varphi(\bar{a}, b)$ and we set $f^{\mathfrak{A}^+}_{\exists y\varphi}(\bar{a}) := b$. Otherwise, we set $f^{\mathfrak{A}^+}_{\exists y\varphi}(\bar{a}) := b$, for an arbitrary element $b \in A$. This ensures that $\mathfrak{A}^+ \models \chi_{\exists y\varphi}$. Since $\mathfrak{A} \models \Phi$ and the function symbols $f_{\exists y\varphi}$ do not appear in Φ we further have $\mathfrak{A}^+ \models \Phi$. Consequently, $\mathfrak{A}^+ \models \Phi^+$. □

In order to obtain a Skolem theory we can iterate this construction.

Theorem 3.6. *Let* κ *be a regular cardinal. Every set* $\Phi \subseteq \mathrm{FO}^{0}_{\kappa\aleph_0}[\Sigma]$ *has a Skolemisation* $\Phi^+ \subseteq \mathrm{FO}^{0}_{\kappa\aleph_0}[\Sigma^+]$ *such that* $|\Phi^+| \leq (|\Sigma| \oplus \aleph_0)^{<\kappa}$ *and* $(\Phi^+)^{\models}$ *is a Skolem theory.*

Proof. We construct an increasing sequence of sets $(\Phi_\alpha)_{\alpha<\kappa}$ with $\Phi_\alpha \subseteq \mathrm{FO}^{0}_{\kappa\aleph_0}[\Sigma_\alpha]$. We set $\Phi_0 := \Phi$ and $\Phi_\delta := \bigcup_{\alpha<\delta} \Phi_\alpha$, for limit ordinals δ. For the successor step, we use Lemma 3.5 to obtain a $\mathrm{FO}^{<\omega}_{\kappa\aleph_0}[\Sigma_\alpha]$-Skolemisation $\Phi_{\alpha+1}$ of Φ_α such that

$$|\Phi_{\alpha+1}| \leq |\Phi_\alpha| \oplus \big|\mathrm{FO}^{<\omega}_{\kappa\aleph_0}[\Sigma_\alpha]\big| \quad \text{and} \quad |\Sigma_{\alpha+1}| \leq |\Sigma_\alpha| \oplus \big|\mathrm{FO}^{<\omega}_{\kappa\aleph_0}[\Sigma_\alpha]\big|.$$

We claim that the union $\Phi^+ := \bigcup_{\alpha<\kappa} \Phi_\alpha$ is the desired Skolemisation. Let $\Sigma^+ := \bigcup_{\alpha<\kappa} \Sigma_\alpha$. First, we show by induction on α that $|\Sigma_\alpha| \leq (|\Sigma| \oplus \aleph_0)^{<\kappa}$. Clearly, this holds for $\Sigma_0 = \Sigma$. For the successor step, we have

$$\begin{aligned} |\Sigma_{\alpha+1}| &\leq |\Sigma_\alpha| \oplus \left|\mathrm{FO}^{<\omega}_{\kappa\aleph_0}[\Sigma_\alpha]\right| \\ &\leq |\Sigma_\alpha| \oplus (|\Sigma_\alpha| \oplus \aleph_0)^{<\kappa} = (|\Sigma_\alpha| \oplus \aleph_0)^{<\kappa} \\ &\leq ((|\Sigma| \oplus \aleph_0)^{<\kappa} \oplus \aleph_0)^{<\kappa} = (|\Sigma| \oplus \aleph_0)^{<\kappa}. \end{aligned}$$

For limit ordinals $\delta < \kappa$, it follows that

$$|\Sigma_\delta| = \sup_{\alpha<\delta} |\Sigma_\alpha| \leq |\delta| \otimes (|\Sigma| \oplus \aleph_0)^{<\kappa} = (|\Sigma| \oplus \aleph_0)^{<\kappa}\,.$$

Consequently, we have

$$|\Sigma^+| = \sup_{\alpha<\kappa} |\Sigma_\alpha| \leq \kappa \otimes (|\Sigma| \oplus \aleph_0)^{<\kappa} = (|\Sigma| \oplus \aleph_0)^{<\kappa},$$

by Corollary A4.4.32. This implies that

$$|\Phi^+| \leq \left|\mathrm{FO}^0_{\kappa\aleph_0}[\Sigma^+]\right| \leq (|\Sigma^+| \oplus \aleph_0)^{<\kappa} \leq (|\Sigma| \oplus \aleph_0)^{<\kappa}.$$

Next, we prove that $(\Phi^+)^{\vDash}$ is a Skolem theory. Let $\exists y\varphi \in \mathrm{FO}^{<\omega}_{\kappa\aleph_0}[\Sigma^+]$. Since κ is regular it follows by induction on φ that $\exists y\varphi \in \mathrm{FO}^{<\omega}_{\kappa\aleph_0}[\Sigma_\alpha]$, for some $\alpha < \kappa$. Hence, there is a $\Sigma_{\alpha+1}$-term that defines a Skolem function for $\exists y\varphi$.

Finally, to show that Φ^+ is a Skolemisation of Φ it remains to prove that every model of Φ can be expanded to one of Φ^+. Let $\mathfrak{A} \vDash \Phi$ be a model of Φ. We construct a sequence $(\mathfrak{A}_\alpha)_{\alpha\leq\kappa}$ of models $\mathfrak{A}_\alpha \vDash \Phi_\alpha$ with $\mathfrak{A}_0 = \mathfrak{A}$ such that, for all $\alpha \leq \beta$, $\mathfrak{A}_\beta$ is an expansion of $\mathfrak{A}_\alpha$. $\mathfrak{A}_\kappa \vDash \Phi^+$ is the desired expansion of $\mathfrak{A}$.

We start with $\mathfrak{A}_0 := \mathfrak{A}$. For the successor step, suppose that $\mathfrak{A}_\alpha$ has already been defined. Since $\Phi_{\alpha+1}$ is a Skolemisation of Φ_α we can expand $\mathfrak{A}_\alpha$ to a $\Sigma_{\alpha+1}$-structure $\mathfrak{A}_{\alpha+1}$ such that $\mathfrak{A}_{\alpha+1} \vDash \Phi_{\alpha+1}$. For limit ordinals δ, we let $\mathfrak{A}_\delta$ be the 'union' of all the $\mathfrak{A}_\alpha$, $\alpha < \delta$, that is, its universe is A and, for each function $f \in \Sigma_\alpha$, we add the function $f^{\mathfrak{A}_\alpha}$ to $\mathfrak{A}_\delta$. (Note that this is well-defined since, if $f \in \Sigma_\alpha$ and $\alpha < \beta$ then $f^{\mathfrak{A}_\alpha} = f^{\mathfrak{A}_\beta}$.) □

An important application of the technique of Skolemisation is the following result.

Theorem 3.7 (Downward Löwenheim-Skolem Theorem).
Let $\Delta \subseteq \mathrm{FO}^{<\omega}_{\kappa\aleph_0}[\Sigma]$, *for a regular cardinal* κ, *and set* $\mu := |\Sigma| \oplus |\Delta| \oplus \kappa^-$ *where* $\kappa^- := \sup\{\lambda \mid \lambda < \kappa\}$.

For each Σ*-structure* $\mathfrak{A}$, *every subset* $X \subseteq A$, *and all cardinals* λ *with* $|X| \oplus \mu \leq \lambda \leq |A|$, *there exists a* Δ*-substructure* $\mathfrak{B} \preceq_\Delta \mathfrak{A}$ *of size* $|B| = \lambda$ *with* $X \subseteq B$.

Proof. Let Γ be the closure of Δ under subformulae, negation, and negation normal form. Since every formula $\varphi \in \mathrm{FO}^{<\omega}_{\kappa\aleph_0}[\Sigma]$ has less that κ subformulae it follows that $|\Gamma| \leq |\Delta| \otimes \kappa^-$. By Lemma 3.5, we can choose a Γ-Skolemisation $T^+ \subseteq \mathrm{FO}^{0}_{\kappa^+\aleph_0}[\Sigma^+]$ of $\mathrm{Th}_\Gamma(\mathfrak{A})$ such that

$$|T^+| \leq |\mathrm{Th}_\Gamma(\mathfrak{A})| \oplus |\Gamma| \quad \text{and} \quad |\Sigma^+| \leq |\Sigma| \oplus |\Gamma| \leq \mu\,.$$

Let $\mathfrak{A}^+$ be a Σ^+-expansion of $\mathfrak{A}$ such that $\mathfrak{A}^+ \vDash T^+$, and choose some set $X \subseteq Z \subseteq A$ of size $|Z| = \lambda$. By Corollary B3.1.11, the substructure $\mathfrak{B}^+ := \langle\!\langle Z \rangle\!\rangle_{\mathfrak{A}^+}$ has cardinality

$$\lambda = |Z| \leq |B^+| \leq |Z| \oplus |\Sigma^+| \oplus \aleph_0 = \lambda\,.$$

By Lemma 3.2, we have $\mathfrak{B}^+ \preceq_\Gamma \mathfrak{A}^+$. Let $\mathfrak{B}$ be the Σ-reduct of $\mathfrak{B}^+$. Then $\mathfrak{B} \preceq_\Delta \mathfrak{A}$, as desired. ☐

Corollary 3.8. *Let* $\mathfrak{A}$ *be a* Σ*-structure. For each set* $X \subseteq A$ *and every cardinal* $|X| \oplus |\Sigma| \oplus \aleph_0 \leq \kappa \leq |A|$, *there exists an elementary substructure* $\mathfrak{B} \preceq \mathfrak{A}$ *of size* $|B| = \kappa$ *such that* $X \subseteq B$.

Example. The field $\mathfrak{R} = \langle \mathbb{R}, +, \cdot, 0, 1\rangle$ of real numbers contains a countable elementary substructure $\mathfrak{R}_0 \prec \mathfrak{R}$.

We can generalise the technique of Skolemisation to $\mathrm{FO}_{\kappa\aleph_0}(\exists^\lambda)$ and $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{wo})$ in a straightforward way. As a result we obtain a variant of the Löwenheim-Skolem Theorem for these logics.

Theorem 3.9. *Let $\Delta \subseteq \mathrm{FO}^{<\omega}_{\kappa\aleph_0}(\exists^\lambda)[\Sigma]$, for a regular cardinal κ, and set $\mu := |\Sigma| \oplus |\Delta| \oplus \lambda \oplus \kappa^-$ where $\kappa^- := \sup\{\lambda \mid \lambda < \kappa\}$.*

For each Σ-structure $\mathfrak{A}$, every subset $X \subseteq A$, and all cardinals ν with $|X| \oplus \mu \leq \nu \leq |A|$, there exists a Δ-substructure $\mathfrak{B} \preceq_\Delta \mathfrak{A}$ of size $|B| = \nu$ with $X \subseteq B$.

Proof. The proof is analogous to that of Theorem 3.7. We adapt the notion of a Skolem function and a Δ-Skolemisation as follows. We say that a sequence $(t_i)_{i<\lambda}$ defines a Skolem function for a formula of the form $\exists^\lambda y\varphi(\bar{x}, y)$ if, for all $i, k < \lambda$ with $i \neq k$,

$$\Phi \vDash \forall\bar{x}(\exists^\lambda y\varphi(\bar{x}, y) \rightarrow \varphi(\bar{x}, t_i(\bar{x}))),$$
$$\Phi \vDash \forall\bar{x}(\exists^\lambda y\varphi(\bar{x}, y) \rightarrow t_i(\bar{x}) \neq t_k(\bar{x})).$$

A Δ-Skolemisation of Φ is a set $\Phi^+ \supseteq \Phi$ such that

- every model of Φ can be extended to one of Φ^+,
- for every formula $\exists y\varphi \in \Delta$, there is a term defining a Skolem function for $\exists y\varphi$,
- for every formula $\exists^\lambda y\varphi \in \Delta$, there is a sequence of terms defining a Skolem function for $\exists^\lambda y\varphi$.

With these definitions it follows as above that if $\mathfrak{A} \vDash \Phi^+$ and $\mathfrak{B} \subseteq \mathfrak{A}$ then $\mathfrak{B} \preceq_\Delta \mathfrak{A}$. Furthermore, for every set Φ, we can find a Δ-Skolemisation of size $|\Phi| \oplus |\Delta| \oplus \lambda$. Consequently, we can repeat the construction in the proof of Theorem 3.7. □

Theorem 3.10. *Let $\Delta \subseteq \mathrm{FO}^{<\omega}_{\kappa\aleph_0}(\mathrm{wo})[\Sigma]$, for a regular cardinal κ, and set $\mu := |\Sigma| \oplus |\Delta| \oplus \kappa^-$ where $\kappa^- := \sup\{\lambda \mid \lambda < \kappa\}$.*

For each Σ-structure $\mathfrak{A}$, every subset $X \subseteq A$, and all cardinals λ with $|X| \oplus \mu \leq \lambda \leq |A|$, there exists a Δ-substructure $\mathfrak{B} \preceq_\Delta \mathfrak{A}$ of size $|B| = \lambda$ with $X \subseteq B$.

Proof. We adapt the notion of a Skolem function and a Δ-Skolemisation as follows. A sequence $(t_n)_{n<\omega}$ defines a Skolem function for the formula

$\mathsf{W}\bar{x}\bar{y}\varphi(\bar{x},\bar{y},\bar{z})$ if, for all $n < \omega$,

$$\Phi \vDash \forall\bar{z}\big[\neg\mathsf{W}\bar{x}\bar{y}\varphi(\bar{x},\bar{y},\bar{z}) \to \varphi(t_{n+1}(\bar{z}), t_n(\bar{z}), \bar{z})\big],$$

that is, the sequence $(t_n)_n$ yields witnesses for the fact that the relation defined by φ is not well-founded.

A Δ-Skolemisation of Φ is a set $\Phi^+ \supseteq \Phi$ such that

- every model of Φ can be extended to one of Φ^+,
- for every formula $\exists y\varphi \in \Delta$, there is a term defining a Skolem function for $\exists y\varphi$,
- for every formula $\mathsf{W}\bar{x}\bar{y}\varphi \in \Delta$, there is a sequence of terms defining a Skolem function for $\mathsf{W}\bar{x}\bar{y}\varphi$.

With these definitions it follows as above that if $\mathfrak{A} \vDash \Phi^+$ and $\mathfrak{B} \subseteq \mathfrak{A}$ then $\mathfrak{B} \preceq_\Delta \mathfrak{A}$. (Note that, if $\varphi(\bar{x},\bar{y},\bar{c})^{\mathfrak{A}}$, for $\bar{c} \subseteq B$, is a well-order of its field then so is $\varphi(\bar{x},\bar{y},\bar{c})^{\mathfrak{A}} \cap B^n = \varphi(\bar{x},\bar{y},\bar{c})^{\mathfrak{B}}$. Conversely, if $\varphi(\bar{x},\bar{y},\bar{c})^{\mathfrak{A}}$ is not a well-order then the Skolem function yields an infinite strictly decreasing sequence of elements of $\mathfrak{B}$. Hence, $\varphi(\bar{x},\bar{y},\bar{c})^{\mathfrak{A}} \cap B^n$ is also not a well-order.)

Furthermore, for every set Φ, we can find a Δ-Skolemisation of size $|\Phi| \oplus |\Delta| \oplus \lambda$. Consequently, we can repeat the construction in the proof of Theorem 3.7. □

Exercise 3.1. Work out the missing details in the above proofs.

4. *The Compactness Theorem*

In this section we introduce an important method to construct models from diagrams. These models $\mathfrak{M}$ will have the additional nice property that every element is denoted by some term, that is, $\mathfrak{M} = \langle\!\langle \varnothing \rangle\!\rangle_{\mathfrak{M}}$.

Definition 4.1. Let $\Phi \subseteq \mathrm{FO}^{\circ}_{\infty\aleph_0}[\Sigma]$. A structure $\mathfrak{H}$ is a *Herbrand model* of Φ if $\mathfrak{H} \vDash \Phi$ and, for every $a \in H$, there is some term $t \in T[\Sigma, \varnothing]$ with $t^{\mathfrak{H}} = a$.

We start by characterising those sets of formulae that contain sufficient information to extract a model.

Definition 4.2. A set $\Phi \subseteq \mathrm{FO}^{\mathrm{o}}_{\infty\aleph_0}[\Sigma]$ is *=-closed* if

- $t = t \in \Phi$, for all $t \in T[\Sigma, \varnothing]$, and
- if $\varphi(x)$ is an atomic formula and $s, t \in T[\Sigma, \varnothing]$ are terms with $s = t \in \Phi$ then we have $\varphi(s) \in \Phi$ iff $\varphi(t) \in \Phi$.

Lemma 4.3. *Let $\Phi \subseteq \mathrm{FO}^{\mathrm{o}}_{\infty\aleph_0}[\Sigma]$ be =-closed. The relation*

$$s \sim t \quad :\text{iff} \quad s = t \in \Phi$$

is a congruence relation of the term algebra $\mathfrak{T}[\Sigma, \varnothing]$.

Proof. $\sim$ is reflexive since $t = t \in \Phi$, for all t. For symmetry, suppose that $s = t \in \Phi$ and set $\varphi(x) := x = s$. It follows that

$$\varphi(s) = s = s \in \Phi \quad \text{implies} \quad \varphi(t) = t = s \in \Phi\,.$$

Similarly, if $r = s \in \Phi$ and $s = t \in \Phi$ then setting $\varphi(x) := r = x$ we see that

$$\varphi(s) = r = s \in \Phi \quad \text{implies} \quad \varphi(t) = r = t \in \Phi\,.$$

Consequently, $\sim$ is an equivalence relation.

Suppose that $s_i \sim t_i$, for $i < n$, and let $f \in \Sigma$ be an n-ary function symbol. In the same way as above we can show, by induction on i, that

$$f s_0 \dots s_i s_{i+1} \dots s_{n-1} = f t_0 \dots t_i s_{i+1} \dots s_{n-1} \in \Phi\,.$$

It follows that $f^{\mathfrak{T}[\Sigma,\varnothing]}(s_0, \dots, s_{n-1}) \sim f^{\mathfrak{T}[\Sigma,\varnothing]}(t_0, \dots, t_{n-1})$, as desired. □

Lemma 4.4. *Every =-closed set of atomic sentences $\Phi \subseteq \mathrm{FO}^{\mathrm{o}}_{\infty\aleph_0}[\Sigma]$ has a Herbrand model $\mathfrak{H}$ such that*

$$\Phi = \{\, \varphi \mid \varphi \text{ atomic and } \mathfrak{H} \vDash \varphi \,\}\,.$$

Proof. By Lemma 4.3, the relation

$$s \sim t \quad :\text{iff} \quad s = t \in \Phi$$

is a congruence relation of the term algebra $\mathfrak{T}[\Sigma, \varnothing]$. Hence, we can take the quotient $\mathfrak{H}_0 := \mathfrak{T}[\Sigma, \varnothing]/\sim$. Let $\mathfrak{H}$ be the expansion of $\mathfrak{H}_0$ by relations

$$R^{\mathfrak{H}} := \{ \langle [t_0]_\sim, \dots, [t_{n-1}]_\sim \rangle \mid Rt_0 \dots t_{n-1} \in \Phi \},$$

for each n-ary relation $R \in \Sigma$. We claim that $\mathfrak{H}$ is the desired model.

Clearly, every element of $\mathfrak{H}$ is denoted by some term. Furthermore, by definition of $\mathfrak{H}$, we have $\mathfrak{H} \vDash \varphi$, for every $\varphi \in \Phi$. Conversely, suppose that $\mathfrak{H} \vDash \varphi$, for some atomic sentence φ. If $\varphi = s = t$ then we have $[s]_\sim = [t]_\sim$ which, by definition of $\mathfrak{H}$, implies that $s = t \in \Phi$. Similarly, if $\varphi = Rt_0 \dots t_{n-1}$ then $\langle [t_0]_\sim, \dots, [t_{n-1}]_\sim \rangle \in R^{\mathfrak{H}}$. Hence, there are terms $s_i \sim t_i$ such that $Rs_0 \dots s_{n-1} \in \Phi$. Since Φ is $=$-closed it follows that $Rt_0 \dots t_{n-1} \in \Phi$. ☐

We have shown how to construct a model for a set of atomic formulae. Next we turn to the case of formulae with quantifiers.

Definition 4.5. A *Hintikka set* is a set $\Phi \subseteq \mathrm{FO}^0_{\infty\aleph_0}[\Sigma]$ of sentences with the following closure properties:

(H1) Φ is $=$-closed.

(H2) If $\varphi \in \Phi$ then $\neg\varphi \notin \Phi$.

(H3) If $\neg\neg\varphi \in \Phi$ then $\varphi \in \Phi$.

(H4) If $\bigwedge \Psi \in \Phi$ then $\Psi \subseteq \Phi$.

(H5) If $\neg \bigwedge \Psi \in \Phi$ then there is some $\psi \in \Psi$ such that $\neg\psi \in \Phi$.

(H6) If $\bigvee \Psi \in \Phi$ then there is some $\psi \in \Psi$ such that $\psi \in \Phi$.

(H7) If $\neg \bigvee \Psi \in \Phi$ then $\neg\psi \in \Phi$, for all $\psi \in \Psi$.

(H8) If $\forall x \varphi(x) \in \Phi$ then $\varphi(t) \in \Phi$, for all $t \in T[\Sigma, \varnothing]$.

(H9) If $\neg\forall x \varphi(x) \in \Phi$ then there is some $t \in T[\Sigma, \varnothing]$ with $\neg\varphi(t) \in \Phi$.

(H10) If $\exists x \varphi(x) \in \Phi$ then there is some $t \in T[\Sigma, \varnothing]$ with $\varphi(t) \in \Phi$.

(H11) If $\neg\exists x \varphi(x) \in \Phi$ then $\neg\varphi(t) \in \Phi$, for all $t \in T[\Sigma, \varnothing]$.

Remark. Every elementary diagram is a Hintikka set.

Lemma 4.6. *Every Hintikka set $\Phi \subseteq \mathrm{FO}^{0}_{\infty\aleph_0}[\Sigma]$ has a Herbrand model.*

Proof. Let $\Phi_0 \subseteq \Phi$ consist of all atomic sentences in Φ. By the definition of a Hintikka set it follows that Φ_0 is =-closed. Hence, we can apply Lemma 4.4 to obtain a Herbrand model $\mathfrak{H}$ of Φ_0. We claim that $\mathfrak{H} \vDash \Phi$.

We prove by induction on the structure of a formula φ that

$$\varphi \in \Phi \text{ implies } \mathfrak{H} \vDash \varphi \quad \text{and} \quad \neg\varphi \in \Phi \text{ implies } \mathfrak{H} \vDash \neg\varphi\,.$$

If φ is atomic then the claim follows by Lemma 4.4.

Suppose that $\varphi = \neg\psi$. If $\varphi \in \Phi$ then we can apply the inductive hypothesis to ψ and it follows that $\mathfrak{H} \vDash \neg\psi$. Similarly, if $\neg\varphi \in \Phi$ then we have $\psi \in \Phi$, which implies that $\mathfrak{H} \vDash \psi$ and $\mathfrak{H} \vDash \neg\varphi$.

Consider the case that $\varphi = \bigwedge \Psi$. If $\bigwedge \Psi \in \Phi$ then $\Psi \subseteq \Phi$ implies that $\mathfrak{H} \vDash \psi$, for all $\psi \in \Psi$, and we have $\mathfrak{H} \vDash \bigwedge \Psi$. Analogously, if $\neg \bigwedge \Psi \in \Phi$ then there is some $\psi \in \Psi$ with $\neg\psi \in \Phi$. By inductive hypothesis it follows that $\mathfrak{H} \vDash \neg\psi$ which implies that $\mathfrak{H} \vDash \neg \bigwedge \Psi$.

Suppose that $\varphi = \forall x \psi(x)$. If $\varphi \in \Phi$ then $\psi(t) \in \Phi$, for all $t \in T[\Sigma, \varnothing]$. Hence, $\mathfrak{H} \vDash \psi(t)$, for all $t \in T[\Sigma, \varnothing]$. Since every element of H is denoted by a term it follows that $\mathfrak{H} \vDash \psi(a)$, for all $a \in H$, that is, $\mathfrak{H} \vDash \forall x \psi(x)$. Finally, if $\neg\forall x \psi(x) \in \Phi$ then there is some $t \in T[\Sigma, \varnothing]$ such that $\neg\psi(t) \in \Phi$. Therefore, we have $\mathfrak{H} \vDash \neg\psi(t)$ which implies that $\mathfrak{H} \vDash \neg\forall x \psi(x)$. The remaining cases are proved analogously. □

It is quite tedious to check that a set Φ satisfies conditions (H1)–(H11). The following lemma provides a simpler criterion for Φ being a Hintikka set.

Lemma 4.7. *Let $\Phi \subseteq \mathrm{FO}^{0}_{\infty\aleph_0}[\Sigma]$ be a set of sentences with the following properties:*

(1) *Every finite subset $\Phi_0 \subseteq \Phi$ is satisfiable.*

(2) *For every sentence* $\varphi \in \mathrm{FO}^0_{\infty\aleph_0}[\Sigma]$ *we have* $\varphi \in \Phi$ *or* $\neg\varphi \in \Phi$.

(3) *If* $\exists x\varphi(x) \in \Phi$ *then there exists some term* $t \in T[\Sigma, \varnothing]$ *such that* $\varphi(t) \in \Phi$.

(4) *If* $\bigvee \Psi \in \Phi$ *where* $|\Psi| \geq \aleph_0$ *then there is some* $\psi \in \Psi$ *with* $\psi \in \Phi$.

(5) *If* $\neg \bigwedge \Psi \in \Phi$ *where* $|\Psi| \geq \aleph_0$ *then there is some* $\psi \in \Psi$ *with* $\neg\psi \in \Phi$.

Then Φ *is a Hintikka set.*

Proof. First we show that

($*$) if $\Phi_0 \subseteq \Phi$ is finite and $\Phi_0 \vDash \varphi$ then $\varphi \in \Phi$.

Suppose otherwise. By (2), $\varphi \notin \Phi$ implies $\neg\varphi \in \Phi$. Hence, (1) implies that $\Phi_0 \cup \{\neg\varphi\}$ is satisfiable, and it follows that $\Phi_0 \nvDash \varphi$. A contradiction.

From ($*$) we can conclude that Φ satisfies (H1), (H3), (H4), (H7), (H8), and (H11). Furthermore, (1) implies (H2), and (3) and ($*$) imply that Φ satisfies (H9) and (H10).

It remains to prove (H5) and (H6). If $\Psi = \{\psi_0, \dots, \psi_{n-1}\}$ is finite then $\psi_0, \dots, \psi_{n-1} \in \Phi$ implies, by ($*$), that $\bigwedge \Psi \in \Phi$. Hence, $\neg \bigwedge \Psi \notin \Phi$. Similarly, If $\neg\psi_0, \dots, \neg\psi_{n-1} \in \Phi$ then it follows that $\bigvee \Psi \notin \Phi$. If, on the other hand, Ψ is infinite then (H5) and (H6) follow immediately from (4) and (5). ☐

Hintikka sets can be used to prove the Compactness Theorem which is the most fundamental result in first-order model theory. Most results in the remainder of this book are based on this theorem. It is frequently used to construct structures with some given properties. To do so, one describes the desired structure by a set of first-order formulae and then uses the Compactness Theorem to prove that this set of axioms is satisfiable.

Theorem 4.8 (Compactness Theorem). *Let* $\Phi \subseteq \mathrm{FO}[\Sigma, X]$ *be a set of first-order formulae and* $\varphi \in \mathrm{FO}[\Sigma, X]$.

(a) Φ *is satisfiable if and only if every finite subset* $\Phi_0 \subseteq \Phi$ *is satisfiable.*

(b) $\Phi \vDash \varphi$ *if and only if there exists a finite subset* $\Phi_0 \subseteq \Phi$ *such that* $\Phi_0 \vDash \varphi$.

Proof. Let us first prove that (a) implies (b). We have

$$\begin{aligned}\Phi \vDash \varphi \text{ iff } & \Phi \cup \{\neg\varphi\} \text{ is inconsistent} \\ \text{iff } & \text{there exists a finite subset } \Phi_0 \subseteq \Phi \text{ such that} \\ & \quad \Phi_0 \cup \{\neg\varphi\} \text{ is inconsistent} \\ \text{iff } & \text{there exists a finite subset } \Phi_0 \subseteq \Phi \text{ such that } \Phi_0 \vDash \varphi\,.\end{aligned}$$

It remains to prove (a). For the nontrivial direction, suppose that every finite subset of Φ is satisfiable. By replacing every free variable in Φ by a constant symbol we may assume that every formula in Φ is a sentence. We have to construct a model of Φ. By Lemma 4.6, it is sufficient to find a Hintikka set $\Psi \supseteq \Phi$.

We construct Ψ in stages. Let $\kappa := |\mathrm{FO}^\circ[\Sigma]| = |\Sigma| \oplus \aleph_0$. Let C be a set containing κ^+ constant symbols of each sort and set $\Sigma_C := \Sigma \cup C$. We fix an enumeration $(\varphi_\alpha)_{\alpha<\kappa^+}$ of $\mathrm{FO}^\circ[\Sigma_C]$ such that, for every $\psi \in \mathrm{FO}^\circ[\Sigma_C]$, the set $\{\, \alpha < \kappa^+ \mid \varphi_\alpha = \psi \,\}$ is cofinal in κ^+.

We construct an increasing sequence $(\Psi_\alpha)_{\alpha<\kappa^+}$ of sets $\Phi \subseteq \Psi_\alpha \subseteq \mathrm{FO}^\circ[\Sigma_C]$ such that every finite subset of Ψ_α is satisfiable and such that the limit $\Psi := \bigcup_\alpha \Psi_\alpha$ is a Hintikka set. By Lemma 4.7 it is sufficient to ensure that

- $\varphi_\alpha \in \Psi_{\alpha+1}$ or $\neg\varphi_\alpha \in \Psi_{\alpha+1}$,
- If $\varphi_\alpha = \exists x \vartheta$ and $\varphi_\alpha \in \Psi_{\alpha+1}$ then $\vartheta(c) \in \Psi_{\alpha+1}$, for some constant $c \in C$.

Set $\Psi_0 := \Phi$. For limit ordinals δ, we set $\Psi_\delta := \bigcup_{\alpha<\delta} \Psi_\alpha$. For the successor step, suppose that Ψ_α has already been defined. If every finite subset of $\Psi_\alpha \cup \{\varphi_\alpha\}$ is satisfiable then set $\psi := \varphi_\alpha$ else set $\psi := \neg\varphi_\alpha$. We claim that every finite subset of $\Psi_\alpha \cup \{\psi\}$ is satisfiable. If $\psi = \varphi_\alpha$ then this holds by choice of ψ. Hence, suppose that $\psi = \neg\varphi_\alpha$ and there is a finite subset $\Gamma_0 \subseteq \Psi_\alpha \cup \{\neg\varphi_\alpha\}$ that is inconsistent. By construction there is also a finite subset $\Gamma_1 \subseteq \Psi_\alpha \cup \{\varphi_\alpha\}$ which is inconsistent. Hence, $\Gamma_0 \vDash \varphi_\alpha$ and $\Gamma_1 \vDash \neg\varphi_\alpha$. It follows that $\Gamma := \Gamma_0 \cup \Gamma_1$ is a finite subset of Ψ_α with $\Gamma \vDash \varphi_\alpha \wedge \neg\varphi_\alpha$. Thus, Γ is inconsistent in contradiction to our assumption on Ψ_α.

We have found a set $\Psi_\alpha \cup \{\psi\}$ that satisfies the first of our conditions. If ψ is not of the form $\exists x \vartheta$ then we can set $\Psi_{\alpha+1} := \Psi_\alpha \cup \{\psi\}$ and we are done. Hence, suppose that $\psi = \exists x \vartheta(x)$. Since $|\Psi_\alpha| \leq \kappa$ we can find a constant symbol $c \in C$ that does not appear in Ψ_α. We define $\Psi_{\alpha+1} := \Psi_\alpha \cup \{\psi, \vartheta(c)\}$. Note that, since every finite subset of $\Psi_\alpha \cup \{\exists x \vartheta\}$ is satisfiable so is every finite subset of $\Psi_{\alpha+1}$. □

Exercise 4.1. Let $\varphi \in$ FO and $\Phi, T \subseteq$ FO. Prove that, if $\varphi \equiv \Phi$ modulo T then there exists a finite subset $\Phi_0 \subseteq \Phi$ such that $\varphi \equiv \bigwedge \Phi_0$ modulo T.

Exercise 4.2. Let $\mathcal{K}_i$, $i \in I$, be a family of first-order axiomatisable classes such that, $\bigcap_{i \in I_0} \mathcal{K}_i \neq \varnothing$, for every finite set $I_0 \subseteq I$. Show that $\bigcap_{i \in I} \mathcal{K}_i \neq \varnothing$.

Exercise 4.3. Let T be a first-order theory and $\mathfrak{A}$ a structure. Prove that $\mathfrak{A}$ can be embedded into some model of T if, and only if, every finitely generated substructure of $\mathfrak{A}$ can be embedded into some model of T.

We conclude this section with some simple applications of the Compactness Theorem. First, we show that first-order logic is not able to count.

Lemma 4.9. *Let Σ be an S-sorted signature and $s \in S$ a sort. There exists no set $\Phi \subseteq \mathrm{FO}^0[\Sigma]$ such that*

$$\mathfrak{A} \vDash \Phi \quad \text{iff} \quad |A_s| < \aleph_0 \,, \qquad \textit{for all } \Sigma\textit{-structures } \mathfrak{A} \,.$$

Proof. For a contradiction, suppose that there is such a set Φ. Let

$$\psi_n := \exists x_0 \cdots \exists x_{n-1} \bigwedge_{i<k} x_i \neq x_k$$

be the sentence expressing that there are at least n elements of sort s. We claim that

$$\Gamma := \Phi \cup \{\, \psi_n \mid n < \omega \,\}$$

is satisfiable. This yields the desired contradiction.

By the Compactness Theorem, we only need to check that every finite subset of Γ is satisfiable. If $\Gamma_0 \subseteq \Gamma$ is finite then there exists a number $k < \omega$ such that

$$\Gamma_0 \subseteq \Phi \cup \{ \psi_n \mid n < k \} .$$

Choose any finite Σ-structure $\mathfrak{A}$ with $|A_s| \geq k$. Since $\mathfrak{A}$ is finite we have $\mathfrak{A} \vDash \Phi$. Furthermore, $\mathfrak{A} \vDash \psi_n$, for all $n < k$. Hence, $\mathfrak{A}$ is a model of Γ_0. □

Example. Let us show that there is no set of first-order formulae expressing that a graph is connected. Suppose that $\Phi \subseteq \mathrm{FO}[E]$ is a set of formulae such that

$$\mathfrak{G} \vDash \Phi \quad \text{iff} \quad \mathfrak{G} \text{ is a connected undirected graph} .$$

We define formulae $\varphi_n(x, y)$ saying that there exists a path of length at most n from x to y by

$$\varphi_0(x, y) := x = y$$
$$\text{and} \quad \varphi_{n+1}(x, y) := \varphi_n(x, y) \vee \exists z(Exz \wedge \varphi_n(z, y)) .$$

Let c, d be new constant symbols and set

$$\Psi := \Phi \cup \{ \neg\varphi_n(c, d) \mid n < \omega \} .$$

Then Ψ is inconsistent since any model would be a connected graph that does not contain a path from c to d. Let $\Psi_0 \subseteq \Psi$ be a finite subset. There is some number k such that

$$\Psi_0 \subseteq \Phi \cup \{ \neg\varphi_n(c, d) \mid n < k \} .$$

Let $\mathfrak{P}_k$ be the graph consisting of a single path with k edges where the endpoints are denoted by c and d.

$$c - \bullet - \cdots - \bullet - d$$

Then we have $\mathfrak{P}_k \vDash \Psi_0$. Hence, every finite subset of Ψ is satisfiable and, by the Compactness Theorem, it follows that Ψ has a model. A contradiction.

Exercise 4.4. (a) Show that the class of all undirected, acyclic graphs is first-order axiomatisable. (A graph is *acyclic* if it does not contain a path $v_0, v_1, \dots, v_{n-1}, v_n, v_0$ where all the v_i are distinct.)

(b) Show that the class of all undirected graph that are not acyclic is not first-order axiomatisable.

(c) Use (b) to prove that the class of all undirected acyclic graphs is not finitely first-order axiomatisable.

Lemma 4.10. *The Compactness Theorem fails for* $\mathrm{FO}_{\kappa\aleph_0}[\Sigma]$ *if* $\kappa > \aleph_0$.

Proof. Let $\varphi_n := \exists x_0 \cdots \exists x_{n-1} \bigwedge_{i \neq k} x_i \neq x_k$ and

$$\varphi_{\mathrm{fin}} := \bigvee \{ \neg\varphi_n \mid n < \omega \} .$$

The set $\Phi := \{\varphi_{\mathrm{fin}}\} \cup \{ \varphi_n \mid n < \omega \}$ is unsatisfiable but each of its finite subsets has a model. □

Lemma 4.11. *Let* $\mathcal{K}$ *be a class of* Σ*-structures. If both* $\mathcal{K}$ *and* $\mathrm{Str}[\Sigma] \smallsetminus \mathcal{K}$ *are first-order axiomatisable then the class* $\mathcal{K}$ *is finitely axiomatisable.*

Proof. Let Φ^+ and Φ^- be sets such that

$$\mathcal{K} = \mathrm{Mod}(\Phi^+) \quad \text{and} \quad \mathrm{Str}[\Sigma] \smallsetminus \mathcal{K} = \mathrm{Mod}(\Phi^-) .$$

Then $\Phi^+ \cup \Phi^-$ is inconsistent. Hence, there are finite subsets $\Phi_0^+ \subseteq \Phi^+$ and $\Phi_0^- \subseteq \Phi^-$ such that $\Phi_0^+ \cup \Phi_0^-$ is inconsistent. Setting $\varphi := \bigwedge \Phi_0^-$ it follows that $\Phi^+ \vDash \neg\varphi_-$. Hence,

$$\mathfrak{A} \vDash \neg\varphi_- , \quad \text{for all } \mathfrak{A} \in \mathcal{K} .$$

Conversely, we have

$$\mathfrak{A} \vDash \varphi_- , \quad \text{for all } \mathfrak{A} \notin \mathcal{K} .$$

Consequently, $\mathrm{Mod}(\neg\varphi_-) = \mathcal{K}$, as desired. □

Generalising the idea behind Lemma 4.9 we obtain a converse to the Downward Löwenheim-Skolem Theorem.

Theorem 4.12 (Upward Löwenheim-Skolem-Tarski Theorem).
Let $T \subseteq \mathrm{FO}^{\circ}[\Sigma]$. If there exists a sort s such that, for every $n < \aleph_0$, T has a model $\mathfrak{A}$ with $|A_s| \geq n$ then T has models $\mathfrak{A}$ where $|A_s|$ has arbitrarily large cardinality.

Proof. Suppose that T has, for every $n < \aleph_0$, a model whose domain of sort s has size at least n. Let κ be an arbitrary cardinal and fix a set $C := \{ c_\alpha \mid \alpha < \kappa \}$ of κ constant symbols of sort s such that Σ and C are disjoint. We claim that the set

$$\Phi := T \cup \{ c \neq d \mid c, d \in C\,,\ c \neq d \}$$

has a model. By the Compactness Theorem, it is sufficient to show that every finite subset $\Phi_0 \subseteq \Phi$ is satisfiable. Since Φ_0 is finite, there exists a finite set $C_0 \subseteq C$ such that

$$\Phi_0 \subseteq T \cup \{ c \neq d \mid c, d \in C_0\,,\ c \neq d \}\,.$$

By assumption, there exists a model $\mathfrak{A} \vDash \varphi$ with $|A_s| \geq |C_0|$. We can turn it into a model of Φ_0 by interpreting the constant symbols $c \in C_0$ by distinct elements of A_s. □

The next example shows that, again, the above theorem fails for $\mathrm{FO}_{\kappa\aleph_0}$ with $\kappa > \aleph_0$. (Another counterexample is given by Lemma C1.1.7.)

Example. Let $\varphi \in \mathrm{FO}$ be a sentence axiomatising the class of ordered fields. The $\mathrm{FO}_{\aleph_1\aleph_0}$-sentence

$$\psi := \varphi \wedge \forall x \bigvee_{n<\omega} x < 1 + \cdots + 1$$

axiomatises the class of all archimedian ordered fields. It follows that ψ has only models of cardinality κ with $\aleph_0 \leq \kappa \leq 2^{\aleph_0}$.

As an immediate consequence of the Upward Löwenheim-Skolem-Tarski Theorem we obtain the result that infinite structures cannot be characterised up to isomorphism in first-order logic.

Corollary 4.13. *If $\mathfrak{A}$ is a structure with at least one infinite domain then there exists no set $\Phi \subseteq \mathrm{FO}$ such that*

$$\mathfrak{B} \vDash \Phi \quad \text{iff} \quad \mathfrak{B} \cong \mathfrak{A} \,.$$

5. *Amalgamation*

We can use the Upward Löwenheim-Skolem-Tarski Theorem to construct elementary extensions of a single structure. In this section we present a way to find a common elementary extension of several structures.

Definition 5.1. Let L be a logic.

(a) For sets $\Phi, \Delta \subseteq L$ of formulae, we define the set of all *Δ-consequences* of Φ by

$$\Phi^{\vDash}_{\Delta} := \Phi^{\vDash} \cap \Delta \,.$$

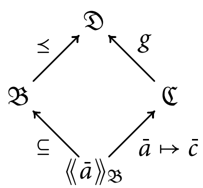
(b) Suppose that L is algebraic. For structures $\mathfrak{A}$ and $\mathfrak{B}$ and tuples $\bar{a} \subseteq A$ and $\bar{b} \subseteq B$, we write

$$\langle \mathfrak{A}, \bar{a} \rangle \leq_{\Delta} \langle \mathfrak{B}, \bar{b} \rangle \quad :\text{iff} \quad \mathfrak{A} \vDash \varphi(\bar{a}) \text{ implies } \mathfrak{B} \vDash \varphi(\bar{b}) \,, \quad \text{for all } \varphi \in \Delta \,.$$

Theorem 5.2 (Amalgamation Theorem). *Let $\mathfrak{B}$ and $\mathfrak{C}$ be Σ-structures, $\Delta \subseteq \mathrm{FO}$, and $\bar{a} \subseteq B$, $\bar{c} \subseteq C$ sequences such that*

$$\langle \mathfrak{C}, \bar{c} \rangle \leq_{\exists\Delta} \langle \mathfrak{B}, \bar{a} \rangle \,.$$

There exists an elementary extensions $\mathfrak{D} \succeq \mathfrak{B}$ and a Δ-map $g : \mathfrak{C} \to \mathfrak{D}$ with $g(\bar{c}) = \bar{a}$.



Proof. By taking an isomorphic copy of $\mathfrak{C}$ we may assume that $\bar{a} = \bar{c}$ and $B \cap C = \bar{a}$. To find the desired structure $\mathfrak{D}$ we prove that

$$T := \mathrm{Th}(\mathfrak{B}_B) \cup \mathrm{Th}_\Delta(\mathfrak{C}_C)$$

is satisfiable. By the Compactness Theorem, it is sufficient to show that every finite subset $T_0 \subseteq T$ has a model. Given $T_0 \subseteq T$ set

$$\varphi(\bar{a}, \bar{d}) := \bigwedge(T_0 \cap \mathrm{Th}_\Delta(\mathfrak{C}_C))$$

where $\bar{d} \subseteq C \setminus \bar{a}$. Suppose, for a contradiction, that

$$\mathrm{Th}(\mathfrak{B}_B) \vDash \neg\exists\bar{y}\varphi(\bar{a}, \bar{y}) .$$

Then we have $\langle\mathfrak{B}, \bar{a}\rangle \vDash \neg\exists\bar{y}\varphi(\bar{a}, \bar{y})$ and, since $\exists\bar{y}\varphi \in \exists\Delta$, it follows that $\langle\mathfrak{C}, \bar{a}\rangle \vDash \neg\exists\bar{y}\varphi(\bar{a}, \bar{y})$. Consequently, we have $\mathrm{Th}_\Delta(\mathfrak{C}_C) \vDash \neg\varphi(\bar{a}, \bar{d})$. Contradiction.

Since $\mathrm{Th}(\mathfrak{B}_B)$ is complete it follows that $\mathrm{Th}(\mathfrak{B}_B) \vDash \exists\bar{y}\varphi(\bar{a}, \bar{y})$. Thus, there exists some tuple $\bar{b} \subseteq B$ such that $\mathfrak{B}_B \vDash \varphi(\bar{a}, \bar{b})$. The structure $\langle\mathfrak{B}_B, \bar{b}\rangle \vDash T_0$ is our desired model.

We have shown that there exists a model $\mathfrak{D} \vDash T$. Since $\mathfrak{D} \vDash \mathrm{Th}(\mathfrak{B}_B)$ there exists an elementary embedding $h : \mathfrak{B} \to \mathfrak{D}$ and, by taking isomorphic copies, we may assume that $\mathfrak{D} \succeq \mathfrak{B}$. We define a function $g : C \to D$ by setting $g(d) := d^{\mathfrak{D}}$, for $d \in C$. ($d^{\mathfrak{D}}$ is the value of the constant symbol d in $\mathfrak{D}$.) Since $\mathfrak{D} \vDash \mathrm{Th}_\Delta(\mathfrak{C}_C)$ it follows that $g : \mathfrak{C} \to \mathfrak{D}$ is a Δ-map. Furthermore, we have $g(\bar{c}) = \bar{c}^{\mathfrak{D}} = \bar{a}$. □

Corollary 5.3. *If $\mathfrak{A} \equiv \mathfrak{B}$ then there exists a structure $\mathfrak{C}$ such that $\mathfrak{A} \preceq \mathfrak{C}$ and $\mathfrak{B} \preceq \mathfrak{C}$.*

Let us record a special instance of the Amalgamation Theorem that will be used in the next section.

Corollary 5.4. *Let $\mathfrak{B}$ and $\mathfrak{C}$ be Σ-structures and $\bar{a} \subseteq B$ a sequence of elements. If $f : \langle\bar{a}\rangle \to C$ is a homomorphism such that*

$$\langle\mathfrak{C}, f\bar{a}\rangle \leq_\exists \langle\mathfrak{B}, \bar{a}\rangle ,$$

then there exists an elementary extension $\mathfrak{D} \succeq \mathfrak{B}$ and an embedding $g : \mathfrak{C} \to \mathfrak{D}$ such that $g f(\bar{a}) = \bar{a}$.

Lemma 5.5. *Let $T, \Delta \subseteq$ FO where Δ is closed under disjunctions. Then $\mathfrak{A} \vDash T_\Delta^\vDash$ if, and only if, there exists a model $\mathfrak{B} \vDash T$ such that $\mathfrak{B} \leq_\Delta \mathfrak{A}$.*

Proof. ($\Leftarrow$) Obviously, $\mathfrak{B} \vDash T_\Delta^\vDash$ and $\mathfrak{B} \leq_\Delta \mathfrak{A}$ implies that $\mathfrak{A} \vDash T_\Delta^\vDash$.

($\Rightarrow$) Set $\Phi := \mathrm{Th}_{\Delta^\neg}(\mathfrak{A})$ where $\Delta^\neg := \{ \neg\varphi \mid \varphi \in \Delta \}$. It is sufficient to find a model $\mathfrak{B}$ of $\Psi := \Phi \cup T$. If Ψ is unsatisfiable then there exists a finite subset $\{\varphi_0, \dots, \varphi_k\} \subseteq \Phi$ such that

$$T \vDash \neg\varphi_0 \vee \dots \vee \neg\varphi_k \,.$$

Suppose that $\varphi_i = \neg\psi_i$, for $\psi_i \in \Delta$. Then $T \vDash \psi_0 \vee \dots \vee \psi_k$ implies that $\psi_0 \vee \dots \vee \psi_k \in T_\Delta^\vDash$ and, hence, $\mathfrak{A} \vDash \psi_0 \vee \dots \vee \psi_k$ in contradiction to $\mathfrak{A} \vDash \varphi_i$, for all $i \leq k$. ☐

Corollary 5.6. *Let $T, \Delta \subseteq$ FO where Δ is closed under disjunctions, and set $\Delta^\neg := \{ \neg\varphi \mid \varphi \in \Delta \}$. For every model $\mathfrak{A} \vDash T_{\forall\Delta}^\vDash$, there exists a model $\mathfrak{B} \vDash T$ and a $\Delta^\neg$-map $g : \mathfrak{A} \to \mathfrak{B}$.*

Proof. Suppose that $\mathfrak{A} \vDash T_{\forall\Delta}^\vDash$. By Lemma 5.5, we can find a model $\mathfrak{C} \vDash T$ such that $\mathfrak{A} \leq_{\exists\Delta^\neg} \mathfrak{C}$. By the Amalgamation Theorem, it follows that there exists some elementary extension $\mathfrak{B} \succeq \mathfrak{C}$ and a $\Delta^\neg$-map $g : \mathfrak{A} \to \mathfrak{B}$. ☐

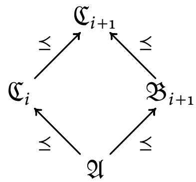
We can amalgamate several structures by iterating the Amalgamation Theorem.

Lemma 5.7. *Let $\mathfrak{B}_i$, $i < \alpha$, be a family of structures and suppose that $\mathfrak{A} \subseteq \mathfrak{B}_i$, for all $i < \alpha$, is a common substructure with universe $A = B_i \cap B_k$, for all $i \neq k$. There exists a structure $\mathfrak{C}$ such that $\mathfrak{B}_i \preceq \mathfrak{C}$, for all $i < \alpha$.*

Proof. We construct an elementary chain $(\mathfrak{C}_i)_{i<\alpha}$ such that $\mathfrak{B}_i \preceq \mathfrak{C}_i$, for $i < \alpha$. The structure $\mathfrak{C} := \bigcup_{i<\alpha} \mathfrak{C}_i$ has the desired properties since $\mathfrak{B}_i \preceq \mathfrak{C}_i \preceq \mathfrak{C}$.

We define $\mathfrak{C}_i$ by induction on i. We start with $\mathfrak{C}_0 := \mathfrak{B}_0$ and, for limit ordinals δ, we set $\mathfrak{C}_\delta := \bigcup_{i<\delta} \mathfrak{C}_i$. For the successor step, we can apply the

Amalgamation Theorem to obtain a common elementary extension $\mathfrak{C}_{i+1}$ of $\mathfrak{C}_i$ and $\mathfrak{B}_{i+1}$.



∎

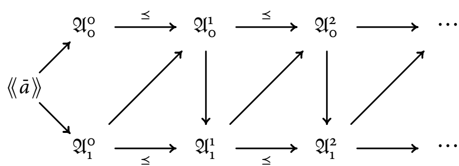
We conclude this section with an amalgamation theorem for expansions instead of extensions. We also record two applications.

Theorem 5.8. *Let Γ_0 and Γ_1 be signatures and set $\Sigma := \Gamma_0 \cap \Gamma_1$. Suppose that $\mathfrak{A}_i$ is a Γ_i-structure, for $i < 2$, and let $\bar{a} \subseteq A_0 \cap A_1$ be a sequence such that*

$$\langle \mathfrak{A}_0|_\Sigma, \bar{a} \rangle \equiv \langle \mathfrak{A}_1|_\Sigma, \bar{a} \rangle .$$

Then there exists a $(\Gamma_0 \cup \Gamma_1)$-structure $\mathfrak{B}$ with $\mathfrak{A}_0 \preceq \mathfrak{B}|_{\Gamma_0}$ and an elementary embedding $g : \mathfrak{A}_1 \to \mathfrak{B}|_{\Gamma_1}$ with $g(\bar{a}) = \bar{a}$.

Proof. We construct structures $\mathfrak{A}_i^n$ for $i < 2$ and $n < \omega$ as follows. We start with $\mathfrak{A}_i^0 := \mathfrak{A}_i$. If $\mathfrak{A}_0^n$ and $\mathfrak{A}_1^n$ are already defined then we apply the Amalgamation Theorem twice. First, we use it to obtain an elementary extension $\mathfrak{A}_0^{n+1} \succeq \mathfrak{A}_0^n$ and an elementary embedding $\mathfrak{A}_1^n|_\Sigma \to \mathfrak{A}_0^{n+1}|_\Sigma$. Then we construct an elementary extension $\mathfrak{A}_1^{n+1} \succeq \mathfrak{A}_1^n$ and an elementary embedding $\mathfrak{A}_0^{n+1}|_\Sigma \to \mathfrak{A}_1^{n+1}|_\Sigma$.



Let $\mathfrak{B}_i := \bigcup_n \mathfrak{A}_i^n$. The elementary embeddings induce an isomorphism $h : \mathfrak{B}_0|_\Sigma \to \mathfrak{B}_1|_\Sigma$. We use h to expand the Γ_0-structure $\mathfrak{B}_0$ to a $(\Gamma_0 \cup \Gamma_1)$-structure $\mathfrak{B}$ by setting

$$\xi^{\mathfrak{B}} := h^{-1}[\xi^{\mathfrak{B}_1}], \qquad \text{for } \xi \in \Gamma_1 \smallsetminus \Sigma .$$

Since $\mathfrak{B}|_{\Gamma_0} \geq \mathfrak{A}_0$ the claim follows. □

Corollary 5.9. *Let $T \subseteq \mathrm{FO}^0[\Sigma]$ and $\mathfrak{A}$ a Σ_0-structure where $\Sigma_0 \subseteq \Sigma$. We have $\mathfrak{A} \vDash T^{\vDash} \cap \mathrm{FO}^0[\Sigma_0]$ if and only if $\mathfrak{A} \leq \mathfrak{B}|_{\Sigma_0}$ for some model $\mathfrak{B}$ of T.*

Proof. ($\Leftarrow$) is trivial. For ($\Rightarrow$), we set $\Delta := \mathrm{FO}^0[\Sigma_0]$ and we assume that $\mathfrak{A} \vDash T^{\vDash}_{\Delta}$. We can use Lemma 5.5 to find a model $\mathfrak{C} \vDash T$ such that $\mathfrak{C} \leq_\Delta \mathfrak{A}$. By choice of Δ this implies that $\mathfrak{C}|_{\Sigma_0} \equiv \mathfrak{A}|_{\Sigma_0}$. Applying Theorem 5.8 we obtain an elementary extension $\mathfrak{B} \geq \mathfrak{C}$ and the desired elementary map $g : \mathfrak{A} \to \mathfrak{B}|_{\Sigma_0}$. □

The second application is the Interpolation Theorem of Craig. We will prove a much more general version in Section C5.5.

Theorem 5.10 (Craig). *Let Γ_0 and Γ_1 be signatures and set $\Sigma := \Gamma_0 \cap \Gamma_1$. Suppose that $\varphi_0 \vDash \varphi_1$ where $\varphi_0 \in \mathrm{FO}^0[\Gamma_0]$ and $\varphi_1 \in \mathrm{FO}^0[\Gamma_1]$. Then there exists a formula $\psi \in \mathrm{FO}^0[\Sigma]$ such that*

$$\varphi_0 \vDash \psi \quad \textit{and} \quad \psi \vDash \varphi_1 .$$

Proof. If φ_0 is inconsistent, we can set $\psi :=$ false. Hence, suppose that φ_0 has a model $\mathfrak{A}_0$ and set $\Psi := \mathrm{Th}(\mathfrak{A}_0|_\Sigma)$.

If $\Psi \vDash \varphi_1$, then we can use the Compactness Theorem to find a finite subset $\Psi_0 \subseteq \Psi$ with $\Psi_0 \vDash \varphi_1$. Hence, $\psi := \bigwedge \Psi_0$ is the desired formula.

Suppose that $\Psi \nvDash \varphi_1$. Then $\Psi \cup \{\neg\varphi_1\}$ has a model $\mathfrak{A}_1$. Since

$$\mathrm{Th}(\mathfrak{A}_1|_\Sigma) = \Psi = \mathrm{Th}(\mathfrak{A}_0|_\Sigma) ,$$

we can use Theorem 5.8, to find a $(\Gamma_0 \cup \Gamma_1)$-structure $\mathfrak{B}$ with

$$\mathrm{Th}(\mathfrak{B}|_{\Gamma_0}) = \mathrm{Th}(\mathfrak{A}_0) \quad \text{and} \quad \mathrm{Th}(\mathfrak{B}|_{\Gamma_1}) = \mathrm{Th}(\mathfrak{A}_1) .$$

In particular, we have $\mathfrak{B} \vDash \varphi_0$ and $\mathfrak{B} \vDash \neg\varphi_1$. Consequently, $\varphi_0 \nvDash \varphi_1$. A contradiction. □

C3. *Types and type spaces*

1. *Types*

In the same way that we can classify structures by their theory, we can distinguish elements of a structure by the formulae they satisfy. Such theories of elements are called *types.*

Definition 1.1. Let L be a logic.

(a) A (*partial*) *L-type* is a satisfiable set of L-formulae.

(b) An L-type $\mathfrak{p}$ is *complete* if it is a complete L-theory.

(c) We denote by $S(L)$ the set of all complete L-types.

(d) For $\Phi \subseteq L$, we define the set

$$\langle \Phi \rangle_L := \{ \mathfrak{p} \in S(L) \mid \Phi \subseteq \mathfrak{p} \}$$

of all types containing Φ. Usually we will omit the index L and just write $\langle \Phi \rangle$. Furthermore, for single formulae φ we write $\langle \varphi \rangle$ instead of $\langle \{\varphi\} \rangle$.

Example. For boolean logic $\mathrm{BL}(\mathfrak{B})$ introduced in Section C1.1, interpretations are ultrafilters and the theory of an ultrafilter $\mathfrak{u}$ is $\mathfrak{u}$ itself. Hence,

$$\begin{aligned} S(\mathrm{BL}(\mathfrak{B})) &= \{ \mathrm{Th}(\mathfrak{u}) \mid \mathfrak{u} \in \mathrm{spec}(\mathfrak{B}) \} \\ &= \{ \mathfrak{u} \mid \mathfrak{u} \in \mathrm{spec}(\mathfrak{B}) \} = \mathrm{spec}(\mathfrak{B}) \,. \end{aligned}$$

Definition 1.2. Let L be an algebraic logic and $\bar{s}$ a sequence of sorts.

(a) Let $\mathfrak{M}$ be a Σ-structure. The *L-type* of a tuple $\bar{a} \in M^{\bar{s}}$ is the set

$$\mathrm{tp}_L(\bar{a}/\mathfrak{M}) := \{ \varphi(\bar{x}) \in L^{\bar{s}}[\Sigma] \mid \mathfrak{M} \vDash \varphi(\bar{a}) \}.$$

If the structure $\mathfrak{M}$ is known from the context we will omit it and simply write $\mathrm{tp}_L(\bar{a})$. Similarly, we omit the index L in case $L = \mathrm{FO}$.

(b) Let $T \subseteq L^{\circ}[\Sigma]$ be an L-theory. An *$\bar{s}$-type of T* is an L-type $\mathfrak{p} \subseteq L^{\bar{s}}[\Sigma]$ such that $\mathfrak{p} \cup T$ is consistent. The set of all complete $\bar{s}$-types of T is

$$S_L^{\bar{s}}(T) := \{ \mathfrak{p} \in S(L^{\bar{s}}[\Sigma]) \mid T \subseteq \mathfrak{p} \}.$$

An *α-type of T*, for an ordinal α, is an $\bar{s}$-type of T where $|\bar{s}| = \alpha$. The set of all complete α-types is

$$S_L^{\alpha}(T) := \bigcup \{ S_L^{\bar{s}}(T) \mid |\bar{s}| = \alpha \}.$$

(c) We also need types with parameters. If $\mathfrak{M}$ is a model of T and $U \subseteq M$ then we say that a type of $T(U)$ is a type *over U*. In particular, the set $\mathrm{tp}_L(\bar{a}/U) := \mathrm{tp}_L(\bar{a}/\mathfrak{M}_U)$ is the L-type of $\bar{a}$ over U. We set

$$S_L^{\bar{s}}(U) := S_L^{\bar{s}}(T(U)).$$

To simplify notation, we define $S_L^{<\omega}(U) := \bigcup_{n<\omega} S_L^n(U)$. Again, we usually omit the index if $L = \mathrm{FO}$.

(d) An $\bar{s}$-type $\mathfrak{p}$ over U is *realised* in $\mathfrak{M}$ if there is some tuple $\bar{a} \in M^{\bar{s}}$ such that $\mathfrak{p} \subseteq \mathrm{tp}_L(\bar{a}/U)$. Otherwise, we say that $\mathfrak{M}$ *omits* $\mathfrak{p}$.

Example. Let $\mathfrak{N} = \langle \omega, s, 0 \rangle$ where $s(n) := n + 1$ is the successor function. We have

$$S^1(\varnothing) = \{\mathfrak{p}_0, \mathfrak{p}_1, \dots, \mathfrak{p}_\infty\}$$

where, for $n < \omega$,

$$\mathfrak{p}_n := \mathrm{tp}(n) \vDash x_0 = s^n(0),$$

and $\mathfrak{p}_\infty \vDash x_0 \neq s^n(0)$, for all n. Hence, $\mathfrak{p}_\infty$ is not realised in $\mathfrak{N}$.

Example. Consider $\langle \mathbb{Q}, < \rangle$. The elements of $S^1(\mathbb{Q})$ correspond to the set of cuts. For every cut $\langle A, B \rangle$ of $\mathbb{Q}$, i.e., every partition $A \cup B = \mathbb{Q}$ such that A is an initial segment and B is a final one, there exists a non-realised type $\mathfrak{p}$ such that

$$\begin{aligned} &\mathfrak{p} \vDash x > a &&\text{for all } a \in A\,, \\ \text{and}\quad &\mathfrak{p} \vDash x < b &&\text{for all } b \in B\,. \end{aligned}$$

It follows that $|S^1(\mathbb{Q})| = 2^{\aleph_0}$. Depending on whether A has a maximal element or B has a minimal one, we obtain the following classification.

(*realised*) For each $a \in \mathbb{Q}$, we have a type $\mathfrak{p} \vDash x = a$, i.e., $\mathfrak{p} = \operatorname{tp}(a/\mathbb{Q})$.

(a^+) For each $a \in \mathbb{Q}$, there exists a type $\mathfrak{p}$ of an element 'immediately above a'. That is,

$$\begin{aligned} &\mathfrak{p} \vDash x > b &&\text{for all } b \leq a\,, \\ \text{and}\quad &\mathfrak{p} \vDash x < b &&\text{for all } b > a\,. \end{aligned}$$

(a^-) Similarly, for each $a \in \mathbb{Q}$, we have the type of an element 'immediately below a'.

($+\infty$) We have one type $\mathfrak{p}$ of an infinite positive element. That is,

$$\mathfrak{p} \vDash x > a \qquad \text{for all } a \in \mathbb{Q}.$$

($-\infty$) Similarly, there is the type of an infinite negative element.

(*irrational*) Finally, for each cut $\langle A, B \rangle$ such that A has no maximal element and B has no minimal one, there is one type $\mathfrak{p}$ such that

$$\begin{aligned} &\mathfrak{p} \vDash x > a &&\text{for all } a \in A\,, \\ \text{and}\quad &\mathfrak{p} \vDash x < b &&\text{for all } b \in B\,. \end{aligned}$$

Exercise 1.1. Let $T := \operatorname{Th}(\mathfrak{Z})$ where $\mathfrak{Z} := \langle \mathbb{Z}, s \rangle$ and $s : x \mapsto x + 1$ is the successor function. Determine $S^n(T)$, for every $n < \omega$. In particular, compute $|S^n(T)|$. *Hint.* Note that, modulo T, every formula is equivalent to a quantifier-free one.

The set of types of $L|_\Phi$ and L/Φ can be computed as follows.

Lemma 1.3. *Let L be a logic and $\Phi \subseteq L$.*

(a) $S(L/\Phi) = \langle\Phi\rangle_L \subseteq S(L)$.

(b) $S(L|_\Phi) = \{\, \mathfrak{p} \cap \Phi \mid \mathfrak{p} \in S(L) \,\}$.

Proof. (a) We have

$$\begin{aligned} \mathfrak{p} \in S(L/\Phi) \quad &\text{iff} \quad \mathfrak{p} = \mathrm{Th}_{L/\Phi}(\mathfrak{J}) \text{ for some } \mathfrak{J} \in \mathrm{Mod}_L(\Phi) \\ &\text{iff} \quad \mathfrak{p} = \mathrm{Th}_L(\mathfrak{J}) \text{ for some } \mathfrak{J} \vDash \Phi \\ &\text{iff} \quad \mathfrak{p} \in S(L) \text{ and } \Phi \subseteq \mathfrak{p} . \end{aligned}$$

(b) We have

$$\begin{aligned} S(L|_\Phi) &= \big\{\, \mathrm{Th}_{L|_\Phi}(\mathfrak{J}) \bigm| \mathfrak{J} \text{ an } L\text{-interpretation} \,\big\} \\ &= \big\{\, \mathrm{Th}_L(\mathfrak{J}) \cap \Phi \bigm| \mathfrak{J} \text{ an } L\text{-interpretation} \,\big\} \\ &= \{\, \mathfrak{p} \cap \Phi \mid \mathfrak{p} \in S(L) \,\} . \end{aligned}$$

□

The relationship between a logic L and its set of types $S(L)$ is similar to that between a boolean algebra $\mathfrak{B}$ and its spectrum $\mathrm{spec}(\mathfrak{B})$. In fact, if L is boolean closed there exists an embedding $S(L) \to \mathrm{spec}(\mathfrak{Lb}(L))$.

Lemma 1.4. *Let L be a logic that is closed under disjunction and conjunction and that contains an unsatisfiable formula.*

(a) *If $\Phi \subseteq L$ then $\Phi^\vDash$ is a filter of $\langle L, \vDash\rangle$ and $\Phi^\vDash/\equiv$ is a filter of $\mathfrak{Lb}(L)$.*

(b) *Every complete L-theory T is an ultrafilter of $\langle L, \vDash\rangle$.*

Proof. Since (a) is obvious, we only need to prove (b). By (a), we know that $T = T^\vDash$ is a filter. Since there is an unsatisfiable formula, this filter is proper.

To prove that T is an ultrafilter consider a disjunction $\varphi \vee \psi \in T$. Since T is complete there exists an interpretation $\mathfrak{J}$ with $\mathrm{Th}_L(\mathfrak{J}) = T$. Hence, $\mathfrak{J} \vDash \varphi \vee \psi$ implies that $\mathfrak{J} \vDash \varphi$ or $\mathfrak{J} \vDash \psi$. In the former case we have $\varphi \in T$ and, otherwise, we have $\psi \in T$. □

Remark. If $\mathfrak{u}$ is a proper filter of $\mathfrak{Lb}(L)$ then the finite intersection property implies that every finite subset of $\mathfrak{u}$ is satisfiable.

In general the converse of statement (b) is not true, but there are some logics where every ultrafilter is a type. We have already seen in Section 1 that this is the case for boolean logic. A more important example of this phenomenon is first-order logic.

Lemma 1.5. *Every ultrafilter $\mathfrak{u}$ of $\langle \mathrm{FO}[\Sigma, X], \vDash \rangle$ is a complete type.*

Proof. If $\mathfrak{u}$ is an ultrafilter, it has the finite intersection property. Hence, every finite subset $\Phi \subseteq \mathfrak{u}$ is satisfiable. By the Compactness Theorem it follows that $\mathfrak{u}$ is satisfiable. Consequently, $\mathfrak{u}$ is a type. Since FO is boolean closed we can use Theorem B2.4.11 and Lemma C1.3.4 (d) to show that $\mathfrak{u}$ is complete. □

Corollary 1.6. *We have*

$$S(\mathrm{FO}[\Sigma, X]) = \mathrm{spec}(\langle \mathrm{FO}[\Sigma, X], \vDash \rangle) .$$

Remark. In the next section we will see that the Stone topology on the spectrum induces a topology on the type space where the closed sets are precisely those of the form $\langle \Phi \rangle$, for $\Phi \subseteq \mathrm{FO}$. The name 'Compactness Theorem' stems from the fact that this theorem implies that the topology obtained in this way is compact.

For logics where the Compactness Theorem fails, there are ultrafilters that do not correspond to types. In fact, the Compactness Theorem is equivalent to the statement of Lemma 1.5.

Example. There are ultrafilters of $\mathfrak{Lb}(\mathrm{FO}_{\aleph_1\aleph_0}[\Sigma])$ which are not types. Let $\psi := \bigwedge_{n<\omega} \varphi_n$ where φ_n is the formula stating that there are at least n elements. The formula $\neg\psi \wedge \varphi_n$ is satisfiable, for every n. Hence, the set $\{\neg\psi\} \cup \{ \varphi_n \mid n < \omega \}$ has the finite intersection property and there exists an ultrafilter

$$\mathfrak{u} \supseteq \{\neg\psi\} \cup \{ \varphi_n \mid n < \omega \} .$$

This ultrafilter is not a type since it is not satisfiable.

This example shows that, for $\kappa > \aleph_0$, the inclusion $S(\mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]) \subset \mathrm{spec}(\langle \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X], \vDash\rangle)$ is proper. We can describe the subset of the spectrum corresponding to $S(\mathrm{FO}_{\kappa\aleph_0}[\Sigma, X])$ as follows. Using Chang's reduction we can find a signature $\Sigma_+ \supseteq \Sigma$ and a first-order theory $T \subseteq \mathrm{FO}[\Sigma_+, X]$ such that

$$\mathrm{spec}(\langle \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X], \vDash\rangle) \cong S(T)\,.$$

Then we can characterise $S(\mathrm{FO}_{\kappa\aleph_0})$ as a subset of $S(T)$ by describing the types in $S(T) \setminus S(\mathrm{FO}_{\kappa\aleph_0})$.

Lemma 1.7. *Let $\varphi \in \mathrm{FO}_{\kappa^+\aleph_0}[\Sigma, X]$ and $|\Sigma| \leq \kappa$. There exists a signature $\Sigma_+ \supseteq \Sigma$ and set C of (partial) $\mathrm{FO}[\Sigma_+, X]$-types such that*

$$\mathfrak{M} \vDash \varphi \quad \text{iff} \quad \textit{there is some } \Sigma_+\textit{-expansion of } \mathfrak{M} \textit{ that omits every type in } C.$$

Furthermore, we can choose Σ_+ and C of size at most κ.

Proof. By Lemma c1.4.12, there exists an $\mathrm{FO}_{\kappa\aleph_0}$-theory T_φ such that $\mathfrak{M} \vDash \varphi$ if, and only if, some expansion of $\mathfrak{M}$ satisfies T_φ. We define a set C of types such that $\mathfrak{M}^+ \vDash T_\varphi$ iff $\mathfrak{M}^+$ omits all types in C.

For every first-order formula $\vartheta \in T_\varphi$, we define the type

$$\mathfrak{p}_\vartheta := \{\neg\vartheta\}\,.$$

Every formula $\vartheta \in T_\varphi \setminus \mathrm{FO}$ is of the form $\vartheta = \forall\bar{x} \bigvee_{i<\lambda} \psi_i$. For these formulae, we set

$$\mathfrak{p}_\vartheta := \{\, \neg\psi_i \mid i < \lambda \,\}\,.$$

By construction, a structure satisfies $\vartheta \in T_\varphi$ if, and only if, it omits $\mathfrak{p}_\vartheta$. Consequently, we can set $C := \{\, \mathfrak{p}_\vartheta \mid \vartheta \in T_\varphi \,\}$. □

2. *Type spaces*

In this section we investigate the analogy between type spaces and spectra. We start by defining a topology on the set of type $S(L)$ that is analogous to the Stone topology of a spectrum.

Definition 2.1. The *type space* of a logic L is the topological space $\mathfrak{S}(L)$ with universe $S(L)$ where the basic closed sets are of the form

$$\langle \varphi_0 \rangle_L \cup \dots \cup \langle \varphi_{n-1} \rangle_L \,, \quad \text{for } n < \omega \text{ and } \varphi_0, \dots, \varphi_{n-1} \in L \,.$$

If L is closed under disjunctions, the closed sets can be written in the simpler form $\langle \Phi \rangle_L$, for $\Phi \subseteq L$.

Lemma 2.2. *If L is closed under disjunctions, every nonempty closed set of $\mathfrak{S}(L)$ is of the form $\langle \Phi \rangle_L$, for $\Phi \subseteq L$.*

Proof. Let $\mathcal{C} := \{\varnothing\} \cup \{ \langle \Phi \rangle_L \mid \Phi \subseteq L \}$. Since $\langle \Phi \rangle_L = \bigcap_{\varphi \in \Phi} \langle \varphi \rangle_L$, every set of $\mathcal{C}$ is closed in $\mathfrak{S}(L)$. To prove the converse, it is sufficient to show that $\mathcal{C}$ forms a topology. Since $\bigcap_i \langle \Phi_i \rangle_L = \langle \bigcup_i \Phi_i \rangle_L$, $\mathcal{C}$ is closed under arbitrary intersections. Furthermore, note that $\varnothing \in \mathcal{C}$ and $S(L) = \langle \varnothing \rangle_L \in \mathcal{C}$.

Hence, it remains to show that $\mathcal{C}$ is closed under finite unions. We claim that

$$\langle \Phi \rangle_L \cup \langle \Psi \rangle_L = \langle \{ \varphi \vee \psi \mid \varphi \in \Phi,\ \psi \in \Psi \} \rangle_L \,.$$

For the non-trivial inclusion, let $\mathfrak{p} \in \langle \{ \varphi \vee \psi \mid \varphi \in \Phi,\ \psi \in \Psi \} \rangle_L$. We have to show that $\mathfrak{p} \in \langle \Phi \rangle_L \cup \langle \Psi \rangle_L$. If $\mathfrak{p} \in \langle \Psi \rangle_L$, we are done. Hence, suppose there is some formula $\psi \in \Psi \setminus \mathfrak{p}$. For every $\varphi \in \Phi$, we have $\varphi \vee \psi \in \mathfrak{p}$. Since $\psi \notin \mathfrak{p}$ it follows as in the proof of Lemma 1.4 that $\varphi \in \mathfrak{p}$. Therefore, $\Phi \subseteq \mathfrak{p}$. ☐

For first-order logic, we have seen in Corollary 1.6 that types and ultrafilters coincide. Since the definitions of the respective topologies are also the same, it follows that the type space of a first-order logic is just its spectrum.

Theorem 2.3. *We have*

$$\mathfrak{S}(\mathrm{FO}[\Sigma, X]) = \operatorname{spec}(\langle \mathrm{FO}[\Sigma, X], \vDash \rangle)\,.$$

In particular, the type space $\mathfrak{S}(\mathrm{FO}[\Sigma, X])$ *is a Stone space.*

Proof. By Corollary 1.6 both spaces have the same universe and, according to Lemma 2.2, the closed sets are also the same. □

Example. Let $T := \operatorname{Th}(\mathfrak{C})$ where $\mathfrak{C} := \langle 2^\omega, (P_n)_{n<\omega} \rangle$ and

$$P_n := \{ \alpha \in 2^\omega \mid \alpha(n) = 1 \}\,.$$

Then $S^1(T) = \{ \mathfrak{p}_\alpha \mid \alpha \in 2^\omega \}$ where

$$\begin{aligned} \mathfrak{p}_\alpha &\vDash P_n x && \text{for } n \in \alpha^{-1}(1)\,, \\ \mathfrak{p}_\alpha &\vDash \neg P_n x && \text{for } n \in \alpha^{-1}(0)\,. \end{aligned}$$

The basic closed sets of $\mathfrak{S}^1(T)$ are of the form

$$\langle P_{i_0} x \wedge \dots \wedge P_{i_k} x \wedge \neg P_{j_0} x \wedge \dots \wedge \neg P_{j_m} x \rangle\,.$$

Since these sets are clopen it follows that the open sets are of the form

$$O_W := \{ \mathfrak{p}_\alpha \mid \text{ there is some } w \prec \alpha \text{ with } w \in W \}$$

with $W \subseteq 2^{<\omega}$. Consequently, the type space $\mathfrak{S}^1(T)$ is homeomorphic to the Cantor discontinuum.

For logics different from first-order logic, the type spaces usually are not Stone spaces.

Definition 2.4. A topological space $\mathfrak{X}$ is a T_0*-space* if, for every pair $x, y \in X$ of distinct points, there exists a closed set C such that

$$x \in C \text{ and } y \in X \setminus C\,, \quad \text{or} \quad x \in X \setminus C \text{ and } y \in C\,.$$

Lemma 2.5. *Let* L *be a logic. The type space* $\mathfrak{S}(L)$ *is a* T_0*-space.*

Proof. If $\mathfrak{p}, \mathfrak{q} \in S(L)$ are distinct types, there exists a formula φ such that $\varphi \in \mathfrak{p} \setminus \mathfrak{q}$ or $\varphi \in \mathfrak{q} \setminus \mathfrak{p}$. Consequently, $\mathfrak{p} \in \langle\varphi\rangle$ and $\mathfrak{q} \in S(L) \setminus \langle\varphi\rangle$, or $\mathfrak{p} \in S(L) \setminus \langle\varphi\rangle$ and $\mathfrak{q} \in \langle\varphi\rangle$. □

As an application of the Stone topology of the type space, consider the question of whether a first-order theory T has a model that realises all types in a given set X but no other ones. This is not possible for every set of types. The next lemma provides a first, topological condition X has to satisfy.

Lemma 2.6. *Let T be a complete first-order theory, $\mathfrak{M}$ a model of T, $U \subseteq M$, $\bar{s}$ a sequence of sorts, and let X be the set of all $\bar{s}$-types over U that are realised in $\mathfrak{M}$. Then X is dense in $\mathfrak{S}^{\bar{s}}(U)$.*

Proof. For a contradiction, suppose that there exists a type $\mathfrak{p} \in S^{\bar{s}}(U)$ with $\mathfrak{p} \notin \mathrm{cl}(X)$. Then we can find some formula $\varphi(\bar{x})$ over U with $\mathfrak{p} \in \langle\varphi\rangle$ and $\langle\varphi\rangle \cap X = \varnothing$. It follows that $\mathfrak{M} \vDash \neg\varphi(\bar{a})$, for all $\bar{a} \in M^{\bar{s}}$. Hence, $\mathfrak{M} \vDash \forall\bar{x}\neg\varphi(\bar{x})$ which implies that $\forall\bar{x}\neg\varphi(\bar{x}) \in T \subseteq \mathfrak{p}$. Consequently, $\varphi(\bar{x}) \wedge \forall\bar{x}\neg\varphi(\bar{x}) \in \mathfrak{p}$ and $\mathfrak{p}$ is inconsistent. Contradiction. □

Example. Let $\mathfrak{N} := \langle\omega, s, 0\rangle$ where $s : n \mapsto n+1$ is the successor function. We have seen on page 528 that the types of $\mathrm{Th}(\mathfrak{N})$ are $\mathfrak{p}_n := \mathrm{tp}(n)$, for $n < \omega$, and the type $\mathfrak{p}_\infty$ of an infinite element. The set of realised types is $X := \{\, \mathfrak{p}_n \mid n < \omega \,\}$, while $\mathfrak{p}_\infty$ is not realised. Note that a set $C \subseteq S(\varnothing)$ with $\mathfrak{p}_\infty \notin C$ is closed if, and only if, it is finite. Hence, $\mathfrak{p}_\infty$ is an accumulation point of X and X is dense in $\mathfrak{S}^1(\varnothing)$.

For most logics, the type space is not a spectrum. But, for a boolean closed logic L, we can at least prove the existence of an embedding $\mathfrak{S}(L) \to \mathrm{spec}(\mathfrak{Lb}(L))$. It turns out that this map is a homeomorphism if, and only if, the type space is compact.

Lemma 2.7. *Let L be a boolean closed logic. The type space $\mathfrak{S}(L)$ is compact if, and only if, every ultrafilter of $\langle L, \vDash\rangle$ is a complete type.*

Proof. ($\Leftarrow$) If every ultrafilter is a type, then $S(L) = \mathrm{spec}(\langle L, \vDash\rangle)$. Since the topologies of both spaces also coincide, they are homeomorphic.

Consequently, the compactness of $\operatorname{spec}(\langle L, \vDash\rangle)$ implies the compactness of $\mathfrak{S}(L)$.

($\Rightarrow$) Let $\mathfrak{u}$ be an ultrafilter of $\langle L, \vDash\rangle$. First, we show that $\mathfrak{u}$ is satisfiable. For a contradiction, suppose otherwise. Then

$$\varnothing = \langle \mathfrak{u}\rangle_L = \bigcap_{\varphi\in\mathfrak{u}} \langle\varphi\rangle_L .$$

Since $\mathfrak{S}(L)$ is compact, there is a finite subset $\Phi_0 \subseteq \mathfrak{u}$ such that

$$\varnothing = \bigcap_{\varphi\in\Phi_0} \langle\varphi\rangle_L .$$

Hence, $\bigwedge \Phi_0 \equiv \bot$ and $\Phi_0 \subseteq \mathfrak{u}$ implies $\bot \in \mathfrak{u}$. A contradiction.

Consequently, there is some model $\mathfrak{J} \vDash \mathfrak{u}$. Since L is boolean closed, it follows that $\operatorname{Th}_L(\mathfrak{J}) = \mathfrak{u}$. Therefore, $\mathfrak{u}$ is a complete type. □

Lemma 2.8. *Let L be a boolean closed logic.*

(a) *The function*

$$h : \mathfrak{S}(L) \to \operatorname{spec}(\mathfrak{Lb}(L)) : \mathfrak{p} \mapsto \mathfrak{p}/\equiv$$

is continuous and injective.

(b) *h is a homeomorphism if, and only if, $\mathfrak{S}(L)$ is compact.*

Proof. (a) First, note that, according to Lemma 1.4, for every $\mathfrak{p} \in S(L)$, $h(\mathfrak{p}) = \mathfrak{p}/\equiv$ is indeed an ultrafilter of $\mathfrak{Lb}(L)$.

For injectivity, consider types $\mathfrak{p} \neq \mathfrak{q}$. By symmetry, we may assume that there is some formula $\varphi \in \mathfrak{p} \setminus \mathfrak{q}$. If $h(\mathfrak{p}) = h(\mathfrak{q})$ then

$$[\varphi]_\equiv \in \mathfrak{p}/\equiv\, = h(\mathfrak{p}) = h(\mathfrak{q}) = \mathfrak{q}/\equiv$$

would imply that $\varphi \in \mathfrak{q}$. A contradiction.

To show that h is continuous, let $\Phi \subseteq \mathfrak{Lb}(L)$. Then

$$\begin{aligned} h^{-1}[\langle\Phi\rangle_{\mathfrak{Lb}(L)}] &= \{\, \mathfrak{p} \in S(L) \mid \Phi \subseteq \mathfrak{p}/\equiv \,\} \\ &= \{\, \mathfrak{p} \in S(L) \mid \bigcup \Phi \subseteq \mathfrak{p} \,\} = \langle \bigcup \Phi\rangle_L \end{aligned}$$

is closed.

(b) ($\Rightarrow$) If h is a homeomorphism, then $\mathfrak{S}(L) \cong \operatorname{spec}(\mathfrak{Lb}(L))$ is a Stone space and, hence, compact.

($\Leftarrow$) By (a), it remains to show that h is closed and surjective. For surjectivity, fix an ultrafilter $\mathfrak{u} \in \operatorname{spec}(\mathfrak{Lb}(L))$. Then $\bigcup \mathfrak{u}$ is an ultrafilter of $\langle L, \vDash \rangle$. Hence, Lemma 2.7 implies that $\bigcup \mathfrak{u} \in S(L)$. Consequently,

$$h(\bigcup \mathfrak{u}) = (\bigcup \mathfrak{u})/\equiv \; = \mathfrak{u},$$

as desired.

It remains to prove that h is closed. By Lemma B5.2.3, it is sufficient to show that $h[\langle \Phi \rangle_L]$ is closed, for every $\Phi \subseteq L$. For $\Phi \subseteq L$, it follows that

$$\begin{aligned} h[\langle \Phi \rangle_L] &= \{ \mathfrak{p}/\equiv \mid \mathfrak{p} \in S(L),\ \Phi \subseteq \mathfrak{p} \} \\ &= \big\{ \mathfrak{p}/\equiv \;\big|\; \mathfrak{p} \in S(L),\ \Phi/\equiv \; \subseteq \mathfrak{p}/\equiv \big\} \\ &= \langle \Phi/\equiv \rangle_{\mathfrak{Lb}(L)} \cap \operatorname{rng} h \\ &= \langle \Phi/\equiv \rangle_{\mathfrak{Lb}(L)} \end{aligned}$$

is closed. □

Corollary 2.9. *Let L be a boolean closed logic. The following conditions are equivalent:*

(1) $\mathfrak{S}(L)$ *is compact.*

(2) $\mathfrak{S}(L) \cong \operatorname{spec}(\mathfrak{Lb}(L))$.

(3) *Every ultrafilter of* $\langle L, \vDash \rangle$ *is a complete type.*

Many results of Section B5.6 on spectra generalise to type spaces. In particular, the type space operation $L \mapsto \mathfrak{S}(L)$ is a functor form the category of logics to the category of topological spaces.

Definition 2.10. Let $\mu := \langle \alpha, \beta \rangle : L_0 \to L_1$ be a morphism of logics. We define a function $\mathfrak{S}(\mu)$ by setting

$$\mathfrak{S}(\mu)(\mathfrak{p}) := \alpha^{-1}[\mathfrak{p}], \quad \text{for } \mathfrak{p} \in S(L_1).$$

Example. For the inclusion morphism $i : L|_\Phi \to L$ and the localisation morphism $\lambda : L \to L/\Phi$ from Lemma c1.6.14, we obtain

$$\mathfrak{S}(i)(\mathfrak{p}) = \mathfrak{p} \cap \Phi \quad \text{and} \quad \mathfrak{S}(\lambda)(\mathfrak{p}) = \mathfrak{p}\,.$$

Proposition 2.11. *Let $\mu := \langle \alpha, \beta \rangle : \langle L_0, \mathcal{K}_0, \vDash \rangle \to \langle L_1, \mathcal{K}_1, \vDash \rangle$ be a morphism of logics.*

(a) *$\mathfrak{S}(\mu)$ is the unique function that makes the following diagram commute:*

$$\begin{array}{ccc} \mathcal{K}_1 & \xrightarrow{\ \beta\ } & \mathcal{K}_0 \\ {\scriptstyle \mathrm{Th}_{L_1}}\downarrow & & \downarrow{\scriptstyle \mathrm{Th}_{L_0}} \\ \mathfrak{S}(L_1) & \xrightarrow[\ \mathfrak{S}(\mu)\ ]{} & \mathfrak{S}(L_0) \end{array}$$

(b) *$\mathfrak{S}(\mu) : \mathfrak{S}(L_1) \to \mathfrak{S}(L_0)$ is continuous.*

(c) *If μ is an embedding then $\mathfrak{S}(\mu)$ is surjective.*

(d) *If α is surjective then $\mathfrak{S}(\mu)$ is injective.*

(e) *If α is surjective and* $\operatorname{rng} \beta = \mathrm{Mod}_{L_0}(\Phi)$, *for some $\Phi \subseteq L_0$, then $\mathfrak{S}(\mu)$ is closed and injective.*

(f) *If $\mathfrak{S}(\mu)$ is surjective, then $\mathfrak{Lb}(\mu) : \mathfrak{Lb}(L_0) \to \mathfrak{Lb}(L_1)$ is injective.*

Proof. (a) We have seen in Lemma c1.5.12 (c) that $\mathfrak{S}(\mu) \circ \mathrm{Th}_{L_1} = \mathrm{Th}_{L_0} \circ \beta$. In particular, $\operatorname{rng} \mathfrak{S}(\mu) \subseteq \operatorname{rng} \mathrm{Th}_{L_0} = S(L_0)$ and the above diagram commutes. For uniqueness, note that $\mathrm{Th}_{L_1} : \mathcal{K}_1 \to S(L_1)$ is surjective. Hence, for every function f making the above diagram commute,

$$\mathfrak{S}(\mu) \circ \mathrm{Th}_{L_1} = \mathrm{Th}_{L_0} \circ \beta = f \circ \mathrm{Th}_{L_1} \quad \text{implies} \quad \mathfrak{S}(\mu) = f\,,$$

by Lemma a2.1.10.

(b) For every $\varphi \in L_0$, we have

$$\begin{aligned} \mathfrak{p} \in \mathfrak{S}(\mu)^{-1}[\langle\varphi\rangle_{L_0}] \quad &\text{iff} \quad \mathfrak{S}(\mu)(\mathfrak{p}) = \alpha^{-1}[\mathfrak{p}] \in \langle\varphi\rangle_{L_0} \\ &\text{iff} \quad \varphi \in \alpha^{-1}[\mathfrak{p}] \\ &\text{iff} \quad \alpha(\varphi) \in \mathfrak{p} \qquad \text{iff} \quad \mathfrak{p} \in \langle\alpha(\varphi)\rangle_{L_1}. \end{aligned}$$

Hence, $\mathfrak{S}(\mu)^{-1}[\langle\varphi\rangle_{L_0}] = \langle\alpha(\varphi)\rangle_{L_1}$. The claim follows by Lemma B5.2.3 since the sets $\langle\varphi\rangle_{L_0}$, for $\varphi \in L_0$, form a closed subbase of the topology of $\mathfrak{S}(L_0)$.

(c) Since β and Th_{L_0} are surjective, so is $\mathrm{Th}_{L_0} \circ \beta = \mathfrak{S}(\mu) \circ \mathrm{Th}_{L_1}$. Consequently, $\mathfrak{S}(\mu)$ is also surjective.

(d) Suppose that α is surjective and let $\mathfrak{p}, \mathfrak{q} \in S(L_1)$ be types such that $\mathfrak{S}(\mu)(\mathfrak{p}) = \mathfrak{S}(\mu)(\mathfrak{q})$. Then $\alpha^{-1}[\mathfrak{p}] = \alpha^{-1}[\mathfrak{q}]$ implies, by Lemma A2.1.10, that

$$\mathfrak{p} = \alpha[\alpha^{-1}[\mathfrak{p}]] = \alpha[\alpha^{-1}[\mathfrak{q}]] = \mathfrak{q}.$$

(e) We have already seen in (d) that $\mathfrak{S}(\mu)$ is injective. To show that it is closed, it is sufficient, by Lemma B5.2.3, to prove that $\mathfrak{S}(\mu)[\langle\varphi\rangle_{L_1}]$ is closed, for every $\varphi \in L_1$. We claim that

$$\mathfrak{S}(\mu)[\langle\varphi\rangle_{L_1}] = \langle\Phi \cup \alpha^{-1}(\varphi)\rangle_{L_0}.$$

($\subseteq$) Let $\mathfrak{p} \in \langle\varphi\rangle_{L_1}$ and fix an L_1-interpretation $\mathfrak{J}$ with $\mathrm{Th}_{L_1}(\mathfrak{J}) = \mathfrak{p}$. Then $\beta(\mathfrak{J}) \in \operatorname{rng}\beta = \mathrm{Mod}_{L_0}(\Phi)$ implies $\Phi \subseteq \mathrm{Th}_{L_0}(\beta(\mathfrak{J})) = \mathfrak{S}(\mu)(\mathfrak{p})$. Furthermore, $\varphi \in \mathfrak{p}$ implies $\alpha^{-1}(\varphi) \subseteq \alpha^{-1}[\mathfrak{p}] = \mathfrak{S}(\mu)(\mathfrak{p})$. Consequently, $\mathfrak{S}(\mu)(\mathfrak{p}) \in \langle\Phi \cup \alpha^{-1}(\varphi)\rangle_{L_0}$.

($\supseteq$) Let $\mathfrak{p} \in \langle\Phi \cup \alpha^{-1}(\varphi)\rangle_{L_0}$ and let $\mathfrak{J}_0$ be an L_0-interpretation with $\mathrm{Th}_{L_0}(\mathfrak{J}_0) = \mathfrak{p}$. Then $\mathfrak{J}_0 \vDash \Phi$ and $\operatorname{rng}\beta = \mathrm{Mod}_{L_0}(\Phi)$ implies that there is some L_1-interpretation $\mathfrak{J}$ with $\beta(\mathfrak{J}) = \mathfrak{J}_0$. Set $\mathfrak{q} := \mathrm{Th}_{L_1}(\mathfrak{J})$. Since α is surjective, we have

$$\begin{aligned} \alpha^{-1}(\varphi) \subseteq \mathfrak{p} \quad &\Rightarrow \quad \beta(\mathfrak{J}) \vDash \alpha^{-1}(\varphi) \\ &\Rightarrow \quad \mathfrak{J} \vDash \alpha[\alpha^{-1}(\varphi)] = \{\varphi\} \quad \Rightarrow \quad \varphi \in \mathfrak{q}. \end{aligned}$$

Hence, $\mathfrak{q} \in \langle\varphi\rangle_{L_1}$ and $\mathfrak{S}(\mu)(\mathfrak{q}) = \mathfrak{p}$.

(f) Let $\varphi, \psi \in L_0$ be formulae with $\alpha(\varphi) \equiv_{L_1} \alpha(\psi)$. We claim that $\varphi \equiv_{L_0} \psi$. By symmetry, it is sufficient to show that $\varphi \vDash \psi$.

Let $\mathfrak{J}$ be an L_0-interpretation with $\mathfrak{J} \vDash \varphi$. Since $\mathfrak{S}(\mu)$ is surjective, there is some type $\mathfrak{p} \in S(L_1)$ with $\mathfrak{S}(\mu)(\mathfrak{p}) = \mathrm{Th}_{L_0}(\mathfrak{J})$. Consequently,

$$\varphi \in \mathrm{Th}_{L_0}(\mathfrak{J}) = \mathfrak{S}(\mu)(\mathfrak{p}) = \alpha^{-1}[\mathfrak{p}] \quad \text{implies} \quad \alpha(\varphi) \in \mathfrak{p}\,.$$

Since $\alpha(\psi) \equiv_{L_1} \alpha(\varphi)$, it follows that $\alpha(\psi) \in \mathfrak{p}$. Hence, $\psi \in \alpha^{-1}(\mathfrak{p}) = \mathrm{Th}_{L_0}(\mathfrak{J})$ and $\mathfrak{J} \vDash \psi$. □

Corollary 2.12. *$\mathfrak{S}$ is a contravariant functor from $\mathfrak{Logic}$ to $\mathfrak{Top}_0$, the category of all T_0-spaces.*

Corollary 2.13. *Let $\mu : L_0 \to L_1$ be a morphism of logics.*

(a) *If μ is an embedding then $\mathfrak{S}(\mu)$ is a continuous surjection.*

(b) *If μ is an epimorphism then $\mathfrak{S}(\mu)$ is a continuous injection.*

(c) *If μ is an isomorphism then $\mathfrak{S}(\mu)$ is a homeomorphism.*

We can strengthen statement (c) of this corollary as follows.

Corollary 2.14. *Let $\mu = \langle\alpha, \beta\rangle : L_0 \to L_1$ be a morphism of logics where α and β are surjective. Then $\mathfrak{S}(\mu) : \mathfrak{S}(L_1) \to \mathfrak{S}(L_0)$ is a homeomorphism.*

Proof. As μ is an embedding, Corollary 2.13 (a) implies that $\mathfrak{S}(\mu)$ is continuous and surjective. Furthermore, $\mathrm{rng}\,\beta = \mathrm{Mod}_{L_0}(\varnothing)$. Therefore, we can use Proposition 2.11 (e) to show that $\mathfrak{S}(\mu)$ is closed and injective. □

Corollary 2.15. *Let L be a logic, $\Phi \subseteq L$, and $\lambda : L \to L/\Phi$ the localisation morphism. The function*

$$\mathfrak{S}(\lambda) : \mathfrak{S}(L/\Phi) \to \mathfrak{S}(L) : \mathfrak{p} \mapsto \mathfrak{p}$$

is continuous, closed, and injective.

Proof. Note that $\lambda = \langle \mathrm{id}, j \rangle$ where $j : \mathrm{Mod}_L(\Phi) \to \mathrm{Mod}_L(\varnothing)$ is the inclusion map. Since $\mathrm{rng}\, j = \mathrm{Mod}_L(\Phi)$, the claim follows by Proposition 2.11 (e). $\square$

Example. In analogy to boolean logic, we define the *Lindenbaum quotient* $\mathcal{Q}(L)$ of a logic L by

$$\mathcal{Q}(L) := \langle \mathfrak{Lb}(L), \mathfrak{S}(L), \vDash \rangle$$

where, for $\mathfrak{p} \in S(L)$ and $\varphi \in L$,

$$\mathfrak{p} \vDash [\varphi]_{\equiv} \quad :\text{iff} \quad \varphi \in \mathfrak{p}\,.$$

We can turn $\mathcal{Q}$ into a functor $\mathcal{Q} : \mathfrak{Logic} \to \mathfrak{Logic}$ by setting, for a morphism $\mu : L_0 \to L_1$,

$$\mathcal{Q}(\mu) := \langle \mathfrak{Lb}(\mu), \mathfrak{S}(\mu) \rangle : \mathcal{Q}(L_0) \to \mathcal{Q}(L_1)\,.$$

The functor $\mathcal{Q}$ is idempotent in the sense that there exists a natural isomorphism $\eta : \mathcal{Q} \circ \mathcal{Q} \to \mathcal{Q}$. This natural isomorphism is defined as follows. For $\mathfrak{p} \in S(L)$, we have

$$\mathrm{Th}_{\mathcal{Q}(L)}(\mathfrak{p}) = \{\, [\varphi]_{\equiv} \mid \varphi \in \mathfrak{p} \,\} = \mathfrak{p}/{\equiv}\,.$$

Hence,

$$\mathfrak{S}(\mathcal{Q}(L)) = \{\, \mathrm{Th}_{\mathcal{Q}(L)}(\mathfrak{p}) \mid \mathfrak{p} \in S(L) \,\} = \{\, \mathfrak{p}/{\equiv} \mid \mathfrak{p} \in S(L) \,\}\,.$$

Since $\mathfrak{p}/{\equiv} = \mathfrak{q}/{\equiv}$ implies $\mathfrak{p} = \mathfrak{q}$, it follows that the function

$$\beta : \mathfrak{S}(L) \to \mathfrak{S}(\mathcal{Q}(L)) : \mathfrak{p} \mapsto \mathfrak{p}/{\equiv}$$

is a homeomorphism. Furthermore, since $\big[[\varphi]_{\equiv}\big]_{\equiv} = \{[\varphi]_{\equiv}\}$, the map

$$\alpha : \mathfrak{Lb}(\mathcal{Q}(L)) \to \mathfrak{Lb}(L) : [[\varphi]_{\equiv}]_{\equiv} \mapsto [\varphi]_{\equiv}$$

is a well-defined isomorphism of partial orders. Consequently, we obtain an isomorphism of logics

$$\eta_L := \langle \alpha, \beta \rangle : \mathcal{Q}(\mathcal{Q}(L)) \to \mathcal{Q}(L)\,.$$

Since, for every morphism $\mu : L_0 \to L_1$, we have

$$\eta_{L_1} \circ \mathcal{Q}(\mathcal{Q}(\mu)) = \mathcal{Q}(\mu) \circ \eta_{L_0}\,,$$

it follows that $(\eta_L)_L$ is a natural isomorphism.

For boolean closed logics where the type space is compact and, hence, homeomorphic to the spectrum of the Lindenbaum algebra, we can strengthen Corollary 2.13 (a) as follows.

Lemma 2.16. *Let L_0 and L_1 be boolean closed logics where $\mathfrak{S}(L_1)$ is compact. If $\mu : L_0 \to L_1$ is an embedding, $\mathfrak{S}(\mu) : \mathfrak{S}(L_1) \to \mathfrak{S}(L_0)$ is continuous, closed, and surjective.*

Proof. We have already seen in Corollary 2.13 that $\mathfrak{S}(\mu)$ is continuous and surjective. Hence, it remains to prove that it is closed. Note that it follows by Lemma b5.3.10 that $\mathfrak{S}(L_0) = \mathfrak{S}(\mu)[S(L_1)]$ is also compact. By Lemma 2.8, there exist homeomorphisms

$$h_i : \mathfrak{S}(L_i) \to \operatorname{spec}(\mathfrak{Lb}(L_i)) : \mathfrak{p} \mapsto \mathfrak{p}/{\equiv}\,, \quad \text{for } i \in [2]\,.$$

Furthermore, we have seen in Lemma c1.6.10 that $\mathfrak{Lb}(\mu)$ is injective. Hence, Lemma b5.6.7 implies that the function $g := \operatorname{spec}(\mathfrak{Lb}(\mu))$ is continuous, closed, and surjective.

$$\begin{array}{ccc}
\mathfrak{S}(L_1) & \xrightarrow{\ \mathfrak{S}(\mu)\ } & \mathfrak{S}(L_0) \\
\downarrow{\scriptstyle h_1} & & \downarrow{\scriptstyle h_2} \\
\operatorname{spec}(\mathfrak{Lb}(L_1)) & \xrightarrow[\ g\ ]{} & \operatorname{spec}(\mathfrak{Lb}(L_0))
\end{array}$$

Since $\mathfrak{S}(\mu) = h_0^{-1} \circ g \circ h_1$, it follows that $\mathfrak{S}(\mu)$ is closed. □

Lemma 2.17. *Let L be a logic and $\Phi \subseteq L$. If $\mathfrak{S}(L)$ is compact, then so are $\mathfrak{S}(L|_\Phi)$ and $\mathfrak{S}(L/\Phi)$.*

Proof. Let $\lambda : L \to L/\Phi$ and $i : L|_\Phi \to L$ be the canonical morphisms. We have seen in Corollary 2.13 that $\mathfrak{S}(i) : \mathfrak{S}(L) \to \mathfrak{S}(L|_\Phi)$ is continuous and surjective. Since $\mathfrak{S}(L)$ is compact, it follows by Lemma B5.3.10 that $\mathfrak{S}(i)[S(L)] = \mathfrak{S}(L|_\Phi)$ is also compact.

By Corollary 2.15, $\mathfrak{S}(\lambda) : \mathfrak{S}(L/\Phi) \to \mathfrak{S}(L)$ is continuous, closed, and injective. Consequently, $\mathfrak{S}(L/\Phi)$ is homeomorphic to a closed subset $\operatorname{rng} \mathfrak{S}(\lambda) \subseteq \mathfrak{S}(L)$ of $\mathfrak{S}(L)$. By Lemma B5.3.9, it follows that $\mathfrak{S}(L/\Phi)$ is compact. □

As a consequence, we obtain the following generalisation of Theorem 2.3.

Theorem 2.18. *For all first-order theories $T \subseteq \mathrm{FO}^0[\Sigma]$,*

$$\mathfrak{S}^{\bar{s}}(T) \cong \operatorname{spec}(\mathfrak{Lb}(\mathrm{FO}^{\bar{s}}[\Sigma]/T))$$

is a Stone space.

Proof. By Lemma 2.17, $\mathfrak{S}^{\bar{s}}(T) = \mathfrak{S}(\mathrm{FO}^{\bar{s}}[\Sigma]/T)$ is compact. Hence, the claim follows by Lemma 2.8 (b). □

For algebraic logics L, every map $\mu : \Sigma \to \Gamma$ between signatures gives rise to a morphism $L[\mu] : L[\Sigma] \to L[\Gamma]$ and a corresponding continuous map $\mathfrak{S}(L[\mu]) : \mathfrak{S}(L[\Gamma]) \to \mathfrak{S}(L[\Sigma])$. In the lemma below, we take a closer look at such a map, where $\mu : \Sigma_U \to \Sigma_V$ corresponds to a renaming of parameters.

Definition 2.19. Let L be an algebraic logic. For a type $\mathfrak{p}$ over U and a function $f : U \to V$, we write

$$f(\mathfrak{p}) := \{\, \varphi(\bar{x}; f(\bar{a})) \mid \varphi(\bar{x}; \bar{a}) \in \mathfrak{p} \,\}.$$

Remark. Suppose that f is a strict L-map and let $\mathfrak{p} := \mathrm{tp}_L(\bar{a}/U)$ where $\bar{a}, U \subseteq \mathrm{dom}\, f$. Then

$$f(\mathfrak{p}) = \mathrm{tp}_L\big(f(\bar{a}) \,/\, f[U]\big)\,.$$

Lemma 2.20. *Suppose that L is an algebraic logic, let $\mathfrak{A}$ and $\mathfrak{B}$ be Σ-structures, and $U \subseteq A$. Every injective, strict L-map $h : U \to B$ induces a homeomorphism*

$$\mathfrak{S}^{\bar{s}}_L(U) \to \mathfrak{S}^{\bar{s}}_L(h[U]) : \mathfrak{p} \mapsto h(\mathfrak{p})\,.$$

Proof. Set $V := h[U]$. Let $\mu : \Sigma_U \to \Sigma_V$ be the morphism of signatures with $\mu \restriction \Sigma = \mathrm{id}_\Sigma$ and $\mu \restriction U = h$, and let

$$\langle \alpha, \beta \rangle := L^{\bar{s}}[\mu] : L^{\bar{s}}[\Sigma_U] \to L^{\bar{s}}[\Sigma_V]$$

be the corresponding morphism of logics. Since μ is bijective so are α and β and we have $L^{\bar{s}}[\mu^{-1}] = \langle \alpha^{-1}, \beta^{-1} \rangle$.

We claim that β induces a bijection $\mathrm{Mod}_L(T(V)) \to \mathrm{Mod}_L(T(U))$. Let $\mathfrak{M} \models T(V)$, $\varphi(\bar{x}) \in L[\Sigma, X]$, and $\bar{c} \subseteq U$. As h is a strict L-map, it follows that

$$\begin{aligned}
\beta(\mathfrak{M}) \models \varphi(\bar{c}) \quad &\text{iff} \quad \mathfrak{M} \models \alpha(\varphi(\bar{c})) = \varphi(h(\bar{c})) \\
&\text{iff} \quad \varphi(h(\bar{c})) \in T(V) \\
&\text{iff} \quad \mathfrak{B} \models \varphi(h(\bar{c})) \\
&\text{iff} \quad \mathfrak{A} \models \varphi(\bar{c}) \\
&\text{iff} \quad \varphi(\bar{c}) \in T(U)\,.
\end{aligned}$$

Similarly, it follows that $\beta^{-1}(\mathfrak{M}) \in \mathrm{Mod}(T(V))$, for every model $\mathfrak{M}$ of $T(U)$. Therefore, $\langle \alpha, \beta \rangle$ induces a morphism

$$\langle \alpha, \beta_0 \rangle : L^{\bar{s}}/T(U) \to L^{\bar{s}}/T(V)$$

where $\beta_0 = \beta \restriction \mathrm{Mod}_L(T(V))$ is bijective. As α is also bijective, it follows by Corollary 2.13 that the induced map

$$\mathfrak{S}^{\bar{s}}_L\langle \alpha, \beta_0 \rangle : \mathfrak{S}^{\bar{s}}_L(V) \to \mathfrak{S}^{\bar{s}}_L(U) : h(\mathfrak{p}) \mapsto \mathfrak{p}$$

is a homeomorphism. □

For first-order type spaces, we can say more on the dependence of a type space on the signature.

Proposition 2.21. *Let* $\Sigma_0 \subseteq \Sigma$ *be signatures,* $T \subseteq \mathrm{FO}^0[\Sigma]$ *a theory, and set* $T_0 := T \cap \mathrm{FO}^0[\Sigma_0]$.

(a) *For every* $\Delta \subseteq \mathrm{FO}^0[\Sigma_0]$, *we have*

$$\mathfrak{S}((\mathrm{FO}^0[\Sigma_0]/T_0)|_\Delta) = \mathfrak{S}((\mathrm{FO}^0[\Sigma]/T)|_\Delta) .$$

(b) *The function*

$$h : \mathfrak{S}(\mathrm{FO}^0[\Sigma]/T) \to \mathfrak{S}(\mathrm{FO}^0[\Sigma_0]/T_0) : \mathfrak{p} \mapsto \mathfrak{p} \cap \mathrm{FO}^0[\Sigma_0]$$

is continuous, closed, and surjective.

Proof. To simplify notation, set $L := \mathrm{FO}^0[\Sigma]$ and $L_0 := \mathrm{FO}^0[\Sigma_0]$.

(a) We start by showing that both type spaces have the same universe.

Let $\mathfrak{p} \in S((L/T)|_\Delta)$. Then there is some $\mathfrak{M} \in \mathrm{Mod}_L(T)$ with $\mathfrak{p} = \mathrm{Th}_\Delta(\mathfrak{M})$. Setting $\mathfrak{M}_0 := \mathfrak{M}|_{\Sigma_0}$ we obtain a model $\mathfrak{M}_0 \in \mathrm{Mod}_{L_0}(T_0)$ with $\mathfrak{p} = \mathrm{Th}_\Delta(\mathfrak{M}_0)$. It follows that $\mathfrak{p} \in S((L_0/T_0)|_\Delta)$.

Conversely, let $\mathfrak{p} \in S((L_0/T_0)|_\Delta)$. Then there is some model $\mathfrak{M}_0 \in \mathrm{Mod}_{L_0}(T_0)$ with $\mathfrak{p} = \mathrm{Th}_\Delta(\mathfrak{M}_0)$. We can use Corollary C2.5.9 to find a model $\mathfrak{M} \in \mathrm{Mod}_L(T)$ such that $\mathfrak{M}_0 \preceq \mathfrak{M}|_\Sigma$. It follows that $\mathfrak{p} = \mathrm{Th}_\Delta(\mathfrak{M})$. Hence, $\mathfrak{p} \in S((L/T)|_\Delta)$.

It remains to show that the two topologies coincide. For $\Phi \subseteq \Delta$, it follows by definition that

$$\begin{aligned} \langle \Phi \rangle_{(L_0/T_0)|_\Delta} &= \{ \mathfrak{p} \in S((L_0/T_0)|_\Delta) \mid \Phi \subseteq \mathfrak{p} \} \\ &= \{ \mathfrak{p} \in S((L/T)|_\Delta) \mid \Phi \subseteq \mathfrak{p} \} = \langle \Phi \rangle_{(L/T)|_\Delta} . \end{aligned}$$

(b) Consider the inclusion map $i : (L/T)|_{L_0} \to L/T$. By Lemma 2.16, the map

$$\mathfrak{S}(i) : \mathfrak{S}(L/T) \to \mathfrak{S}((L/T)|_{L_0}) : \mathfrak{p} \mapsto \mathfrak{p} \cap L$$

is continuous, closed, and surjective. Furthermore, we have seen in (a) that the identity map

$$\mathrm{id} : \mathfrak{S}((L/T)|_{L_0}) \to \mathfrak{S}(L_0/T_0)$$

is a homeomorphism. It follows that the composition $h = \mathrm{id} \circ \mathfrak{S}(i)$ is continuous, closed, and surjective. □

A special case of Proposition 2.21 (b) is worth singling out.

Corollary 2.22. *Let $T \subseteq \mathrm{FO}^{\circ}[\Sigma]$ be a first-order theory, $U \subseteq V$ sets of parameters, and*

$$i : \mathrm{FO}^{\bar{s}}[\Sigma_U]/T(U) \to \mathrm{FO}^{\bar{s}}[\Sigma_V]/T(V)$$

the inclusion morphism. The induced map

$$\mathfrak{S}(i) : \mathfrak{S}^{\bar{s}}(V) \to \mathfrak{S}^{\bar{s}}(U)$$

is continuous, closed, and surjective.

3. Retracts

For every fragment Δ of a logic L, we have seen above that the inclusion morphism $i : \Delta \to L$ induces a surjective, continuous map

$$\mathfrak{S}(i) : \mathfrak{S}(L) \to \mathfrak{S}(\Delta) : \mathfrak{p} \mapsto \mathfrak{p} \cap \Delta\,.$$

It follows that the type space of Δ is a quotient of the type space of L. In this section we take a closer look at the relationship between these two type spaces.

Definition 3.1. Let L be a logic, $L_0 \subseteq L$ a fragment, and $i : L_0 \to L$ the inclusion morphism.

(a) A morphism $r : L \to L_0$ is a *retraction* if $r \circ i = \mathrm{id}$.

(b) L_o is a *retract* of L if there exists a retraction $L \to L_o$.

The type space of a retract is homeomorphic to the type space of the full logic.

Lemma 3.2. *Let $r = \langle \alpha, \beta \rangle : L \to L_o$ be a retraction and $i : L_o \to L$ the inclusion morphism.*

(a) $\beta = \mathrm{id}$.

(b) $\alpha(\varphi) \equiv_L \varphi$, *for every* $\varphi \in L$.

(c) $\mathfrak{S}(r) = \mathfrak{S}(i)^{-1}$.

Proof. (a) Note that $i = \langle \iota, \mathrm{id} \rangle$, where $\iota : L_o \to L$ is the inclusion function. Hence, $r \circ i = \mathrm{id}$ implies that $\mathrm{id} \circ \beta = \mathrm{id}$.

(b) To show that $\alpha(\varphi) \equiv_L \varphi$, let $\mathfrak{J}$ be an L-interpretation. We have seen in (a) that $\beta(\mathfrak{J}) = \mathfrak{J}$. Since r is a morphism of logics, it follows that

$$\mathfrak{J} \vDash \varphi \quad \text{iff} \quad \mathfrak{J} \vDash \alpha(\varphi)\,.$$

(c) Note that $r \circ i = \mathrm{id}$ implies $\mathfrak{S}(i) \circ \mathfrak{S}(r) = \mathrm{id}$. Hence, it remains to show that $\mathfrak{S}(r) \circ \mathfrak{S}(i) = \mathrm{id}$. Consider $\mathfrak{p} \in S(L)$. By (b), it follows that

$$\varphi \in \mathfrak{p} \quad \text{iff} \quad \alpha(\varphi) \in \mathfrak{p}\,, \quad \text{for all } \varphi \in L\,.$$

Hence, $\mathfrak{p} = \alpha^{-1}[\mathfrak{p}] = \mathfrak{S}(i \circ r)(\mathfrak{p}) = (\mathfrak{S}(r) \circ \mathfrak{S}(i))(\mathfrak{p})$. ☐

Corollary 3.3. *Let $r : L \to L_o$ be a retraction and $i : L_o \to L$ the inclusion morphism.*

(a) $\mathfrak{S}(i) : \mathfrak{S}(L) \to \mathfrak{S}(L_o)$ *is a homeomorphism.*

(b) $\mathfrak{S}(r) : \mathfrak{S}(L_o) \to \mathfrak{S}(L)$ *is a homeomorphism.*

Proof. Both statements follow from Lemma 3.2 (c). ☐

Lemma 3.4. *Let L be a logic, $L_o \subseteq L$ a fragment, and $i = \langle \iota, \mathrm{id} \rangle : L_o \to L$ the inclusion morphism. The following statements are equivalent:*

(1) L_o *is a retract of* L.

(2) *For every formula* $\varphi \in L$*, there is a formula* $\varphi_0 \in L_0$ *such that* $\varphi \equiv_L \varphi_0$.
(3) *The function* $\mathfrak{Lb}(i) : \mathfrak{Lb}(L_0) \to \mathfrak{Lb}(L)$ *is an isomorphism.*

Proof. (1) ⇒ (2) follows immediately by Lemma 3.2 (b).

(2) ⇒ (3) We have seen in Lemma C1.6.10 that $\mathfrak{Lb}(i)$ is an embedding. Hence, it remains to show that it is surjective. Let $[\varphi]_\equiv \in \mathfrak{Lb}(L)$. By (2), there is some formula $\varphi_0 \in L_0$ with $\varphi_0 \equiv_L \varphi$. It follows that

$$\mathfrak{Lb}(i)([\varphi_0]_\equiv) = [\varphi_0]_\equiv = [\varphi]_\equiv .$$

(3) ⇒ (1) We define a function $\alpha : L \to L_0$ as follows. For $\varphi \in L_0$, we set $\alpha(\varphi) := \varphi$. For $\varphi \in L \setminus L_0$, we choose an arbitrary formula ψ such that $[\psi]_\equiv \in \mathfrak{Lb}(i)^{-1}([\varphi]_\equiv)$ and set $\alpha(\varphi) := \psi$. Note that, for every $\varphi \in L$,

$$\alpha(\varphi) \in \mathfrak{Lb}(i)^{-1}([\varphi]_\equiv)$$

implies that

$$[\varphi]_\equiv = \mathfrak{Lb}(i)([\alpha(\varphi)]_\equiv) = [\alpha(\varphi)]_\equiv .$$

Hence, $\alpha(\varphi) \equiv_L \varphi$, for all $\varphi \in L$. By definition of α, we further have

$$(\alpha \circ \iota)(\varphi) = \alpha(\varphi) = \varphi , \quad \text{for all } \varphi \in L_0 .$$

Hence, to show that $r := \langle \alpha, \mathrm{id} \rangle$ is a left inverse of i it remains to prove that r is a morphism of logics. Let $\varphi \in L$ be a formula and $\mathfrak{J}$ an L-interpretation. Since $\varphi \equiv_L \alpha(\varphi)$, we have

$$\mathfrak{J} \vDash \varphi \quad \text{iff} \quad \mathfrak{J} \vDash \alpha(\varphi) ,$$

as desired. □

Below we will present several results that assume Δ to be boolean closed. The following lemma can sometimes be used to replace this restriction by the requirement that Δ is closed under negation.

Lemma 3.5. *Let L be a boolean closed logic, $\Delta \subseteq L$ closed under negation, and let $i : \Delta \to L$ be the inclusion morphism. If every formula in L is equivalent to a finite boolean combination of formulae in Δ, then*

$$\mathfrak{S}(i) : \mathfrak{S}(L) \to \mathfrak{S}(\Delta)$$

is a homeomorphism.

Proof. By Corollary 2.13, $\mathfrak{S}(i)$ is continuous and surjective.

For injectivity, suppose that $\mathfrak{S}(i)(\mathfrak{p}) = \mathfrak{S}(i)(\mathfrak{q})$. Then $\mathfrak{p} \cap \Delta = \mathfrak{q} \cap \Delta$. Since every formula in L is equivalent to a boolean combination of formulae in Δ, it follows that $\mathfrak{p} = \mathfrak{q}$.

It remains to show that $\mathfrak{S}(i)$ is closed. By Lemma B5.2.3, it is sufficient to prove that $\mathfrak{S}(i)[\langle\varphi\rangle_L]$ is closed, for every $\varphi \in L$. Fix $\varphi \in L$. By assumption on Δ and L, there are sets $\Psi_0, \dots, \Psi_{n-1} \subseteq \Delta$ such that $\varphi \equiv_L \bigvee_{k<n} \bigwedge \Psi_k$. Since, trivially, $\Psi_k \equiv_L \Psi_k$, it follows by Lemma 3.9 that

$$\mathfrak{S}(i)[\langle\Psi_k\rangle_L] = \langle\Psi_k\rangle_\Delta .$$

Consequently,

$$\mathfrak{S}(i)[\langle\varphi\rangle_L] = \mathfrak{S}(i)\Big[\bigcup_{k<n}\langle\Psi_k\rangle_L\Big] = \bigcup_{k<n}\mathfrak{S}(i)\big[\langle\Psi_k\rangle_L\big] = \bigcup_{k<n}\langle\Psi_k\rangle_\Delta$$

is closed. □

Exercise 3.1. Show that the preceding lemma may fail if Δ is not closed under negation.

Corollary 3.6. *Let L be a boolean closed logic such that $\mathfrak{S}(L)$ is compact, let $\Delta_0 \subseteq \Delta \subseteq L$ be closed under negation, and let $i : \Delta_0 \to \Delta$ be the inclusion morphism. The induced map*

$$\mathfrak{S}(i) : \mathfrak{S}(\Delta) \to \mathfrak{S}(\Delta_0)$$

is continuous, closed, and surjective.

Proof. Let $\Delta_0^+ \subseteq \Delta^+ \subseteq L$ be the boolean closures of Δ_0 and Δ, and let $j_0 : \Delta_0 \to \Delta_0^+$, $j : \Delta \to \Delta^+$, and $i_+ : \Delta_0^+ \to \Delta^+$ be the corresponding inclusion morphisms. By Lemma 3.5, $\mathfrak{S}(j_0)$ and $\mathfrak{S}(j)$ are homeomorphisms. Hence,

$$j \circ i = i_+ \circ j_0 \quad \text{implies} \quad \mathfrak{S}(i) = \mathfrak{S}(j_0) \circ \mathfrak{S}(i_+) \circ \mathfrak{S}(j)^{-1}.$$

Since, by Lemma 2.16, the functions on the right-hand side are continuous, closed, and surjective, so is $\mathfrak{S}(i)$. □

In the remainder of this section we consider to which extend the reverse of Corollary 3.3 (a) holds: in which cases is $\mathfrak{S}(i)$ being a homeomorphism sufficient for Δ to have the same expressive power as L.

Lemma 3.7. *Let L_0 and L_1 be logics and $\mu : \mathfrak{S}(L_0) \to \mathfrak{S}(L_1)$ a homeomorphism. Then*

$$\mathfrak{p} \subseteq \mathfrak{q} \quad \text{iff} \quad \mu(\mathfrak{p}) \subseteq \mu(\mathfrak{q}), \quad \textit{for all } \mathfrak{p}, \mathfrak{q} \in S(L_0).$$

Proof. It is sufficient to prove that $\mathfrak{p} \subseteq \mathfrak{q}$ implies $\mu(\mathfrak{p}) \subseteq \mu(\mathfrak{q})$. Then we can prove the converse implication, by considering the homeomorphism μ^{-1}. Note that we have

$$\begin{aligned} &\mathfrak{p} \subseteq \mathfrak{q} &&\text{iff} \quad \text{for all } \Phi, \quad \mathfrak{p} \in \langle \Phi \rangle_{L_0} \Rightarrow \mathfrak{q} \in \langle \Phi \rangle_{L_0}, \\ \text{and} \quad &\mu(\mathfrak{p}) \subseteq \mu(\mathfrak{q}) &&\text{iff} \quad \text{for all } \Psi, \ \mu(\mathfrak{p}) \in \langle \Psi \rangle_{L_1} \Rightarrow \mu(\mathfrak{q}) \in \langle \Psi \rangle_{L_1}. \end{aligned}$$

Let us show that the condition

$$\mathfrak{p} \in \langle \Phi \rangle_{L_0} \Rightarrow \mathfrak{q} \in \langle \Phi \rangle_{L_0}, \quad \text{for all } \Phi \subseteq L_0$$

is equivalent to

$$\mathfrak{p} \in C \Rightarrow \mathfrak{q} \in C, \quad \text{for all closed } C \subseteq S(L_0).$$

Clearly, if the implication holds for all closed sets C, it in particular holds for closed sets of the form $\langle \Phi \rangle_{L_0}$. Hence, it is sufficient to prove the converse. Suppose that every set $\langle \Phi \rangle_{L_0}$ containing $\mathfrak{p}$ also contains $\mathfrak{q}$ and let C be a closed set with $\mathfrak{p} \in C$. By definition, there is a family $(\Psi_i)_{i \in I}$ of finite sets $\Psi_i \subseteq L_0$ such that

$$C = \bigcap_{i \in I} \bigcup_{\psi \in \Psi_i} \langle \psi \rangle_{L_0} .$$

Since $\mathfrak{p} \in C$, there are formulae $\psi_i \in \Psi_i$, for $i \in I$, such that $\mathfrak{p} \in \langle \psi_i \rangle_{L_0}$. By assumption, this implies that $\mathfrak{q} \in \langle \psi_i \rangle_{L_0}$. Hence,

$$\mathfrak{q} \in \bigcap_{i \in I} \langle \psi_i \rangle_{L_0} \subseteq \bigcap_{i \in I} \bigcup_{\psi \in \Psi_i} \langle \psi \rangle_{L_0} = C .$$

To prove the lemma, suppose that $\mathfrak{p} \subseteq \mathfrak{q}$. We have just seen that this implies that

$$\mathfrak{p} \in C \Rightarrow \mathfrak{q} \in C , \quad \text{for all closed } C \subseteq S(L_0) .$$

Hence,

$$\mu(\mathfrak{p}) \in \mu[C] \Rightarrow \mu(\mathfrak{q}) \in \mu[C] , \quad \text{for all closed } C \subseteq S(L_0) .$$

Since μ is a homeomorphism, it follows that

$$\mu(\mathfrak{p}) \in D \Rightarrow \mu(\mathfrak{q}) \in D , \quad \text{for all closed } D \subseteq S(L_1) .$$

As we have seen above, this implies that

$$\mathfrak{p} \in \langle \Psi \rangle_{L_1} \Rightarrow \mathfrak{q} \in \langle \Psi \rangle_{L_1} , \quad \text{for all } \Psi \subseteq L_1 .$$

Consequently, we have $\mu(\mathfrak{p}) \subseteq \mu(\mathfrak{q})$. □

Corollary 3.8. *Let L be a logic, $\Delta \subseteq L$, and $i : \Delta \to L$ the inclusion morphism. If $\mathfrak{S}(i) : \mathfrak{S}(L) \to \mathfrak{S}(\Delta)$ is a homeomorphism, then*

$$\mathfrak{p} \cap \Delta \vDash \mathfrak{p} , \quad \textit{for all } \mathfrak{p} \in S(L) .$$

Proof. Suppose that $\mathfrak{J} \vDash \mathfrak{p} \cap \Delta$. To show that $\mathfrak{J} \vDash \mathfrak{p}$, consider $\mathfrak{q} := \mathrm{Th}_L(\mathfrak{J})$. Since $\mathfrak{S}(i)$ is bijective we have

$$\begin{aligned} &\mathfrak{S}(i)^{-1}(\mathfrak{p} \cap \Delta) = \mathfrak{S}(i)^{-1}(\mathfrak{S}(i)(\mathfrak{p})) = \mathfrak{p} \\ \text{and} \quad &\mathfrak{S}(i)^{-1}(\mathfrak{q} \cap \Delta) = \mathfrak{S}(i)^{-1}(\mathfrak{S}(i)(\mathfrak{q})) = \mathfrak{q}\,. \end{aligned}$$

Hence, $\mathfrak{p} \cap \Delta \subseteq \mathrm{Th}_\Delta(\mathfrak{J}) = \mathfrak{q} \cap \Delta$ implies, by Lemma 3.7, that

$$\mathfrak{p} = \mathfrak{S}(i)^{-1}(\mathfrak{p} \cap \Delta) \subseteq \mathfrak{S}(i)^{-1}(\mathfrak{q} \cap \Delta) = \mathfrak{q} = \mathrm{Th}_L(\mathfrak{J})\,.$$

Consequently, $\mathfrak{J} \vDash \mathfrak{p}$. ☐

Below we will provide several characterisations of when Δ has the same expressive power as L. We start with a technical lemma containing a condition on when two sets Φ, Ψ of formulae are equivalent.

Lemma 3.9. *Let L be a logic, $\Delta \subseteq L$, and $i : \Delta \to L$ the inclusion morphism. If $\mathfrak{S}(i) : \mathfrak{S}(L) \to \mathfrak{S}(\Delta)$ is bijective, we have*

$$\Phi \equiv_L \Psi \quad \text{iff} \quad \mathfrak{S}(i)[\langle \Phi \rangle_L] = \langle \Psi \rangle_\Delta\,,$$

for all sets $\Phi \subseteq L$ and $\Psi \subseteq \Delta$.

Proof. ($\Rightarrow$) Suppose that $\Phi \equiv_L \Psi$. For ($\subseteq$), let $\mathfrak{p} \in \langle \Phi \rangle_L$. Then $\Phi \subseteq \mathfrak{p}$ implies

$$\Psi \subseteq \Phi^{\vDash} \cap \Delta \subseteq \mathfrak{p} \cap \Delta = \mathfrak{S}(i)(\mathfrak{p})\,.$$

Hence, $\mathfrak{S}(i)(\mathfrak{p}) \in \langle \Psi \rangle_\Delta$.

For ($\supseteq$), let $\mathfrak{p} \in \langle \Psi \rangle_\Delta$. Since $\mathfrak{S}(i)$ is surjective, there is some $\mathfrak{q} \in S(L)$ with $\mathfrak{S}(i)(\mathfrak{q}) = \mathfrak{p}$. Hence,

$$\Psi \subseteq \mathfrak{p} = \mathfrak{S}(i)(\mathfrak{q}) = \mathfrak{q} \cap \Delta \subseteq \mathfrak{q}\,.$$

Since $\Psi \vDash \Phi$, it follows that $\Phi \subseteq \mathfrak{q}$. Consequently, we have $\mathfrak{q} \in \langle \Phi \rangle_L$, which implies that $\mathfrak{p} = \mathfrak{S}(i)(\mathfrak{q}) \in \mathfrak{S}(i)[\langle \Phi \rangle_L]$.

($\Leftarrow$) We have to show that $\Phi \equiv_L \Psi$. First, suppose that $\mathfrak{J} \vDash \Phi$ and let $\mathfrak{p} := \mathrm{Th}_\Delta(\mathfrak{J})$. Then $\mathfrak{p} \in \langle \Phi \rangle_L$ implies that

$$\mathrm{Th}_\Delta(\mathfrak{J}) = \mathfrak{p} \cap \Delta = \mathfrak{S}(i)(\mathfrak{p}) \in \mathfrak{S}(i)[\langle \Phi \rangle_L] = \langle \Psi \rangle_\Delta .$$

Hence, $\mathfrak{J} \vDash \Psi$. Conversely, suppose that $\mathfrak{J} \vDash \Psi$ and let $\mathfrak{p} := \mathrm{Th}_L(\mathfrak{J})$. Then $\mathfrak{S}(i)(\mathfrak{p}) = \mathfrak{p} \cap \Delta \in \langle \Psi \rangle_\Delta$. Since $\mathfrak{S}(i)$ is injective, we have

$$\mathfrak{p} \in \mathfrak{S}(i)^{-1}(\mathfrak{p} \cap \Delta) \subseteq \mathfrak{S}(i)^{-1}[\langle \Psi \rangle_\Delta] = \langle \Phi \rangle_L$$

and, therefore, $\mathfrak{J} \vDash \Phi$. □

For fragments $\Delta \subseteq L$ that are closed under disjunctions, we obtain the following characterisation of when every L-formula is equivalent to a set of Δ-formulae.

Proposition 3.10. *Let L be a logic, $\Delta \subseteq L$, and let $i : \Delta \to L$ be the inclusion morphism. If Δ is closed under disjunctions, the following statements are equivalent.*

(1) *For every $\Phi \subseteq L$, there is some $\Psi \subseteq \Delta$ such that $\Phi \equiv_L \Psi$.*

(2) *$\Phi \equiv_L \Phi^\vDash \cap \Delta$, for all $\Phi \subseteq L$.*

(3) *$\mathfrak{S}(i) : \mathfrak{S}(L) \to \mathfrak{S}(\Delta)$ is a homeomorphism.*

Proof. (1) $\Rightarrow$ (2) Let $\Phi \subseteq L$. Clearly, $\Phi \vDash \Phi^\vDash \cap \Delta$. Hence, we only need to prove that $\Phi^\vDash \cap \Delta \vDash \Phi$. By (1), there is a set $\Psi \subseteq \Delta$ such that $\Psi \equiv_L \Phi$. Hence, $\Phi \vDash \Psi$ implies that $\Psi \subseteq \Phi^\vDash \cap \Delta$. Since $\Psi \vDash \Phi$, it therefore follows that $\Phi^\vDash \cap \Delta \vDash \Phi$.

(2) $\Rightarrow$ (3) Suppose that every $\Phi \subseteq L$ is equivalent to $\Phi^\vDash \cap \Delta$. We have to prove that $\mathfrak{S}(i)$ is continuous, closed, and bijective. Continuity and surjectivity follow from Corollary 2.13.

For injectivity, suppose that $\mathfrak{p}, \mathfrak{q} \in S(L)$ are two types with $\mathfrak{S}(i)(\mathfrak{p}) = \mathfrak{S}(i)(\mathfrak{q})$. By (2), $\mathfrak{p} \equiv_L \mathfrak{p} \cap \Delta$ and $\mathfrak{q} \equiv_L \mathfrak{q} \cap \Delta$. Consequently, we have

$$\mathfrak{p} \equiv_L \mathfrak{p} \cap \Delta = \mathfrak{S}(i)(\mathfrak{p}) = \mathfrak{S}(i)(\mathfrak{q}) = \mathfrak{q} \cap \Delta \equiv_L \mathfrak{q} .$$

It follows that $\mathfrak{p} = \mathfrak{q}$, as desired.

It remains to prove that $\mathfrak{S}(i)$ is closed. Since $\mathfrak{S}(i)$ is injective, it is sufficient, by Lemma B5.2.3, to prove that $\mathfrak{S}(i)[\langle \Phi \rangle_L]$ is closed, for every $\Phi \subseteq L$. By (2), $\Phi \equiv_L \Phi^{\vDash} \cap \Delta$. Hence, it follows by Lemma 3.9 that the set $\mathfrak{S}(i)[\langle \Phi \rangle_L] = \langle \Phi^{\vDash} \cap \Delta \rangle_\Delta$ is closed.

(3) ⇒ (1) Suppose that $\mathfrak{S}(i)$ is a homeomorphism. To show that every $\Phi \subseteq L$ is equivalent to some $\Psi \subseteq \Delta$, we fix $\Phi \subseteq L$. Since $\langle \Phi \rangle_L$ is closed in $\mathfrak{S}(L)$, it follows that $C := \mathfrak{S}(i)[\langle \Phi \rangle_L]$ is a closed subset of $\mathfrak{S}(\Delta)$. By Lemma 2.2, there exists a set $\Psi \subseteq \Delta$ such that $C = \langle \Psi \rangle_\Delta$. Hence, $\mathfrak{S}(i)[\langle \Phi \rangle_L] = \langle \Psi \rangle_\Delta$ implies, by Lemma 3.9, that $\Phi \equiv_L \Psi$. □

Exercise 3.2. Show that the preceding lemma may fail if Δ is not closed under disjunctions.

For logics with compact type space, we can strengthen this proposition as follows.

Proposition 3.11. *Let L be a boolean closed logic such that $\mathfrak{S}(L)$ is compact, let $\Delta \subseteq L$, and let $i : \Delta \to L$ be the inclusion morphism. The following statements are equivalent:*

(1) *For every $\varphi \in L$, there is some $\psi \in \Delta$ with $\psi \equiv_L \varphi$.*

(2) *Δ is a retract of L.*

(3) *Δ is boolean closed and*

$$\operatorname{spec}(\mathfrak{Lb}(i)) : \operatorname{spec}(\mathfrak{Lb}(L)) \to \operatorname{spec}(\mathfrak{Lb}(\Delta))$$

is a homeomorphism.

(4) *Δ is boolean closed and $\mathfrak{S}(i) : \mathfrak{S}(L) \to \mathfrak{S}(\Delta)$ is a homeomorphism.*

Proof. (1) ⇔ (2) was already proved in Lemma 3.4.

(3) ⇒ (2) According to Lemma B5.6.7, $\mathfrak{Lb}(i) : \mathfrak{Lb}(\Delta) \to \mathfrak{Lb}(L)$ is an isomorphism. Hence, the claim follows by Lemma 3.4.

(1) ⇒ (4) $\mathfrak{S}(i)$ is a homeomorphism by Corollary 3.3 (a). Therefore, we only need to show that Δ is boolean closed. Let $\varphi, \vartheta \in \Delta$. Then $\varphi \wedge \vartheta$,

$\varphi \vee \vartheta$, and $\neg\varphi$ are L-formulae. By (1), there are formulae $\psi_0, \psi_1, \psi_2 \in \Delta$ with

$$\psi_0 \equiv_L \varphi \wedge \vartheta\,, \quad \psi_1 \equiv_L \varphi \vee \vartheta\,, \quad \text{and} \quad \psi_2 \equiv_L \neg\varphi\,.$$

Hence, Δ is boolean closed.

(4) ⇒ (3) According to Lemma 2.17, $\mathfrak{S}(L)$ and $\mathfrak{S}(\Delta)$ are both compact. Therefore, we can use Lemma 2.8 to obtain homeomorphisms

$$h : \mathfrak{S}(L) \to \operatorname{spec}(\mathfrak{Lb}(L)) \quad \text{and} \quad h_0 : \mathfrak{S}(\Delta) \to \operatorname{spec}(\mathfrak{Lb}(\Delta))\,.$$

If $\mathfrak{S}(i)$ is a homeomorphism, then so is $\operatorname{spec}(\mathfrak{Lb}(i)) = h_0 \circ \mathfrak{S}(i) \circ h^{-1}$. □

Corollary 3.12. *Let L be a boolean closed logic such that $\mathfrak{S}(L)$ is compact, and let $\Delta \subseteq \Phi \subseteq L$. The following statements are equivalent.*

(1) *Every formula in Φ is equivalent to a finite boolean combination of formulae in Δ.*

(2) $\mathfrak{p} \cap \Delta = \mathfrak{q} \cap \Delta$ *implies* $\mathfrak{p} \cap \Phi = \mathfrak{q} \cap \Phi$, *for all* $\mathfrak{p}, \mathfrak{q} \in S(L)$.

Proof. (1) ⇒ (2) is obvious. For (2) ⇒ (1), let Δ_+ and Φ_+ be the boolean closures of, respectively, Δ and Φ and let $i : \Delta_+ \to \Phi_+$ be the inclusion morphism. By Proposition 3.11, it is sufficient to show that $\mathfrak{S}(i) : \mathfrak{S}(\Phi_+) \to \mathfrak{S}(\Delta_+)$ is a homeomorphism.

According to Lemma 2.16, $\mathfrak{S}(i)$ is continuous, closed, and surjective. Hence, it remains to prove that it is injective. Suppose that $\mathfrak{S}(i)(\mathfrak{p}) = \mathfrak{S}(i)(\mathfrak{q})$. Fix models $\mathfrak{J}_0 \vDash \mathfrak{p}$ and $\mathfrak{J}_1 \vDash \mathfrak{q}$, and set $\mathfrak{p}_+ := \operatorname{Th}_L(\mathfrak{J}_0)$ and $\mathfrak{q}_+ := \operatorname{Th}_L(\mathfrak{J}_1)$. Then

$$\mathfrak{p}_+ \cap \Delta_+ = \mathfrak{p} \cap \Delta_+ = \mathfrak{S}(i)(\mathfrak{p}) = \mathfrak{S}(i)(\mathfrak{q}) = \mathfrak{q} \cap \Delta_+ = \mathfrak{q}_+ \cap \Delta_+\,.$$

In particular, we have $\mathfrak{p}_+ \cap \Delta = \mathfrak{q}_+ \cap \Delta$. By (2), we obtain $\mathfrak{p}_+ \cap \Phi = \mathfrak{q}_+ \cap \Phi$, which implies that

$$\mathfrak{p} = \mathfrak{p}_+ \cap \Phi_+ = \mathfrak{q}_+ \cap \Phi_+ = \mathfrak{q}\,.$$ □

As an application we prove the intuitively obvious fact that, if there are more formulae than types, many formulae have to be equivalent.

Proposition 3.13. *Let L be a boolean closed logic such that $\mathfrak{S}(L)$ is compact. There exists a retract L_0 of L of size $|L_0| \leq |S(L)| \oplus \aleph_0$.*

Proof. Let $(\mathfrak{p}_\alpha)_{\alpha<\kappa}$ be an enumeration of $S(L)$ without repetitions. For every pair of indices $\alpha, \beta < \kappa$, $\alpha \neq \beta$, fix a formula $\psi_{\alpha\beta} \in \mathfrak{p}_\alpha \setminus \mathfrak{p}_\beta$. Set $\Psi := \{ \psi_{\alpha\beta} \mid \alpha, \beta < \kappa \}$ and let L_0 be the set of all finite boolean combinations of formulae in Ψ. Then $|L_0| \leq \kappa \otimes \kappa \otimes \aleph_0 \leq \kappa \oplus \aleph_0$ and

$$\mathfrak{p}_\alpha \cap \Psi = \mathfrak{p}_\beta \cap \Psi \quad \text{implies} \quad \mathfrak{p}_\alpha = \mathfrak{p}_\beta \,.$$

Therefore, Corollary 3.12 implies that L_0 is a retract of L. □

Corollary 3.14. *Let $T \subseteq \mathrm{FO}^0[\Sigma]$ be a first-order theory. There exists a subset $\Sigma_0 \subseteq \Sigma$ of size $|\Sigma_0| \leq |S^{<\omega}(T)|$ and a family of formulae $\varphi_\xi(\bar{x})$, for $\xi \in \Sigma \setminus \Sigma_0$, such that, for every model $\mathfrak{M}$ of T,*

$$\xi^{\mathfrak{M}} = \varphi_\xi^{\mathfrak{M}|_{\Sigma_0}} \,, \quad \textit{for all } \xi \in \Sigma \setminus \Sigma_0 \,.$$

Proof. For each finite tuple $\bar{s}$ of sorts, we can use Proposition 3.13 to obtain a retract $\Delta_{\bar{s}}$ of $\mathrm{FO}^{\bar{s}}[\Sigma]/T$ such that $|\Delta_{\bar{s}}| \leq |S^{\bar{s}}(T)|$. Let Σ_0 be the set of all symbols from Σ that appear in some $\Delta_{\bar{s}}$. Note that $S^{\bar{s}}(T) \neq \emptyset$ implies that

$$|S^{<\omega}(T)| = \Big|\bigcup_{\bar{s}} S^{\bar{s}}(T)\Big| \geq \aleph_0 \,.$$

Hence,

$$|\Sigma_0| \leq \sum_{\bar{s}} |\Delta_{\bar{s}}| \oplus \aleph_0 = |S^{<\omega}(T)| \oplus \aleph_0 = |S^{<\omega}(T)| \,.$$

Furthermore, for every relation symbol $R \in \Sigma \setminus \Sigma_0$ of type $\bar{s}$, there exists a formula $\varphi_R(\bar{x}) \in \Delta_{\bar{s}} \subseteq \mathrm{FO}^{\bar{s}}[\Sigma_0]$ such that $R\bar{x} \equiv \varphi_R(\bar{x})$. Similarly, for every function symbol $f \in \Sigma \setminus \Sigma_0$ of type $\bar{s} \to t$, there exists a formula $\varphi_f(\bar{x}, y) \in \Delta_{\bar{s}t} \subseteq \mathrm{FO}^{\bar{s}t}[\Sigma_0]$ such that $f\bar{x} = y \equiv \varphi_f(\bar{x}, y)$. □

4. *Local type spaces*

For technical reasons we will consider in the next section certain quotients of first-order type spaces $\mathfrak{S}^{\bar{s}}(U)$. To define these quotients we consider a restriction $L|_\Delta$ of some logic L and we equip the corresponding set of types $S(L|_\Delta)$ with a topology that is finer than the usual one. Our aim is to show that, for first-order logic, this topology coincides with the usual one. For simplicity, we only consider logics L that are closed under disjunction.

Definition 4.1. Let L be a logic that is closed under disjunction, $\Delta \subseteq L$ a fragment, and let $i : \Delta \to L$ be the inclusion morphism. We denote by $\mathfrak{S}_\Delta(L)$ the topological space with universe $S(\Delta)$ where the topology consists of all sets

$$\langle \Phi \rangle_\Delta := \mathfrak{S}(i)[\langle \Phi \rangle_L], \quad \text{for } \Phi \subseteq L .$$

Lemma 4.2. *Let L be a logic that is closed under disjunctions, $\Delta \subseteq L$, and let $i : \Delta \to L$ be the inclusion morphism.*

(a) *The restriction function*

$$\rho_\Delta : \mathfrak{S}(L) \to \mathfrak{S}_\Delta(L) : \mathfrak{p} \mapsto \mathfrak{p} \cap \Delta$$

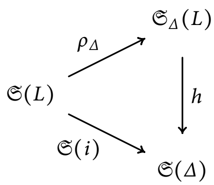
is closed and surjective.

(b) *The identity function*

$$h : \mathfrak{S}_\Delta(L) \to \mathfrak{S}(\Delta) : \mathfrak{p} \mapsto \mathfrak{p}$$

is continuous and bijective.

(c) $\mathfrak{S}_\Delta(L) = \mathfrak{S}(\Delta)$ *if, and only if,* $\mathfrak{S}(i) : \mathfrak{S}(L) \to \mathfrak{S}(\Delta)$ *is closed and* ρ_Δ *is continuous.*



Proof. First, note that $\mathfrak{S}(i) = h \circ \rho_\Delta$ since

$$h(\rho_\Delta(\mathfrak{p})) = h(\mathfrak{p} \cap \Delta) = \mathfrak{p} \cap \Delta = \mathfrak{S}(i)(\mathfrak{p}), \quad \text{for every } \mathfrak{p} \in S(L).$$

(a) Since L is closed under disjunctions, each closed set of $\mathfrak{S}(L)$ is of the form $\langle \Phi \rangle_L$, for some $\Phi \subseteq L$. The function ρ_Δ is closed since, for every $\Phi \subseteq L$,

$$\rho_\Delta[\langle \Phi \rangle_L] = \{ \mathfrak{p} \cap \Delta \mid \mathfrak{p} \in \langle \Phi \rangle_L \} = \{ \mathfrak{S}(i)(\mathfrak{p}) \mid \mathfrak{p} \in \langle \Phi \rangle_L \} = \langle \Phi \rangle_\Delta$$

is a closed set of $\mathfrak{S}(\Delta)$.

For surjectivity, note that h^{-1} and $\mathfrak{S}(i)$ are both surjective. Therefore, so is $\rho_\Delta = h^{-1} \circ \mathfrak{S}(i)$.

(b) h is clearly bijective. For continuity, note that $h \circ \rho_\Delta = \mathfrak{S}(i)$. Since $\mathfrak{S}(i)$ is surjective, it follows by Lemma A2.1.10 for a closed set $C \subseteq S(\Delta)$ that

$$\begin{aligned} h^{-1}[C] &= h^{-1}[\mathfrak{S}(i)[\mathfrak{S}(i)^{-1}[C]]] \\ &= h^{-1}[h[\rho_\Delta[\mathfrak{S}(i)^{-1}[C]]]] = \rho_\Delta[\mathfrak{S}(i)^{-1}[C]]. \end{aligned}$$

This set is closed, since $\mathfrak{S}(i)$ is continuous and ρ_Δ is closed.

(c) ($\Rightarrow$) If $\mathfrak{S}_\Delta(L) = \mathfrak{S}(\Delta)$, then h is a homeomorphism. Hence, $\rho_\Delta = h^{-1} \circ \mathfrak{S}(i)$ is a composition of continuous functions and, therefore, continuous. Similarly, $\mathfrak{S}(i) = h \circ \rho_\Delta$ is a composition of closed functions and, therefore, closed.

($\Leftarrow$) It is sufficient to show that the identity function

$$h : \mathfrak{S}_\Delta(L) \to \mathfrak{S}(\Delta) : \mathfrak{p} \mapsto \mathfrak{p}$$

is a homeomorphism. We have already seen in (b) that it is bijective and continuous. Hence, it remains to prove that h is closed.

By assumption, ρ_Δ is continuous and $\mathfrak{S}(i)$ is closed. It follows as in (b) that

$$h[C] = \mathfrak{S}(i)[\rho_\Delta^{-1}[C]]$$

is closed, for every closed set $C \subseteq S_\Delta(L)$. □

In the applications below we are interested in the case where $\mathfrak{S}(L)$ is compact and Δ closed under negation. In this situation the topologies of $\mathfrak{S}_\Delta(L)$ and $\mathfrak{S}(\Delta)$ coincide.

Theorem 4.3. *Let L be a boolean closed logic such that $\mathfrak{S}(L)$ is compact and let $\Delta \subseteq L$.*

(a) *The restriction function*

$$\rho_\Delta : \mathfrak{S}(L) \to \mathfrak{S}_\Delta(L) : \mathfrak{p} \mapsto \mathfrak{p} \cap \Delta$$

is continuous, closed, and surjective.

(b) *If Δ is closed under negation, then $\mathfrak{S}_\Delta(L) = \mathfrak{S}(\Delta)$.*

Proof. (a) We have already seen in Lemma 4.2 (a) that ρ_Δ is closed and surjective. Hence, it remains to prove that it is continuous.

Let Δ_+ be the set of all finite boolean combinations of formulae in Δ. We claim that

$$\rho_\Delta^{-1}[\langle \Phi \rangle_\Delta] = \langle \Phi^\vDash \cap \Delta_+ \rangle_L \,.$$

($\subseteq$) Let $\mathfrak{p} \in \rho_\Delta^{-1}[\langle \Phi \rangle_\Delta]$. Then $\mathfrak{p} \cap \Delta = \rho_\Delta(\mathfrak{p}) \in \langle \Phi \rangle_\Delta$ and there is some type $\mathfrak{q} \in \langle \Phi \rangle_L$ with $\mathfrak{q} \cap \Delta = \mathfrak{p} \cap \Delta$. Since every formula in $\mathfrak{q} \cap \Delta_+$ is a boolean combination of formulae in $\mathfrak{q} \cap \Delta$, it follows that $\mathfrak{q} \cap \Delta_+ = \mathfrak{p} \cap \Delta_+$. Hence,

$$\Phi^\vDash \subseteq \mathfrak{q} \quad \text{implies} \quad \Phi^\vDash \cap \Delta_+ \subseteq \mathfrak{q} \cap \Delta_+ = \mathfrak{p} \cap \Delta_+ \,.$$

Consequently, $\mathfrak{p} \in \langle \Phi^\vDash \cap \Delta_+ \rangle_L$.

($\supseteq$) Let $\mathfrak{p} \in \langle \Phi^\vDash \cap \Delta_+ \rangle_L$ and set $\mathfrak{p}_0 := \mathfrak{p} \cap \Delta_+$. If there is some $\mathfrak{q} \in S(L)$ with $\Phi \cup \mathfrak{p}_0 \subseteq \mathfrak{q}$, then

$$\mathfrak{q} \cap \Delta_+ = \mathfrak{p}_0 \quad \text{implies} \quad \rho_\Delta(\mathfrak{p}) = \mathfrak{p} \cap \Delta = \mathfrak{p}_0 \cap \Delta = \mathfrak{q} \cap \Delta \in \langle \Phi \rangle_\Delta \,.$$

Hence, $\mathfrak{p} \in \rho_\Delta^{-1}[\langle \Phi \rangle_\Delta]$. Consequently, it remains to show that $\Phi \cup \mathfrak{p}_0$ is satisfiable.

For a contradiction, suppose otherwise. Then

$$\langle \Phi \rangle_L \cap \bigcap_{\psi \in \mathfrak{p}_0} \langle \psi \rangle_L = \langle \Phi \cup \mathfrak{p}_0 \rangle_L = \varnothing\,.$$

Since $\mathfrak{S}(L)$ is compact, we can find a finite subset $\Psi \subseteq \mathfrak{p}_0$ such that

$$\langle \Phi \rangle_L \cap \bigcap_{\psi \in \Psi} \langle \psi \rangle_L = \varnothing\,.$$

Hence, $\Phi \vDash \neg \bigwedge \Psi$. Note that $\Psi \subseteq \Delta_+$ implies $\neg \bigwedge \Psi \in \Delta_+$. Hence, $\neg \bigwedge \Psi \in \Phi^{\vDash} \cap \Delta_+ \subseteq \mathfrak{p}_0$ and $\mathfrak{p}_0$ is inconsistent. A contradiction.

(b) We have seen in (a) that ρ_Δ is continuous. By Lemma 4.2 (c), it is therefore sufficient to show that $\mathfrak{S}(i) : \mathfrak{S}(L) \to \mathfrak{S}(\Delta)$ is closed. Let $\Delta_+ \subseteq L$ be the set of all finite boolean combinations of formulae in Δ, and let $i_0 : \Delta \to \Delta_+$ and $i_+ : \Delta_+ \to L$ be the corresponding inclusion morphisms. Then $\mathfrak{S}(i_+)$ is closed by Lemma 2.16, and $\mathfrak{S}(i_0)$ is closed by Lemma 3.5. Hence, $\mathfrak{S}(i) = \mathfrak{S}(i_0) \circ \mathfrak{S}(i_+)$ is also closed. □

We will mainly use type spaces of the form $\mathfrak{S}_\Delta(L)$ in the case of first-order logic. In this case the definitions are as follows.

Definition 4.4. Let $T \subseteq \mathrm{FO}^0[\Sigma]$ be a theory and $\Delta \subseteq \mathrm{FO}[\Sigma, X \cup Y]$ a set of formulae where X and Y are disjoint sets of variables. For a set U of parameters, we set

$$\begin{aligned}\Delta_U^{\neg} := {} & \{\, \varphi(\bar x; \bar c) \mid \varphi(\bar x; \bar y) \in \Delta,\ \bar x \subseteq X,\ \bar y \subseteq Y,\ \bar c \subseteq U \,\} \\ & \cup \{\, \neg\varphi(\bar x; \bar c) \mid \varphi(\bar x; \bar y) \in \Delta,\ \bar x \subseteq X,\ \bar y \subseteq Y,\ \bar c \subseteq U \,\}\,.\end{aligned}$$

(a) A partial type $\mathfrak{p}$ over a set U is a *Δ-type* if $\mathfrak{p} \subseteq \Delta_U^{\neg}$. For $\Delta = \{\varphi\}$ we simply speak of a φ-type.

(b) The *restriction* of a partial type $\mathfrak{p}$ is the type

$$\mathfrak{p}|_\Delta := \mathfrak{p} \cap \Delta_U^{\neg}\,.$$

(c) Let $\mathfrak{M}$ be a structure. The *Δ-type* of a tuple $\bar a \subseteq M$ over a set $U \subseteq M$ is

$$\mathrm{tp}_\Delta(\bar a/U) := \mathrm{tp}(\bar a/U)|_\Delta\,.$$

(d) A Δ-type $\mathfrak{p}$ over U is *complete* if, for every formula $\varphi(\bar{x};\bar{y}) \in \Delta$ and each tuple $\bar{c} \subseteq U$, we have $\varphi(\bar{x};\bar{c}) \in \mathfrak{p}$ or $\neg\varphi(\bar{x};\bar{c}) \in \mathfrak{p}$.

(e) The space of all complete Δ-types over U is

$$\mathfrak{S}_\Delta(U) := \mathfrak{S}_{\Delta_U^\neg}\big(\mathrm{FO}[\Sigma_U, X]/T(U)\big)\,.$$

As usual we also write $\mathfrak{S}_\Delta(T)$ for $\mathfrak{S}_\Delta(\varnothing)$.

Since first-order type spaces are compact, it follows by the above results that $\mathfrak{S}_\Delta(U)$ is equal to $\mathfrak{S}((\mathrm{FO}[\Sigma_U, X]/T(U))|_{\Delta_U^\neg})$. Our aim is to show that this definition does not depend on the signature Σ.

Theorem 4.5. *Let $T \subseteq \mathrm{FO}^\circ[\Sigma]$ be a theory and $\Delta \subseteq \mathrm{FO}[\Sigma, X \cup Y]$ a set of formulae where X and Y are disjoint sets of variables. For a set U of parameters, set*

$$\Delta_U^\neg := \{\, \varphi(\bar{x};\bar{c}) \mid \varphi(\bar{x};\bar{y}) \in \Delta,\ \bar{c} \subseteq U \,\} \\ \cup \{\, \neg\varphi(\bar{x};\bar{c}) \mid \varphi(\bar{x};\bar{y}) \in \Delta,\ \bar{c} \subseteq U \,\}\,.$$

(a) $\mathfrak{S}_\Delta(U) = \mathfrak{S}\big((\mathrm{FO}[\Sigma_U, X]/T(U))\,\big|_{\Delta_U^\neg}\big)\,.$

(b) *If $\Delta \subseteq \mathrm{FO}[\Sigma_0, X_0 \cup Y]$, for some $\Sigma_0 \subseteq \Sigma$ and $X_0 \subseteq X$, then*

$$\mathfrak{S}_\Delta(T) = \mathfrak{S}_\Delta(T \cap \mathrm{FO}^\circ[\Sigma_0])\,,$$

where the local type space on the left-hand side is with respect to the logic $\mathrm{FO}[\Sigma, X]$ and the one one the right-hand side with respect to $\mathrm{FO}[\Sigma_0, X_0]$.

Proof. (a) follows by Theorem 4.3 (b), while (b) follows from (a) and Proposition 2.21 (a) (treating the free variables from X and X_0 as constant symbols). □

Corollary 4.6. *Let $T \subseteq \mathrm{FO}^\circ[\Sigma]$ be a theory, $\Delta \subseteq \mathrm{FO}[\Sigma, X \cup Y]$ a set of formulae, U a set of parameters, and $\Delta_U^\neg$ the set from Definition 4.4. Then*

$$\mathfrak{S}_\Delta(U) \cong \mathrm{spec}(\mathfrak{Lb}(\Delta_U^\neg))\,,$$

where $\mathfrak{Lb}(\Delta_U^\neg)$ denotes the subalgebra of $\mathfrak{Lb}(\mathrm{FO}[\Sigma_U, X]/T(U))$ generated by $\Delta_U^\neg$.

Proof. Set $L := \mathrm{FO}[\Sigma_U, X]/T(U)$ and let Δ^+ be the boolean closure of $\Delta_U^\neg$. By Theorems 4.5 (a) and 2.18 and Lemma 3.5, it follows that

$$\begin{aligned}\mathfrak{S}_\Delta(U) = \mathfrak{S}(\Delta_U^\neg) &\cong \mathfrak{S}(\Delta^+) \\ &\cong \operatorname{spec}(\mathfrak{Lb}(\Delta^+)) = \operatorname{spec}(\mathfrak{Lb}(\Delta_U^\neg)) . \end{aligned} \qquad \square$$

5. Stable theories

In this section we consider the size of first-order type spaces. First, let us state two trivial bounds.

Lemma 5.1. *Let T be a complete first-order theory and $\bar{s}$ a sequence of sorts. Then*

$$|U| \leq |S^{\bar{s}}(U)| \leq 2^{|T| \oplus |U| \oplus |\bar{s}|}, \quad \textit{for every set } U \textit{ of parameters.}$$

One situation where the size of a type space is important is when we want to construct a model realising all types. First note that we can use the Compactness Theorem to show that, for every structure $\mathfrak{A}$, we can add a tuple realising any given type $\mathfrak{p}$ over a subset $U \subseteq A$.

Lemma 5.2. *Let $\mathfrak{A}$ be a Σ-structure, $U \subseteq A$, and $\mathfrak{p} \in S^\alpha(U)$. There exists an elementary extension $\mathfrak{B} \succeq \mathfrak{A}$ of size $|B| \leq |A| \oplus |\Sigma| \oplus |\alpha| \oplus \aleph_0$ in which $\mathfrak{p}$ is realised.*

Proof. Let $\Phi := \mathfrak{p} \cup \mathrm{Th}(\mathfrak{A}_A)$. We regard the free variables x_i, $i < \alpha$, of $\mathfrak{p}$ as constant symbols. If Φ is satisfiable then, by the Theorem of Löwenheim and Skolem, there exists a model $\mathfrak{B} \models \Phi$ of size $|B| \leq |A| \oplus |\Sigma| \oplus |\alpha| \oplus \aleph_0$. Furthermore, we have $\mathfrak{B} \succeq \mathfrak{A}$, by Lemma C2.2.3, and there exists some $\bar{a} \in B^\alpha$ with $\mathrm{tp}(\bar{a}/U) = \mathfrak{p}$.

Hence, it is sufficient to show that Φ is satisfiable. Let $\Phi_0 \subseteq \Phi$ be finite. We write

$$\bigwedge \Phi_0 = \varphi(\bar{x}, \bar{a}) \wedge \psi(\bar{a}, \bar{b})$$

where $\bar{a} \subseteq U$, $\bar{b} \subseteq A \setminus U$, $\mathfrak{p} \vDash \varphi(\bar{x}, \bar{a})$, and $\mathfrak{A} \vDash \psi(\bar{a}, \bar{b})$. The last statement implies that $\exists \bar{y} \psi(\bar{a}, \bar{y}) \in \mathrm{Th}(\mathfrak{A}_U)$. By definition of a type, there exists a model $\mathfrak{C} \vDash \mathfrak{p} \cup \mathrm{Th}(\mathfrak{A}_U)$. In particular, we have

$$\mathfrak{C} \vDash \varphi(\bar{x}, \bar{a}) \wedge \exists \bar{y} \psi(\bar{a}, \bar{y}) .$$

Choose a tuple $\bar{c} \subseteq C$ such that $\mathfrak{C} \vDash \psi(\bar{a}, \bar{c})$. We obtain a model of Φ_0 by interpreting the constant symbol b_i by the element c_i, for every i. □

Corollary 5.3. *For every Σ-structure $\mathfrak{A}$, there exists an elementary extension $\mathfrak{B} \succeq \mathfrak{A}$ of size at most $|S^{<\omega}(A)|$ in which every type $\mathfrak{p} \in S^{<\omega}(A)$ is realised.*

Proof. According to Corollary 3.14, we can find a signature $\Sigma_0 \subseteq \Sigma_A$ of size $|\Sigma_0| \leq |S^{<\omega}(A)|$ such that there exists a retraction

$$\langle \alpha, \beta \rangle : \mathrm{FO}^{<\omega}[\Sigma_A]/T(A) \to \mathrm{FO}^{<\omega}[\Sigma_0]/T_0 ,$$

where $T_0 := T(A) \cap \mathrm{FO}^0[\Sigma_0]$. If we can show that there exists a model $\mathfrak{B}$ of T_0 realising every type in $S^{<\omega}(T_0)$, it follows that its expansion $\beta(\mathfrak{B})$ is a model of $T(A)$ realising every type in $S^{<\omega}(A)$. Therefore, we may assume without loss of generality that $|\Sigma| \leq |S^{<\omega}(A)|$.

Fix an enumeration $(\mathfrak{p}_\alpha)_{\alpha<\kappa}$ of $S^{<\omega}(A)$. We can use Lemma 5.2 to find, for every $\alpha < \kappa$, an elementary extension $\mathfrak{C}_\alpha \succeq \mathfrak{A}$ realising $\mathfrak{p}_\alpha$. By Lemma C2.5.7, there exists a common elementary extension $\mathfrak{C}$ of all $\mathfrak{C}_\alpha$. It follows that $\mathfrak{C}$ realises every type $\mathfrak{p}_\alpha$. By the Theorem of Löwenheim and Skolem, we can find an elementary substructure $\mathfrak{A} \preceq \mathfrak{B} \preceq \mathfrak{C}$ of size at most $|S^{<\omega}(A)| \oplus |\Sigma| \oplus \aleph_0$ such that every $\mathfrak{p}_\alpha$ is realised in $\mathfrak{B}$. Since $S^{<\omega}(A) = \bigcup_{n<\omega} S^n(A)$ is infinite, we have $|S^{<\omega}(A)| \oplus |\Sigma| \oplus \aleph_0 = |S^{<\omega}(A)|$ and the claim follows. □

The number of different types a theory possesses also serves as a rough measure of its complexity. Intuitively, if there are only a few types the number of different configurations that can appear in a model is small. Before considering full type spaces $\mathfrak{S}^{\bar{s}}(U)$, we start by looking at those of the form $\mathfrak{S}_\varphi(U)$.

Definition 5.4. Let T be a complete first-order theory and κ an infinite cardinal. A formula $\varphi(\bar{x};\bar{y})$ is *κ-stable* (with respect to T) if we have $|S_\varphi(U)| \leq \kappa$, for all sets U of size $|U| \leq \kappa$. We call $\varphi(\bar{x};\bar{y})$ *stable* if it is κ-stable, for some infinite cardinal κ. Otherwise, $\varphi(\bar{x};\bar{y})$ is *unstable.*

Example. If $\sim$ is an equivalence relation with infinitely many classes, then the formula $x \sim y$ is κ-stable, for all infinite κ, since

$$|S_{x\sim y}(U)| = |U/{\sim}| \oplus 1 \leq |U| \oplus 1\,.$$

The definition does not tell us much about stable formulae. We will therefore present three equivalent characterisations, two combinatorial ones that can be checked more easily, and one logical characterisation.

The equivalence proofs rest on two combinatorial results. The first one is a special case of the Theorem of Ramsey. We will prove the full version in Section E5.1 below.

Lemma 5.5. *Let $(a_n)_{n<\omega}$ be a sequence of elements and let $(B_n)_{n<\omega}$ be a sequence of sets. There exists an infinite set $I \subseteq \omega$ such that either*

$$a_i \in B_k\,, \quad \text{for all } i < k \text{ in } I\,,$$

or

$$a_i \notin B_k\,, \quad \text{for all } i < k \text{ in } I\,.$$

Proof. We construct an increasing sequence $n_0 < n_1 < \cdots$ of indices, a sequence $m_0, m_1, \ldots \in [2]$ of numbers, and a decreasing sequence $J_0 \supseteq J_1 \supseteq \cdots$ of infinite sets such that, for every $i < \omega$, we have $n_i \in J_i$ and either

$$m_i = 0 \quad \text{and} \quad a_{n_i} \notin B_k\,, \quad \text{for all } k \in J_{i+1}\,,$$

or

$$m_i = 1 \quad \text{and} \quad a_{n_i} \in B_k\,, \quad \text{for all } k \in J_{i+1}\,.$$

We start with $n_0 := 0$ and $J_0 := \omega$. By induction, suppose that we have already defined n_i and J_i. Set

$$L_0 := \{\, k \in J_i \mid a_{n_i} \notin B_k \,\} \quad \text{and} \quad L_1 := \{\, k \in J_i \mid a_{n_i} \in B_k \,\}\,.$$

Then $J_i = L_0 \cup L_1$. As J_i is infinite, at least one of L_0 and L_1 must also be infinite. Choose $m_i < 2$ such that L_{m_i} is infinite. We set

$$J_{i+1} := L_{m_i} \smallsetminus [n_i + 1] \quad \text{and} \quad n_{i+1} := \min J_{i+1} .$$

Having defined $(n_i)_{i<\omega}$, $(m_i)_{i<\omega}$, and $(J_i)_{i<\omega}$, we consider the sets

$$M_0 := \{\, i < \omega \mid m_i = 0 \,\} \quad \text{and} \quad M_1 := \{\, i < \omega \mid m_i = 1 \,\} .$$

Note that $n_j \in J_j \subseteq J_i$ implies that

$$\begin{aligned} & a_{n_i} \notin B_{n_j} , \quad \text{for all } i < j \text{ in } M_0 , \\ \text{and} \quad & a_{n_i} \in B_{n_j} , \quad \text{for all } i < j \text{ in } M_1 . \end{aligned}$$

Since $M_0 \cup M_1 = \omega$, at least one of M_0 and M_1 is infinite. If M_0 is infinite, we can therefore set $I := \{\, n_i \mid i \in M_0 \,\}$. Otherwise, we use $I := \{\, n_i \mid i \in M_1 \,\}$. ☐

Theorem 5.6 (Erdős, Makkai). *Let X be an infinite set and $S \subseteq \mathscr{P}(X)$ a family of size $|S| > |X|$. Then there are sequences $(a_i)_{i<\omega}$ in X and $(B_i)_{i<\omega}$ in S such that either*

$$\begin{aligned} & a_i \in B_k \quad \text{iff} \quad i \le k , \quad \textit{for all } i, k < \omega , \\ \textit{or} \quad & a_i \in B_k \quad \text{iff} \quad i \ge k , \quad \textit{for all } i, k < \omega . \end{aligned}$$

Proof. For every pair of disjoint finite subsets $Y, Z \subseteq X$, choose, if possible, a set $B \in S$ with $Y \subseteq B$ and $Z \subseteq X \smallsetminus B$. Let $S_0 \subseteq S$ be the set of the chosen subsets B. As there are only $|X^{<\omega} \times X^{<\omega}| = |X|$ pairs of finite subsets, it follows that $|S_0| \le |X| < |S|$. Consequently, there exists a set $A \in S$ that cannot be expressed as a finite boolean combination of sets from S_0. (We allow empty boolean combinations, so that A is different from $\varnothing$ and X.)

We inductively construct sequences $(c_n)_{n<\omega}$ in A, $(d_n)_{n<\omega}$ in $X \smallsetminus A$, and $(B_n)_{n<\omega}$ in S_0 such that, for all n,

- $\{c_0, \dots, c_n\} \subseteq B_n$,

- $\{d_0, \dots, d_n\} \subseteq X \setminus B_n$, and
- $c_i \in B_n \Leftrightarrow d_i \in B_n$, for all $i > n$.

For the inductive step, suppose that we have already defined elements $c_0, \dots, c_{n-1}, d_0, \dots, d_{n-1}$, and sets $B_0, \dots, B_{n-1}$. Since A is not a boolean combination of $B_0, \dots, B_{n-1}$, there are elements $c_n \in A$ and $d_n \in X \setminus A$ such that

$$c_n \in B_k \quad \text{iff} \quad d_n \in B_k\,, \quad \text{for all } k < n\,.$$

Then $\{c_0, \dots, c_n\} \subseteq A$ and $\{d_0, \dots, d_n\} \subseteq X \setminus A$. By choice of S_0, it follows that we can choose a set $B_n \in S_0$ with $\{c_0, \dots, c_n\} \subseteq B_n$ and $\{d_0, \dots, d_n\} \subseteq X \setminus B_n$. This concludes the inductive step.

We have constructed sequences such that

$$c_i \in B_k \quad \text{and} \quad d_i \notin B_k\,, \quad \text{for } i \leq k\,,$$

$$c_i \in B_k \Leftrightarrow d_i \in B_k\,, \qquad \text{for } i > k\,.$$

By Lemma 5.5, there exists an infinite subset $I \subseteq \omega$ such that either

- $c_i \notin B_k$, for all indices $i > k$ in I, or
- $c_i \in B_k$, for all indices $i > k$ in I.

In the first case, the sequences $(c_n)_{n \in I}$ and $(B_n)_{n \in I}$ satisfy

$$c_i \in B_k \quad \text{iff} \quad i \leq k\,, \quad \text{for all } i, k \in I\,.$$

In the second case, the sequences $(d_n)_{n \in I}$ and $(B_n)_{n \in I}$ satisfy

$$d_i \in B_k \quad \text{iff} \quad i > k\,, \quad \text{for all } i, k \in I\,.$$

Shifting the sequence $(d_i)_{i \in I}$ by one, we obtain the desired sequences $(a_i)_{i < \omega}$ and $(B_i)_{i < \omega}$. □

Using these two results, we can present our characterisations. We introduce each in turn, before proving that they are all equivalent to (un-)stability. The first combinatorial characterisation is based on the non-existence of a definable linear order.

Definition 5.7. Let T be a theory. A formula $\varphi(\bar{x}, \bar{y})$ has the *order property* (with respect to T) if there exists a model $\mathfrak{M} \vDash T$ containing two sequences $(\bar{a}^n)_{n<\omega}$ and $(\bar{b}^n)_{n<\omega}$ such that

$$\mathfrak{M} \vDash \varphi(\bar{a}^i, \bar{b}^k) \quad \text{iff} \quad i \leq k\,.$$

Using compactness we obtain several equivalent definitions of the order property.

Lemma 5.8. *Let T be a complete first-order theory and $\varphi(\bar{x}, \bar{y})$ a formula. The following statements are equivalent.*

(1) *φ has the order property with respect to T.*

(2) *For every linear order $\langle I, \leq \rangle$, there exists a model $\mathfrak{M}$ of T that contains sequences $(\bar{a}^i)_{i \in I}$ and $(\bar{b}^i)_{i \in I}$ such that*

$$\mathfrak{M} \vDash \varphi(\bar{a}^i, \bar{b}^k) \quad \text{iff} \quad i \leq k\,.$$

(3) *For every model $\mathfrak{M}$ of T and all finite linear orders $\langle I, \leq \rangle$, there are sequences $(\bar{a}^i)_{i \in I}$ and $(\bar{b}^i)_{i \in I}$ in M such that*

$$\mathfrak{M} \vDash \varphi(\bar{a}^i, \bar{b}^k) \quad \text{iff} \quad i \leq k\,.$$

Proof. (2) ⇒ (1) The claim follows from (2) if we set $I = \omega$.

(3) ⇒ (2) This is a direct application of the Compactness Theorem. Given I, choose new constant symbols $\bar{c}^i$ and $\bar{d}^i$, for $i \in I$, and define

$$\begin{aligned} \Phi := T &\cup \{\, \varphi(\bar{c}^i, \bar{d}^k) \mid i, k \in I,\ i \leq k \,\} \\ &\cup \{\, \neg\varphi(\bar{c}^i, \bar{d}^k) \mid i, k \in I,\ i > k \,\}\,. \end{aligned}$$

Clearly, every model of Φ contains two sequences with the desired properties. Hence, it remains to prove that Φ is satisfiable. By the Compactness Theorem, we only have to show that every finite subset of Φ has a model. Let $\Phi_0 \subseteq \Phi$ be finite. Then there exists a finite subset $I_0 \subseteq I$ such that

$$\begin{aligned} \Phi_0 \subseteq T &\cup \{\, \varphi(\bar{c}^i, \bar{d}^k) \mid i, k \in I_0,\ i \leq k \,\} \\ &\cup \{\, \neg\varphi(\bar{c}^i, \bar{d}^k) \mid i, k \in I_0,\ i > k \,\}\,. \end{aligned}$$

Let $\mathfrak{M}$ be an arbitrary model of T. By (3), we can find sequences $(\bar{a}^i)_{i \in I_0}$ and $(\bar{b}^i)_{i \in I_0}$ in M such that

$$\mathfrak{M} \vDash \varphi(\bar{a}^i, \bar{b}^k) \quad \text{iff} \quad i \leq k\,.$$

Consequently, we can satisfy Φ_0 in the structure $\mathfrak{M}$ if we interpret the constants $\bar{c}^i$ by $\bar{a}^i$ and the constants $\bar{d}^i$ by $\bar{b}^i$.

(1) ⇒ (3) Fix a model $\mathfrak{N}$ of T that contains sequences $(\bar{a}^n)_{n<\omega}$ and $(\bar{b}^n)_{n<\omega}$ such that

$$\mathfrak{N} \vDash \varphi(\bar{a}^i, \bar{b}^k) \quad \text{iff} \quad i \leq k\,.$$

Consider the formula

$$\psi_m := \exists \bar{x}_0 \cdots \exists \bar{x}_{m-1} \exists \bar{y}_0 \cdots \exists \bar{y}_{m-1} \bigwedge_{i<k} [\varphi(\bar{x}_i, \bar{y}_k) \wedge \varphi(\bar{x}_i, \bar{x}_i) \wedge \neg\varphi(\bar{x}_k, \bar{y}_i)]\,.$$

Suppose that $|I| = m < \omega$ and let $\mathfrak{M}$ be an arbitrary model of T. Since $\mathfrak{N} \vDash \varphi_m$ we have $T \vDash \varphi_m$ which, in turn, implies that $\mathfrak{M} \vDash \varphi_m$. Consequently, $\mathfrak{M}$ contains two finite sequences $(\bar{a}^n)_{n<m}$ and $(\bar{b}^n)_{n<m}$ with the desired properties. □

The second combinatorial characterisation is based on the non-existence of certain trees.

Definition 5.9. Let T be a complete first-order theory, $\varphi(\bar{x}; \bar{y})$ a formula, U a set of parameters, and γ an ordinal. A *φ-tree of height γ over U* is a family $(\bar{c}_w)_{w \in 2^{<\gamma}}$ of parameters $\bar{c}_w \subseteq U$ such that, for every $\eta \in 2^\gamma$, the set

$$T(U) \cup \{\, \varphi^{\eta(\alpha)}(\bar{x}; \bar{c}_{\eta \restriction \alpha}) \mid \alpha < \gamma \,\}$$

is consistent, where

$$\varphi^0(\bar{x}; \bar{y}) := \varphi(\bar{x}; \bar{y}) \quad \text{and} \quad \varphi^1(\bar{x}; \bar{y}) := \neg\varphi(\bar{x}; \bar{y})\,.$$

Lemma 5.10. *Let T be a first-order theory and $\varphi(\bar{x};\bar{y})$ a formula such that, for every $n < \omega$, there exists a model of T containing a φ-tree of height n. Then, for every ordinal γ, there exists a model of T containing a φ-tree of height γ.*

Proof. Given γ, set

$$\Phi_\gamma := T \cup \{\, \varphi^{\eta(\alpha)}(\bar{x}_\eta; \bar{y}_{\eta\restriction\alpha}) \mid \alpha < \gamma,\ \eta \in 2^\gamma \,\}.$$

If this set is satisfiable, there exists a model of T containing elements $\bar{a}_\eta$ and $\bar{c}_w$, for $\eta \in 2^\gamma$ and $w \in 2^{<\gamma}$, such that every $\bar{a}_\eta$ satisfies

$$T(\textstyle\bigcup_w \bar{c}_w) \cup \{\, \varphi^{\eta(\alpha)}(\bar{x}; \bar{c}_{\eta\restriction\alpha}) \mid \alpha < \gamma \,\}.$$

Hence, $(\bar{c}_w)_{w\in 2^{<\gamma}}$ is a φ-tree of height γ.

It therefore remains to show that Φ_γ is satisfiable. By the Compactness Theorem, it is sufficient to prove that every finite subset is satisfiable. Hence, consider a finite set $\Psi \subseteq \Phi_\gamma$. Let $\alpha_0 < \dots < \alpha_{n-1}$ be an enumeration of all ordinals α such that Ψ contains a formula of the form $\varphi^{\eta(\alpha)}(\bar{x}_\eta; \bar{y}_{\eta\restriction\alpha})$ and let $\sigma : 2^{\leq\gamma} \to 2^{\leq n}$ be the function mapping a sequence $\eta \in 2^\beta$ of length $\beta \leq \gamma$ to its restriction $\langle \eta(\alpha_0), \dots, \eta(\alpha_k)\rangle$, where $k < n$ is the maximal index such that $\alpha_k < \beta$. By assumption, there exists a φ-tree $(\bar{d}_w)_{w\in 2^{<n}}$ of height n. For each branch $\zeta \in 2^n$, fix a tuple $\bar{a}_\zeta$ satisfying

$$T(\textstyle\bigcup_w d_w) \cup \{\, \varphi^{\zeta(i)}(\bar{x}; \bar{d}_{\zeta\restriction i}) \mid i < n \,\}.$$

Then Ψ is satisfied if we assign the value $\bar{a}_{\sigma(\eta)}$ to the variable $\bar{x}_\eta$ and the value $\bar{d}_{\sigma(w)}$ to the variable $\bar{y}_w$. □

The existence of large φ-trees implies that the local type spaces are also large. In particular, formulae with large φ-trees are unstable.

Lemma 5.11. *Let T be a complete theory and $\varphi(\bar{x};\bar{y})$ a formula such that there are φ-trees of height n, for all $n < \omega$. For every infinite cardinal κ there exists a set U of parameters such that $|S_\varphi(U)| > \kappa = |U|$.*

Proof. Let μ be the minimal cardinal such that $2^\mu > \kappa$. By Lemma 5.10, there exists a φ-tree $(\bar{c}_w)_{w \in 2^{<\mu}}$ of height μ. Since $2^{<\mu} \leq \kappa$, we can choose a set U of size $|U| = \kappa$ containing all parameters $\bar{c}_w$, for $w \in 2^{<\mu}$. For every branch $\eta \in 2^\mu$, fix a tuple $\bar{a}_\eta$ satisfying

$$\{ \varphi^{\eta(\alpha)}(\bar{x}; \bar{c}_{\eta \restriction \alpha}) \mid \alpha < \mu \}.$$

For $\eta \neq \zeta$, it follows that $\operatorname{tp}_\varphi(\bar{a}_\eta/U) \neq \operatorname{tp}_\varphi(\bar{a}_\zeta/U)$. Hence,

$$|S_\varphi(U)| \geq 2^\mu > \kappa = |U|. \qquad \square$$

Before proving the converse, let us present a third, logical characterisation of stability.

Definition 5.12. Let $\mathfrak{M}$ be a structure, $C, U \subseteq M$ sets of parameters, Δ a set of formulae, and $\varphi(\bar{x}; \bar{y})$ a formula.

(a) A *φ-definition* of a type $\mathfrak{p} \in S_\varphi(U)$ over C is a formula $\delta(\bar{y})$ over C such that

$$\varphi(\bar{x}; \bar{c}) \in \mathfrak{p} \quad \text{iff} \quad \mathfrak{M} \vDash \delta(\bar{c}), \quad \text{for all } \bar{c} \subseteq U.$$

(b) A complete type $\mathfrak{p} \in S_\Delta(U)$ is *definable* over C if, for every $\varphi \in \Delta$, the type $\mathfrak{p}|_\varphi$ has a φ-definition over C.

Example. Recall the example on page 529, where we described $S^1(\mathbb{Q})$ for the theory $T := \operatorname{Th}(\langle \mathbb{Q}, < \rangle)$. The definable types are those of the form (a^+), (a^-), $(+\infty)$, $(-\infty)$, and all realised types. The irrational types are not definable. For instance, for (a^+) and $\varphi(x; y) := x < y$, we can use the definition $\delta(y) := y > a$.

The number of definable types is always small.

Lemma 5.13. *Let $\varphi(\bar{x}; \bar{y}) \in \mathrm{FO}[\Sigma, X \cup Y]$. Then $S_\varphi(U)$ contains at most*

$$|\Sigma| \oplus |C| \oplus \aleph_0$$

types that are definable over C.

Proof. W.l.o.g. we may assume that X and Y are finite. Then there are $|\Sigma| \oplus |C| \oplus \aleph_0$ first-order formulae over C and, hence, at most that many φ-definitions. Furthermore, if $\mathfrak{p}, \mathfrak{q} \in S_\varphi(U)$ are types with the same φ-definition then $\mathfrak{p} = \mathfrak{q}$. □

Lemma 5.14. *Let U be a set of parameters and let $\varphi(\bar{x}; \bar{y})$ be a first-order formula that has no φ-tree of height $N < \omega$. Then every φ-type in $S_\varphi(U)$ is definable over U.*

Proof. For a formula $\psi(\bar{x})$ over U, let $D_\varphi(\psi)$ be the maximal number n such that there exists a φ-tree $(\bar{c}_w)_{w \in 2^{<n}}$ of height n such that, for every $\eta \in 2^n$, the set

$$T(U) \cup \{\psi(\bar{x})\} \cup \{\, \varphi^{\eta(i)}(\bar{x}; \bar{c}_{\eta \restriction i}) \mid i < n \,\}$$

is consistent. By assumption, $D_\varphi(\psi) < N$. In particular, the maximum is well-defined. Furthermore, D_φ is monotone in the sense that

$$\psi \vDash \vartheta \quad \text{implies} \quad D_\varphi(\psi) \leq D_\varphi(\vartheta)\,.$$

Given $\mathfrak{p} \in S_\varphi(U)$, choose a finite subset $\Phi \subseteq \mathfrak{p}$ such that $D_\varphi(\bigwedge \Phi)$ is minimal. By choice of Φ and monotonicity of D_φ, it follows for every $\bar{c} \subseteq U$ that

$$\varphi(\bar{x}; \bar{c}) \in \mathfrak{p} \quad \text{iff} \quad D_\varphi\big(\bigwedge \Phi(\bar{x}) \wedge \varphi(\bar{x}; \bar{c})\big) = D_\varphi\big(\bigwedge \Phi(\bar{x})\big)\,.$$

Since the non-existence of a φ-tree of height n with the above property is definable in first-order logic, it follows that, for every $n < \omega$ and every formula $\psi(\bar{x}; \bar{y})$ over U, there is a formula $\delta^n_\psi(\bar{y})$ over U such that

$$\mathfrak{M} \vDash \delta^n_\psi(\bar{c}) \quad \text{iff} \quad D_\varphi(\psi(\bar{x}; \bar{c})) < n\,.$$

Hence, we can use the formula $\delta^n_{\bigwedge \Phi \wedge \varphi}$ with $n := D_\varphi(\bigwedge \Phi) + 1$ to define $\mathfrak{p}$. □

After having introduced three properties of formulae, we can show that they are all equivalent to (un-)stability.

Theorem 5.15. *Let T be a complete first-order theory and $\varphi(\bar{x};\bar{y})$ a formula. The following statements are equivalent:*

(1) *φ is stable.*

(2) *φ is κ-stable, for all infinite cardinals κ.*

(3) *φ does not have the order property.*

(4) *There exists some $n < \omega$ such that there is no φ-tree of height n.*

(5) *Every complete φ-type is definable over its domain.*

Proof. (2) ⇒ (1) is trivial and (1) ⇒ (4) ⇒ (5) ⇒ (1) were already proved in, respectively, Lemmas 5.11, 5.14, and 5.13.

(4) ⇒ (3) Suppose that φ has the order property. Let $\leq$ be the infix ordering on $I := 2^{\leq\omega}$, which is defined by

$$\begin{aligned} u < v \quad &:\text{iff} \quad v = u1x\,, && \text{for some } x \in 2^{\leq\omega}, \\ &\quad\ \text{or } u = w0x \text{ and } v = w1y\,, && \text{for some } w \in 2^{<\omega} \text{ and } \\ & && x, y \in 2^{\leq\omega}. \end{aligned}$$

By Lemma 5.8, we can find a model $\mathfrak{M}$ of T that contains sequences $(\bar{a}_w)_{w\in I}$ and $(\bar{b}_w)_{w\in I}$ such that

$$\mathfrak{M} \vDash \varphi(\bar{a}_u, \bar{b}_v) \quad \text{iff} \quad u \leq v\,.$$

For $\eta \in 2^\omega$ and $n < \omega$, it follows that

$$\mathfrak{M} \vDash \varphi(\bar{a}_\eta, \bar{b}_{\eta\restriction n}) \quad \text{iff} \quad \eta \leq \eta \restriction n \quad \text{iff} \quad \eta(n) = 0\,.$$

Consequently, for every $\eta \in 2^\omega$ and every $n < \omega$, the tuple $\bar{a}_\eta$ satisfies

$$T(U) \cup \{\, \varphi^{\eta(i)}(\bar{x}; \bar{b}_{\eta\restriction i}) \mid i < n \,\}\,,$$

and $(\bar{b}_w)_{w\in 2^{<n}}$ is a φ-tree of height n.

(3) ⇒ (2) Suppose that there is an infinite set U with $|S_\varphi(U)| > |U|$. Fix a model $\mathfrak{M}$ containing realisations of every φ-type over U. Let $\bar{s}$ be

the sorts of those variables in $\bar{x}$ that actually appear in φ and let $\bar{t}$ be those in $\bar{y}$. For $\bar{a} \in M^{\bar{s}}$, we set

$$S(\bar{a}) := \{ \bar{c} \in U^{\bar{t}} \mid \mathfrak{M} \vDash \varphi(\bar{a}; \bar{c}) \} .$$

Note that $\operatorname{tp}_\varphi(\bar{a}/U) \neq \operatorname{tp}_\varphi(\bar{b}/U)$ implies $S(\bar{a}) \neq S(\bar{b})$. Hence,

$$\mathcal{S} := \{ S(\bar{a}) \mid \bar{a} \in M^{\bar{s}} \} \subseteq \wp(U^{\bar{t}})$$

is a family of size $|\mathcal{S}| = |S_\varphi(U)| > |U| = |U^{\bar{t}}|$. By Theorem 5.6, there exist sequences $(\bar{c}_i)_{i<\omega}$ in $U^{\bar{t}}$ and $(\bar{a}_i)_{i<\omega}$ in $M^{\bar{s}}$ such that either

$$\begin{aligned} &\bar{c}_i \in S(\bar{a}_k) \quad \text{iff} \quad i \leq k\,, \\ \text{or} \quad &\bar{c}_i \in S(\bar{a}_k) \quad \text{iff} \quad i \geq k\,. \end{aligned}$$

It follows that

$$\begin{aligned} &\mathfrak{M} \vDash \varphi(\bar{a}_i; \bar{c}_k) \quad \text{iff} \quad i \leq k \\ \text{or} \quad &\mathfrak{M} \vDash \varphi(\bar{a}_i; \bar{c}_k) \quad \text{iff} \quad i \geq k\,. \end{aligned}$$

In the first case, φ has the order property and we are done. In the second case, we can take, for every $n < \omega$, a prefix of length n of these two sequences and reverse their ordering to obtain sequences $(\bar{a}'_i)_{i<n}$ and $(\bar{c}'_i)_{i<n}$ such that

$$\mathfrak{M} \vDash \varphi(\bar{a}'_i; \bar{c}'_k) \quad \text{iff} \quad i \leq k\,.$$

Consequently, it follows by Lemma 5.8 (3) that φ has the order property. ☐

Having characterised stable formulae, we turn to theories and their type spaces.

Definition 5.16. (a) A complete first-order theory T is *κ-stable* if we have $|S^{\bar{s}}(U)| \leq \kappa$, for all finite tuples of sorts $\bar{s}$ and every set U of size

$|U| \leq \kappa$. We call T *stable* if it is κ-stable, for some infinite cardinal κ. Otherwise, T is *unstable.*

(b) A complete first-order theory T is *totally transcendental* if

$$\mathrm{rk}_{\mathrm{CB}}(\mathfrak{S}^{\bar{s}}(U)) < \infty \quad \text{for all sets } U \text{ and all finite tuples } \bar{s}\,.$$

We obtain equivalent characterisations to those of Theorem 5.15.

Theorem 5.17. *Let T be a complete first-order theory. The following statements are equivalent:*

(1) *T is stable.*

(2) *T is κ-stable, for every cardinal κ such that $\kappa^{|T|} = \kappa$.*

(3) *Every first-order formula is stable.*

(4) *Every complete type is definable over its domain.*

(5) $\mathrm{rk}_{\mathrm{CB}}(\mathfrak{S}_\Delta(U)) < \infty$*, for all sets U and all finite sets Δ.*

(6) $|S_\Delta(U)| \leq \kappa$*, for all infinite cardinals κ, all finite sets Δ, and all sets U of size $|U| \leq \kappa$.*

Proof. (2) ⇒ (1) is trivial.

(1) ⇒ (3) Suppose that some formula $\varphi(\bar{x}, \bar{y})$ is not stable. By Theorem 5.15, it follows that, for every infinite cardinal κ, there exists a set U of size $|U| \leq \kappa$ such that

$$\kappa < |S_\varphi(U)| \leq |S^{\bar{s}}(U)|\,,$$

where $\bar{s}$ are the sorts of $\bar{x}$. Consequently, T is not κ-stable, for any $\kappa \geq \aleph_0$.

(3) ⇒ (4) Every type $\mathfrak{p} \in S^{\bar{s}}(U)$ is definable over its domain since, by Theorem 5.15, all of its restrictions $\mathfrak{p}|_\varphi$ are definable.

(4) ⇒ (6) Let Δ be a finite set of formulae and U a set of size $|U| \leq \kappa$. There exists an injective function $S_\Delta(U) \to \prod_{\varphi \in \Delta} S_\varphi(U)$ mapping a Δ-type $\mathfrak{p}$ to the tuple of its restrictions $(\mathfrak{p}|_\varphi)_{\varphi \in \Delta}$. If every type is definable over its domain, it follows by Theorem 5.15 that

$$|S_\Delta(U)| \leq \prod_{\varphi \in \Delta} |S_\varphi(U)| \leq \kappa^{|\Delta|} = \kappa\,.$$

(6) ⇒ (2) Let κ be a cardinal with $\kappa^{|T|} = \kappa$ and let U be a set of size $|U| \leq \kappa$. Since there exists an injective function

$$S^{<\omega}(U) \to \prod_\varphi S_\varphi(U) : \mathfrak{p} \mapsto (\mathfrak{p}|_\varphi)_\varphi ,$$

it follows that

$$|S^{<\omega}(U)| \leq \prod_\varphi |S_\varphi(U)| \leq \kappa^{|T|} = \kappa .$$

(6) ⇒ (5) Suppose that $\mathrm{rk}_{\mathrm{CB}}(\mathfrak{S}_\Delta(U)) = \infty$. We have seen in Corollary 4.6 that

$$\mathfrak{S}_\Delta(U) \cong \mathrm{spec}(\mathfrak{Lb}(\Delta_U^\neg)) .$$

By Lemma B2.5.15, there exists an embedding $(\psi_w)_{w\in 2^{<\omega}}$ of $2^{<\omega}$ into $\mathfrak{Lb}(\Delta_U^\neg)$. Let $U_0 \subseteq U$ be the set of all parameters appearing in these formulae ψ_w. Then U_0 is countable and $(\psi_w)_{w\in 2^{<\omega}}$ is an embedding of $2^{<\omega}$ into $\mathfrak{Lb}(\Delta_{U_0}^\neg)$. Consequently,

$$\mathrm{rk}_{\mathrm{CB}}(\mathfrak{S}_\Delta(U_0)) = \mathrm{rk}_{\mathrm{CB}}(\mathrm{spec}(\mathfrak{Lb}(\Delta_{U_0}^\neg))) = \infty .$$

It follows by Corollary B5.7.4 that $|S_\Delta(U_0)| \geq 2^{\aleph_0} > |U_0|$. This contradicts (6).

(5) ⇒ (6) Suppose that there is some infinite set U with $|S_\Delta(U)| > |U|$. We have seen in Corollary 4.6 that

$$\mathfrak{S}_\Delta(U) \cong \mathrm{spec}(\mathfrak{Lb}(\Delta_U^\neg)) .$$

Consequently, $|\mathrm{spec}(\mathfrak{Lb}(\Delta_U^\neg))| > |\mathfrak{Lb}(\Delta_U^\neg)|$ implies, by Corollary B2.5.22, that

$$\mathrm{rk}_{\mathrm{CB}}(\mathfrak{S}_\Delta(U)) = \mathrm{rk}_{\mathrm{CB}}(\mathrm{spec}(\mathfrak{Lb}(\Delta_U^\neg))) = \infty . \qquad \square$$

$\aleph_0$-stable theories are particularly simple. They are κ-stable, for every cardinal κ, and not only the local type spaces $\mathfrak{S}_\Delta(U)$, but even the full type space $\mathfrak{S}^{<\omega}(U)$ has a Cantor-Bendixson rank.

Theorem 5.18. *Let T be a complete theory over a countable signature. The following statements are equivalent:*

(1) *T is $\aleph_0$-stable.*

(2) *T is κ-stable, for all infinite cardinals κ.*

(3) *T is totally transcendental.*

Proof. By Theorem 2.18, we have

$$\mathfrak{S}^{\bar{s}}(U) \cong \operatorname{spec}(\mathfrak{B}(U)) \quad \text{where} \quad \mathfrak{B}(U) := \mathfrak{Lb}(\mathrm{FO}^{\bar{s}}[\Sigma_U]/T(U)) .$$

(2) ⇒ (1) is trivial.

(3) ⇒ (2) Suppose that there is some infinite cardinal κ such that T is not κ-stable, that is, we have $|S^{\bar{s}}(U)| > |U|$, for some set U of size $|U| = \kappa$. By Corollary B2.5.22 there is some type $\mathfrak{p} \in S^{\bar{s}}(U)$ with $\mathrm{rk}_{\mathrm{P}}(\varphi) = \infty$. Hence, Theorem B5.7.8 implies that $\mathrm{rk}_{\mathrm{CB}}(\mathfrak{S}^{\bar{s}}(U)) = \infty$.

(1) ⇒ (3) Suppose that $\mathrm{rk}_{\mathrm{CB}}(\mathfrak{S}^{\bar{s}}(U)) = \infty$, for some set U and some finite tuple $\bar{s}$. By Theorem B5.7.8, there is some formula $\varphi \in \mathfrak{B}(U)$ with $\mathrm{rk}_{\mathrm{P}}(\varphi) = \infty$. Hence, we an use Lemma B2.5.15 to find an embedding $(\psi_w)_{w \in 2^{<\omega}}$ of $2^{<\omega}$ into $\mathfrak{B}(U)$. Let $U_0 \subseteq U$ be the set of all parameters appearing in these formulae ψ_w. Then U_0 is countable and $(\psi_w)_{w \in 2^{<\omega}}$ is an embedding of $2^{<\omega}$ into $\mathfrak{B}(U_0)$. By Lemma B5.7.3, it follows that

$$|S^{\bar{s}}(U_0)| = |\operatorname{spec}(\mathfrak{B}(U_0))| \geq 2^{\aleph_0} > |U| .$$

Hence, T is not $\aleph_0$-stable. □

C4. *Back-and-forth equivalence*

1. *Partial isomorphisms*

In many constructions and proofs we will have to find two sequences $\bar a$ and $\bar b$ that cannot be told apart by any formula of a given logic, i.e., we are interested in the relation $\langle \mathfrak{A}, \bar a\rangle \equiv_L \langle \mathfrak{B}, \bar b\rangle$. In the present chapter we take a closer look at such relations for $L = \mathrm{FO}_{\infty\aleph_0}$ and $L = \mathrm{FO}$.

Definition 1.1. Let $\mathfrak{A}$ and $\mathfrak{B}$ be Σ-structures, $\bar a \subseteq A$ and $\bar b \subseteq B$ sequences of the same length, and α an ordinal.

(a) We write $\langle \mathfrak{A}, \bar a\rangle \equiv_\alpha \langle \mathfrak{B}, \bar b\rangle$ iff

$$\mathfrak{A} \vDash \varphi(\bar a) \quad \text{iff} \quad \mathfrak{B} \vDash \varphi(\bar b)\,,$$

for all formulae $\varphi \in \mathrm{FO}_{\infty\aleph_0}[\Sigma]$ of quantifier rank $\mathrm{qr}(\varphi) \leq \alpha$. If $\mathfrak{A} \equiv_\alpha \mathfrak{B}$ we say that $\mathfrak{A}$ is *α-equivalent* to $\mathfrak{B}$.

(b) We write $\langle \mathfrak{A}, \bar a\rangle \equiv_\infty \langle \mathfrak{B}, \bar b\rangle$ iff $\langle \mathfrak{A}, \bar a\rangle \equiv_\alpha \langle \mathfrak{B}, \bar b\rangle$, for all ordinals α. Hence, we have

$$\langle \mathfrak{A}, \bar a\rangle \equiv_\infty \langle \mathfrak{B}, \bar b\rangle \quad \text{iff} \quad \langle \mathfrak{A}, \bar a\rangle \equiv_{\mathrm{FO}_{\infty\aleph_0}} \langle \mathfrak{B}, \bar b\rangle\,.$$

The relations $\equiv_\alpha$ can be computed by induction on α. Note that we have $\langle \mathfrak{A}, \bar a\rangle \equiv_0 \langle \mathfrak{B}, \bar b\rangle$ if and only if the function $a_i \mapsto b_i$ induces an isomorphism $\langle\!\langle \bar a\rangle\!\rangle_{\mathfrak{A}} \cong \langle\!\langle \bar b\rangle\!\rangle_{\mathfrak{B}}$.

Definition 1.2. Let $\mathfrak{A}$ and $\mathfrak{B}$ be Σ-structures. A *partial isomorphism* from $\mathfrak{A}$ to $\mathfrak{B}$ is a function p with $\operatorname{dom} p \subseteq A$ and $\operatorname{rng} p \subseteq B$ such that p can be extended to an isomorphism

$$\langle\!\langle \operatorname{dom} p\rangle\!\rangle_{\mathfrak{A}} \cong \langle\!\langle \operatorname{rng} p\rangle\!\rangle_{\mathfrak{B}}\,.$$

We denote the set of all partial isomorphisms from $\mathfrak{A}$ to $\mathfrak{B}$ whose domains have cardinality less than κ by $\mathrm{pIso}_\kappa(\mathfrak{A}, \mathfrak{B})$. The union for all cardinals κ is $\mathrm{pIso}(\mathfrak{A}, \mathfrak{B}) := \bigcup_\kappa \mathrm{pIso}_\kappa(\mathfrak{A}, \mathfrak{B})$.

For sequences $(a_i)_{i<\alpha}$ and $(b_i)_{i<\alpha}$ we simplify notation by writing $p : \bar{a} \mapsto \bar{b}$ for the function $p = \{ \langle a_i, b_i \rangle \mid i < \alpha \}$. (Note that, if we reorder the sequences $\bar{a}$ and $\bar{b}$ then we obtain the same function p.)

Remark. (a) Note that, by Theorem B3.1.9, in the above definition the isomorphism

$$\pi : \langle\!\langle \mathrm{dom}\, p \rangle\!\rangle_{\mathfrak{A}} \to \langle\!\langle \mathrm{rng}\, p \rangle\!\rangle_{\mathfrak{B}}$$

extending p is unique, if it exists.

(b) If Σ is a relational signature then $\langle\!\langle X \rangle\!\rangle_{\mathfrak{A}} = X$ and a function p is a partial isomorphism iff $p : \mathrm{dom}\, p \cong \mathrm{rng}\, p$.

(c) Finally, note that $\langle\rangle \mapsto \langle\rangle = \varnothing$ is the unique function p with $\mathrm{dom}\, p = \varnothing$ and $\mathrm{rng}\, p = \varnothing$. It is a partial isomorphism iff $\langle\!\langle \varnothing \rangle\!\rangle_{\mathfrak{A}} \cong \langle\!\langle \varnothing \rangle\!\rangle_{\mathfrak{B}}$, that is, if the substructures generated by the constants of $\mathfrak{A}$ and $\mathfrak{B}$ are isomorphic and if the same relations of arity 0 hold in $\mathfrak{A}$ and $\mathfrak{B}$.

Definition 1.3. Let $\mathfrak{A}$ and $\mathfrak{B}$ be Σ-structures.

(a) A partial isomorphism $p \in \mathrm{pIso}(\mathfrak{A}, \mathfrak{B})$ has the *back-and-forth property* with respect to a set $I \subseteq \mathrm{pIso}(\mathfrak{A}, \mathfrak{B})$ of partial isomorphisms if the following conditions are satisfied:

Forth. For all $a \in A$, there is some $q \in I$ such that $p \subseteq q$ and $a \in \mathrm{dom}\, q$.

Back. For all $b \in B$, there is some $q \in I$ such that $p \subseteq q$ and $b \in \mathrm{rng}\, q$.

A set $J \subseteq \mathrm{pIso}(\mathfrak{A}, \mathfrak{B})$ of partial isomorphisms has the back-and-forth property with respect to I if every element of J has the back-and-forth property.

(b) A *back-and-forth system* between $\mathfrak{A}$ and $\mathfrak{B}$ is a sequence $(I_\alpha)_\alpha$ of sets $I_\alpha \subseteq \mathrm{pIso}(\mathfrak{A}, \mathfrak{B})$ such that

- for every α, $I_{\alpha+1}$ has the back-and-forth property with respect to I_α, and
- $I_\delta \subseteq \bigcap_{\alpha<\delta} I_\alpha$, for limit ordinals δ.

The *canonical back-and-forth system* $(I_\alpha(\mathfrak{A}, \mathfrak{B}))_\alpha$ between $\mathfrak{A}$ and $\mathfrak{B}$ is defined inductively by

$$\begin{aligned} I_0(\mathfrak{A}, \mathfrak{B}) &:= \mathrm{pIso}(\mathfrak{A}, \mathfrak{B}) , \\ I_{\alpha+1}(\mathfrak{A}, \mathfrak{B}) &:= \{ p \in I_\alpha(\mathfrak{A}, \mathfrak{B}) \mid p \text{ has the back-and-forth} \\ &\qquad\qquad \text{property w.r.t. } I_\alpha(\mathfrak{A}, \mathfrak{B}) \} , \\ \text{and} \quad I_\delta(\mathfrak{A}, \mathfrak{B}) &:= \bigcap_{\alpha<\delta} I_\alpha(\mathfrak{A}, \mathfrak{B}) , \quad \text{for limit ordinals } \delta . \end{aligned}$$

We will also need the restrictions

$$I^\kappa_\alpha(\mathfrak{A}, \mathfrak{B}) := I_\alpha(\mathfrak{A}, \mathfrak{B}) \cap \mathrm{pIso}_\kappa(\mathfrak{A}, \mathfrak{B})$$

to domains of size less than κ.

Example. Let $\mathfrak{A} = \langle \mathbb{Z}, < \rangle$ and $\mathfrak{B} = \langle \mathbb{Q}, < \rangle$. We have

$$\begin{aligned} I_0(\mathfrak{A}, \mathfrak{B}) &= \{ \bar a \mapsto \bar b \mid a_i < a_k \Leftrightarrow b_i < b_k \} , \\ I_1(\mathfrak{A}, \mathfrak{B}) &= \{ \bar a \mapsto \bar b \mid a_i < a_k \Leftrightarrow b_i < b_k \text{ and } |a_i - a_k| \neq 1 \} \\ I_2(\mathfrak{A}, \mathfrak{B}) &= \{ \langle\rangle \mapsto \langle\rangle \} , \\ I_3(\mathfrak{A}, \mathfrak{B}) &= \varnothing . \end{aligned}$$

Recall that an open dense linear order is a linear order without first and last element such that between any two elements there is a third one.

Lemma 1.4. *If* $\mathfrak{A} = \langle A, < \rangle$ *and* $\mathfrak{B} = \langle B, < \rangle$ *are open dense linear orders then* $\mathrm{pIso}_{\aleph_0}(\mathfrak{A}, \mathfrak{B})$ *has the back-and-forth property with respect to itself.*

Proof. Suppose that $\bar a \mapsto \bar b \in \mathrm{pIso}_{\aleph_0}(\mathfrak{A}, \mathfrak{B})$ where w.l.o.g. we may assume that $a_0 \leq \dots \leq a_{n-1}$. By symmetry it is sufficient to prove the forth property. Let $c \in A$. If $c = a_i$, for some i, then $\bar a c \mapsto \bar b b_i$ is a partial isomorphism and we are done. Suppose that there is some i such that $a_i < c < a_{i+1}$. Since $\mathfrak{B}$ is dense we can select an arbitrary element $b_i < d < b_{i+1}$ and the mapping $\bar a c \mapsto \bar b d$ is a partial isomorphism. Similarly, if $c < a_0$ or $c > a_{n-1}$ then we can take any element $d < b_0$ or $d > b_{n-1}$ to obtain a partial isomorphism $\bar a c \mapsto \bar b d$. □

Theorem 1.5 (Cantor). *Any two countable open dense linear orders are isomorphic.*

Proof. Let $\mathfrak{A} = \langle A, < \rangle$ and $\mathfrak{B} = \langle B, < \rangle$ be countable open dense linear orders and fix enumerations $(a_i)_{i<\omega}$ and $(b_i)_{i<\omega}$ of A and B, respectively. Let $I := \mathrm{pIso}_{\aleph_0}(\mathfrak{A}, \mathfrak{B})$. We construct an increasing chain $p_0 \subseteq p_1 \subseteq \dots$ of partial isomorphisms $p_i \in I$ such that $a_i \in \mathrm{dom}\, p_{2i+1}$ and $b_i \in \mathrm{dom}\, p_{2i+2}$. Their union $p := \bigcup_i p_i$ is a partial isomorphism with domain $\mathrm{dom}\, p = A$ and range $\mathrm{rng}\, p = B$, that is, it is the desired total isomorphism $p : \mathfrak{A} \cong \mathfrak{B}$.

We define p_i by induction on i. Let $p_0 := \varnothing$. Suppose that $p_i \in I$ has already been defined and that $i = 2n$ is even. Since I has the forth property with respect to itself we can find some $p_{i+1} \in I$ extending p_i such that $a_n \in \mathrm{dom}\, p_{i+1}$. Similarly, if $i = 2n + 1$ is odd then we use the back property to find a partial isomorphism $p_{i+1} \in I$ extending p_i with $b_n \in \mathrm{rng}\, p_{i+1}$. □

Exercise 1.1. Let $\mathfrak{R} = \langle \mathbb{R}, + \rangle$ be the additive group of real numbers. Show that $\mathrm{pIso}_{\aleph_0}(\mathfrak{R}, \mathfrak{R})$ has the back-and-forth property with respect to itself.

Exercise 1.2. Prove that any two countable atomless boolean algebras are isomorphic.

Remark. (a) The canonical back-and-forth system $(I_\alpha(\mathfrak{A}, \mathfrak{B}))_\alpha$ is maximal, that is, for any back-and-forth system $(I_\alpha)_\alpha$ we have $I_\alpha \subseteq I_\alpha(\mathfrak{A}, \mathfrak{B})$, for all α.

(b) Obviously, a back-and-forth system forms a descending chain

$$I_0 \supseteq I_1 \supseteq \dots \supseteq I_\alpha \supseteq \dots .$$

Furthermore, if there is some ordinal α such that

$$I_\alpha(\mathfrak{A}, \mathfrak{B}) = I_{\alpha+1}(\mathfrak{A}, \mathfrak{B})$$

then $I_\alpha(\mathfrak{A}, \mathfrak{B}) = I_\beta(\mathfrak{A}, \mathfrak{B})$, for all $\beta \geq \alpha$. Hence, there always exists an ordinal $\alpha < |I_0(\mathfrak{A}, \mathfrak{B})|^+$ such that

$$I_\alpha(\mathfrak{A}, \mathfrak{B}) = I_\beta(\mathfrak{A}, \mathfrak{B}), \quad \text{for all } \beta \geq \alpha .$$

Definition 1.6. Let α be the minimal ordinal such that

$$I_\alpha(\mathfrak{A}, \mathfrak{B}) = I_{\alpha+1}(\mathfrak{A}, \mathfrak{B}) .$$

We denote this limit by $I_\infty(\mathfrak{A}, \mathfrak{B}) := I_\alpha(\mathfrak{A}, \mathfrak{B})$ and the corresponding restrictions by $I^\kappa_\infty(\mathfrak{A}, \mathfrak{B}) := I^\kappa_\alpha(\mathfrak{A}, \mathfrak{B})$.

Remark. $I_\infty(\mathfrak{A}, \mathfrak{B})$ has the back-and-forth property with respect to itself.

Exercise 1.3. Let $\mathfrak{A}$ and $\mathfrak{B}$ be finite structures with $|A|, |B| \leq n$. Prove that $I_n(\mathfrak{A}, \mathfrak{B}) = I_\infty(\mathfrak{A}, \mathfrak{B})$.

Lemma 1.7. *If $p \in I_\alpha(\mathfrak{A}, \mathfrak{B})$ and $q \subseteq p$ then $q \in I_\alpha(\mathfrak{A}, \mathfrak{B})$.*

Proof. The claim follows by a straightforward induction on α. □

Corollary 1.8. $I_\alpha(\mathfrak{A}, \mathfrak{B}) \neq \varnothing$ iff $\langle\rangle \mapsto \langle\rangle \in I_\alpha(\mathfrak{A}, \mathfrak{B})$.

Lemma 1.9. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be structures and κ an infinite cardinal. The sequence $(I^\kappa_\alpha(\mathfrak{A}, \mathfrak{B}))_\alpha$ is a back-and-forth system.*

Proof. The claim follows by induction on α since, if

$$\bar{a} \mapsto \bar{b} \in I^\kappa_{\alpha+1}(\mathfrak{A}, \mathfrak{B}) \quad \text{and} \quad \bar{a}c \mapsto \bar{b}d \in I_\alpha(\mathfrak{A}, \mathfrak{B})$$

then the set $\bar{a}c$ has cardinality less than κ. Therefore,

$$\bar{a}c \mapsto \bar{b}d \in \mathrm{pIso}_\kappa(\mathfrak{A}, \mathfrak{B})$$

which implies that $\bar{a}c \mapsto \bar{b}d \in I^\kappa_\alpha(\mathfrak{A}, \mathfrak{B})$. □

Definition 1.10. Let $\mathfrak{A}$ and $\mathfrak{B}$ be Σ-structures, $\bar{a} \subseteq A$, $\bar{b} \subseteq B$, and α an ordinal. We define

$$\langle\mathfrak{A}, \bar{a}\rangle \cong_\alpha \langle\mathfrak{B}, \bar{b}\rangle \quad :\text{iff} \quad \bar{a} \mapsto \bar{b} \in I_\alpha(\mathfrak{A}, \mathfrak{B}) ,$$
$$\langle\mathfrak{A}, \bar{a}\rangle \cong_\infty \langle\mathfrak{B}, \bar{b}\rangle \quad :\text{iff} \quad \bar{a} \mapsto \bar{b} \in I_\infty(\mathfrak{A}, \mathfrak{B}) .$$

If $\mathfrak{A} \cong_\alpha \mathfrak{B}$ we say that $\mathfrak{A}$ is *α-isomorphic* to $\mathfrak{B}$. For an arbitrary back-and-forth system $(I_\beta)_\beta$ we write

$$(I_\beta)_\beta : \langle\mathfrak{A}, \bar{a}\rangle \cong_\alpha \langle\mathfrak{B}, \bar{b}\rangle \quad :\text{iff} \quad \bar{a} \mapsto \bar{b} \subseteq p \,, \text{for some } p \in I_\alpha .$$

Example. Let $\Sigma = \{ P_i \mid i < n \}$ be a signature consisting of n unary predicates. For a Σ-structure $\mathfrak{A} = \langle A, \bar{P} \rangle$ and a set $I \subseteq [n]$, we set

$$P_I^{\mathfrak{A}} := \{ a \in A \mid a \in P_i^{\mathfrak{A}} \text{ iff } i \in I \} .$$

For $k, l, m < \omega$, define

$$k =_m l \quad :\text{iff} \quad k = l \text{ or } k, l \geq m .$$

We claim that $\langle \mathfrak{A}, \bar{a} \rangle \cong_m \langle \mathfrak{B}, \bar{b} \rangle$ if and only if $\bar{a} \mapsto \bar{b} \in \mathrm{pIso}(\mathfrak{A}, \mathfrak{B})$ and

$$|P_I^{\mathfrak{A}} \setminus \bar{a}| =_m |P_I^{\mathfrak{B}} \setminus \bar{b}| , \quad \text{for all } I \subseteq [n] .$$

We prove the claim by induction on m. If $m = 0$ then $\bar{a} \mapsto \bar{b} \in I_0(\mathfrak{A}, \mathfrak{B})$ iff $\bar{a} \mapsto \bar{b}$ is a partial isomorphism. Suppose that $m > 0$.

For one direction, assume that there is some I such that

$$|P_I^{\mathfrak{A}} \setminus \bar{a}| \neq_m |P_I^{\mathfrak{B}} \setminus \bar{b}| .$$

By symmetry we may assume that $|P_I^{\mathfrak{A}} \setminus \bar{a}| > |P_I^{\mathfrak{B}} \setminus \bar{b}|$. If $c \in P_I^{\mathfrak{A}} \setminus \bar{a}$ then we have

$$|P_I^{\mathfrak{A}} \setminus \bar{a}c| \neq_{m-1} |P_I^{\mathfrak{B}} \setminus \bar{b}d| , \qquad \text{for every } d \in P_I^{\mathfrak{B}} \setminus \bar{b} .$$

By inductive hypothesis it follows that $\bar{a}c \mapsto \bar{b}d \notin I_{m-1}(\mathfrak{A}, \mathfrak{B})$, for all $d \in B$. Consequently, $\bar{a} \mapsto \bar{b} \notin I_m(\mathfrak{A}, \mathfrak{B})$.

For the other direction, let $\bar{a} \mapsto \bar{b}$ be a partial isomorphism such that

$$|P_I^{\mathfrak{A}} \setminus \bar{a}| =_m |P_I^{\mathfrak{B}} \setminus \bar{b}| , \quad \text{for all } I \subseteq [n] ,$$

and let $c \in A \setminus \bar{a}$. Set $I := \{ i < n \mid c \in P_i^{\mathfrak{A}} \}$ and choose an arbitrary element $d \in P_I^{\mathfrak{B}} \setminus \bar{b}$. It follows that

$$|P_I^{\mathfrak{A}} \setminus \bar{a}c| =_{m-1} |P_I^{\mathfrak{B}} \setminus \bar{b}d| .$$

By inductive hypothesis, this implies that $\bar{a}c \mapsto \bar{b}d \in I_{m-1}(\mathfrak{A}, \mathfrak{B})$, as desired. The back property follows by symmetry.

We will show below that the relations $\cong_\alpha$ and $\equiv_\alpha$ coincide. Hence, we can determine whether $\mathfrak{A} \equiv_\alpha \mathfrak{B}$ holds by defining a back-and-forth system $(I_\beta)_\beta : \mathfrak{A} \cong_\alpha \mathfrak{B}$ with $I_\alpha \neq \emptyset$.

Lemma 1.11. *We have* $\mathfrak{A} \cong_\infty \mathfrak{B}$ *if and only if there exists a nonempty set* $I \subseteq \mathrm{pIso}_{\aleph_0}(\mathfrak{A}, \mathfrak{B})$ *that has the back-and-forth property with respect to itself.*

Proof. ($\Rightarrow$) By Lemma 1.9, we can set $I := I_\infty^{\aleph_0}(\mathfrak{A}, \mathfrak{B})$.

($\Leftarrow$) We prove by induction on α that $I \subseteq I_\alpha(\mathfrak{A}, \mathfrak{B})$, for all α. Then we have $I \subseteq I_\infty(\mathfrak{A}, \mathfrak{B})$ which implies that $I_\infty(\mathfrak{A}, \mathfrak{B}) \neq \emptyset$.

Clearly, $I \subseteq \mathrm{pIso}(\mathfrak{A}, \mathfrak{B}) = I_0(\mathfrak{A}, \mathfrak{B})$. Suppose that $I \subseteq I_\alpha(\mathfrak{A}, \mathfrak{B})$. Each $p \in I$ has the back-and-forth property with respect to I and, therefore, also with respect to $I_\alpha(\mathfrak{A}, \mathfrak{B}) \supseteq I$. Hence, $p \in I_{\alpha+1}(\mathfrak{A}, \mathfrak{B})$. Finally, if δ is a limit ordinal and $I \subseteq I_\alpha(\mathfrak{A}, \mathfrak{B})$, for all $\alpha < \delta$, then

$$I \subseteq \bigcap_{\alpha<\delta} I_\alpha(\mathfrak{A}, \mathfrak{B}) = I_\delta(\mathfrak{A}, \mathfrak{B}) .$$

□

As an application we consider discrete linear orders.

Definition 1.12. Let $\mathfrak{A} = \langle A, \leq \rangle$ be a linear order.

(a) $\mathfrak{A}$ is *discrete* if every element of $\mathfrak{A}$ that is not the least one has an immediate predecessor, and every element that is not the greatest one has an immediate successor. We say that $\mathfrak{A}$ is *bounded* if it has a least and a greatest element.

(b) We define the *distance* $d(a, b)$ of two elements $a, b \in A$ by

$$d(a, b) := |\{ c \in A \mid a \leq c < b \text{ or } b \leq c < a \}| .$$

Furthermore, we set

$$d(-\infty, b) := |{\Downarrow} b| ,$$
$$d(a, \infty) := |{\Uparrow} a| ,$$
and $$d(-\infty, \infty) := |A| \oplus 1 .$$

(c) For numbers $m, n, k < \omega$, we define

$$m =_k n \quad :\text{iff} \quad m = n \quad \text{or} \quad m, n \geq k .$$

Lemma 1.13. *Let $\mathfrak{A} = \langle A, \leq \rangle$ and $\mathfrak{B} = \langle B, \leq \rangle$ be bounded discrete linear orders, $\bar{a} \in A^m$ and $\bar{b} \in B^m$, and $n < \omega$. We have*

$$\langle \mathfrak{A}, \bar{a} \rangle \cong_n \langle \mathfrak{B}, \bar{b} \rangle$$

if and only if $\bar{a} \mapsto \bar{b}$ is a partial isomorphism such that, for all i, k,

$$d(a_i, a_k) =_{2^n} d(b_i, b_k), \qquad d(a_i, \infty) =_{2^n} d(b_i, \infty),$$
$$d(-\infty, a_k) =_{2^n} d(-\infty, b_k), \quad d(-\infty, \infty) =_{2^n} d(-\infty, \infty).$$

Proof. ($\Rightarrow$) We prove the claim by induction on n. Let $m := |\bar{a}|$. To avoid case distinctions we add new least and greatest elements $-\infty$ and ∞ to $\mathfrak{A}$ and $\mathfrak{B}$ and we set $a_{-1} := -\infty$ and $a_m := \infty$, and similarly for b_{-1} and b_m.

For $n = 0$, we have

$$\langle \mathfrak{A}, \bar{a} \rangle \cong_0 \langle \mathfrak{B}, \bar{b} \rangle \quad \text{iff} \quad \bar{a} \mapsto \bar{b} \in \mathrm{pIso}(\mathfrak{A}, \mathfrak{B}) .$$

Note that every partial automorphism trivially satisfies the condition $d(a_i, a_k) =_1 d(b_i, b_k)$.

Consider the case that $n > 0$ and suppose that $\langle \mathfrak{A}, \bar{a} \rangle \cong_n \langle \mathfrak{B}, \bar{b} \rangle$. Clearly, the first condition is satisfied since $\bar{a} \mapsto \bar{b}$ is a partial isomorphism. Therefore, it remains to show that

$$d(a_i, a_k) =_{2^n} d(b_i, b_k), \quad \text{for all } -1 \leq i, k \leq m .$$

For a contradiction, suppose that there are i and k such that

$$d(a_i, a_k) \neq_{2^n} d(b_i, b_k) .$$

By symmetry we may assume that $a_i < a_k$ and $d(a_i, a_k) < d(b_i, b_k)$. In particular, we have $d(a_i, a_k) < 2^n$. Furthermore, by inductive hypothesis, we have

$$d(a_i, a_k) =_{2^{n-1}} d(b_i, b_k),$$

which is only possible if $d(a_i, a_k) \geq 2^{n-1}$. Hence, there exists some element $b_i < d \leq b_k$ with $d(b_i, d) = 2^{n-1}$. By the back-and-forth property, we can find an element $c \in A$ such that

$$\langle \mathfrak{A}, \bar{a}c \rangle \cong_{n-1} \langle \mathfrak{B}, \bar{b}d \rangle .$$

By inductive hypothesis, we have $d(a_i, c) =_{2^{n-1}} d(b_i, d)$ which implies that $d(a_i, c) \geq 2^{n-1} = d(b_i, d)$. Consequently, we have

$$d(c, a_k) = d(a_i, a_k) - d(a_i, c) \leq 2^n - 1 - 2^{n-1} = 2^{n-1} - 1$$

which implies that $d(c, a_k) = d(d, b_k)$. Together, it follows that that

$$\begin{aligned} d(a_i, a_k) &= d(a_i, c) + d(c, a_k) \\ &\geq d(b_i, d) + d(d, b_k) = d(b_i, b_k) . \end{aligned}$$

A contradiction.

($\Leftarrow$) Let I_n be the set of all partial functions $\bar{a} \mapsto \bar{b}$ where the tuples $\bar{a}$ and $\bar{b}$ satisfy the above conditions. We claim that $(I_n)_{n<\omega}$ is a back-and-forth system. Clearly, every $\bar{a} \mapsto \bar{b} \in I_0$ is a partial isomorphism. It remains to check the back-and-forth property. By symmetry, we only need to prove one direction. Let $\bar{a} \mapsto \bar{b} \in I_n$ and $c \in A$. Fix indices i and k such that $a_i \leq c \leq a_k$ and there is no index l with $a_i < a_l < a_k$.

We distinguish three cases. If $d(a_i, c) < 2^{n-1}$ then let $d \in B$ be the element such that $b_i \leq d \leq b_k$ and $d(b_i, d) = d(a_i, c)$. If $d(a_i, a_k) = d(b_i, b_k)$ then we clearly have $d(c, a_k) = d(d, b_k)$. If, on the other hand, $d(a_i, a_k), d(b_i, b_k) \geq 2^n$ then $d(c, a_k) \geq 2^{n-1}$ and $d(d, b_k) \geq 2^{n-1}$. Hence, in both cases we have $d(d, b_k) =_{2^{n-1}} d(c, a_k)$.

Similarly, if $d(a_i, c) \geq 2^{n-1}$ but $d(c, a_k) < 2^{n-1}$ then we choose $d \in B$ such that $b_i \leq d \leq b_k$ and $d(d, b_k) = d(c, a_k)$. As above it follows that $d(a_i, c) =_{2^{n-1}} d(b_i, d)$.

Finally, suppose that $d(a_i, c), d(c, a_k) \geq 2^{n-1}$. Then we select an element $b_i < d < b_k$ such that $d(b_i, d) = 2^{n-1}$. Since $d(a_i, a_k), d(b_i, b_k) \geq 2^n$ it follows that $d(d, b_k) = d(b_i, b_k) - d(b_i, c) \geq 2^{n-1}$. □

Corollary 1.14. *For discrete linear orders $\mathfrak{A}$ and $\mathfrak{B}$ and $n < \omega$, we have*

$$\mathfrak{A} \cong_n \mathfrak{B} \quad \text{iff} \quad |A| =_{2^n-1} |B|\,.$$

Lemma 1.15. *Let $\mathfrak{A}_i = \langle A_i, <, \bar{P}\rangle$ and $\mathfrak{B}_i = \langle B_i, <, \bar{P}\rangle$, for $i \in [2]$, be linear orders expanded by unary predicates $\bar{P}$.*

$$\mathfrak{A}_0 \cong_\alpha \mathfrak{B}_0 \quad \textit{and} \quad \mathfrak{A}_1 \cong_\alpha \mathfrak{B}_1 \quad \textit{implies} \quad \mathfrak{A}_0 + \mathfrak{A}_1 \cong_\alpha \mathfrak{B}_0 + \mathfrak{B}_1\,.$$

Proof. Fix back-and-forth systems $(I^i_\beta)_{\beta\leq\alpha} : \mathfrak{A}_i \cong_\alpha \mathfrak{B}_i$. We claim that

$$(J_\beta)_{\beta\leq\alpha} : \mathfrak{A}_0 + \mathfrak{A}_1 \cong_\alpha \mathfrak{B}_0 + \mathfrak{B}_1$$

where

$$J_\beta := \{\, \bar{a}\bar{c} \mapsto \bar{b}\bar{d} \mid \bar{a} \mapsto \bar{b} \in I^0_\beta \text{ and } \bar{c} \mapsto \bar{d} \in I^1_\beta \,\}\,.$$

We have $J_\alpha \neq \varnothing$ since $I^i_\alpha \neq \varnothing$, for both i. Furthermore, $J_\delta = \bigcap_{\beta<\delta} J_\beta$, for limit ordinals δ. It remains to prove the back-and-forth property. Suppose that $\bar{a}\bar{c} \mapsto \bar{b}\bar{d} \in J_{\beta+1}$ and $e \in A$. If $e \in A_0$ then there is some $f \in B_0$ with $\bar{a}e \mapsto \bar{b}f \in I^0_\beta$. Hence, it follows that $\bar{a}e\bar{c} \mapsto \bar{b}f\bar{d} \in J_\beta$. If $e \in A_1$ then the same argument provides a suitable element $f \in B_1$. The back property follows analogously. □

2. Hintikka formulae

The relations $\cong_\alpha$ are definable in $\mathrm{FO}_{\infty\aleph_0}$ by a formula of quantifier rank α. Consequently, we have $\equiv_\alpha \subseteq \cong_\alpha$. The other inclusion will be shown in Section 3.

Lemma 2.1. *Let $\mathfrak{A}$ be a Σ-structure, $\bar{a} \subseteq A$, and α an ordinal. There exists a formula $\varphi^\alpha_{\mathfrak{A},\bar{a}}(\bar{x}) \in \mathrm{FO}_{\infty\aleph_0}[\Sigma]$ of quantifier rank $\mathrm{qr}(\varphi^\alpha_{\mathfrak{A},\bar{a}}) = \alpha$ such that*

$$\mathfrak{B} \vDash \varphi^\alpha_{\mathfrak{A},\bar{a}}(\bar{b}) \quad \text{iff} \quad \bar{a} \mapsto \bar{b} \in I_\alpha(\mathfrak{A}, \mathfrak{B})\,,$$

for all Σ-structures $\mathfrak{B}$ and every $\bar{b} \subseteq B$.

Proof. We construct $\varphi^{\alpha}_{\mathfrak{A},\bar{a}}$ by induction on α.

($\alpha = 0$) Let Φ be the set of all literals $\psi(\bar{x})$ such that $\mathfrak{A} \vDash \psi(\bar{a})$. We set $\varphi^{0}_{\mathfrak{A},\bar{a}} := \bigwedge \Phi$.

($\alpha = \beta + 1$) We have to express the back-and-forth property.

$$\varphi^{\beta+1}_{\mathfrak{A},\bar{a}}(\bar{x}) := \varphi^{\beta}_{\mathfrak{A},\bar{a}}(\bar{x}) \wedge \bigwedge_{c \in A} \exists y \varphi^{\beta}_{\mathfrak{A},\bar{a}c}(\bar{x}y) \wedge \forall y \bigvee_{c \in A} \varphi^{\beta}_{\mathfrak{A},\bar{a}c}(\bar{x}y) .$$

(α limit) We take the conjunction over all $\beta < \alpha$.

$$\varphi^{\alpha}_{\mathfrak{A},\bar{a}}(\bar{x}) := \bigwedge_{\beta < \alpha} \varphi^{\beta}_{\mathfrak{A},\bar{a}}(\bar{x}) .$$

□

Remark. Formulae of the form $\varphi^{\alpha}_{\mathfrak{A},\bar{a}}$ are called *Hintikka formulae.* Note that $\varphi^{\alpha}_{\mathfrak{A},\bar{a}} \in \mathrm{FO}_{\kappa^+\aleph_0}[\Sigma]$ where $\kappa := |A| \oplus |\Sigma| \oplus |\alpha| \oplus \aleph_0$. If Σ, $\bar{a}$, and α are finite then it follows by induction on α that there are only finitely many formulae of the form $\varphi^{\alpha}_{\mathfrak{A},\bar{a}}$ and that we can choose them to be in $\mathrm{FO}^{<\omega}[\Sigma]$.

Since $\cong_\infty = \cong_\alpha$, for some ordinal α, we can also define the relation $\cong_\infty$.

Definition 2.2. Let $\mathfrak{A}$ be a structure. The *Scott height* of $\mathfrak{A}$ is the least ordinal α such that $I^{\aleph_0}_{\infty}(\mathfrak{A},\mathfrak{A}) = I^{\aleph_0}_{\alpha}(\mathfrak{A},\mathfrak{A})$. The *Scott sentence* $\varphi^{\infty}_{\mathfrak{A}}$ of $\mathfrak{A}$ is defined by

$$\varphi^{\infty}_{\mathfrak{A}} := \varphi^{\alpha}_{\mathfrak{A},\langle\rangle} \wedge \bigwedge_{\bar{a} \in A^{<\omega}} \forall \bar{x} \big[\varphi^{\alpha}_{\mathfrak{A},\bar{a}}(\bar{x}) \rightarrow \varphi^{\alpha+1}_{\mathfrak{A},\bar{a}}(\bar{x}) \big] ,$$

where α is the Scott height of $\mathfrak{A}$.

Lemma 2.3. *The Scott height of $\mathfrak{A}$ is less than $|A|^+$.*

Proof. If A is finite then $I_{|A|}(\mathfrak{A},\mathfrak{A}) = I_\infty(\mathfrak{A},\mathfrak{A})$ and the Scott height is at most $|A| < \aleph_0$. Similarly, if A is infinite then there exists some ordinal

$$\alpha < |I^{\aleph_0}_{0}(\mathfrak{A},\mathfrak{A})|^+ \leq (|A|^{<\aleph_0})^+ = |A|^+$$

such that $I^{\aleph_0}_{\alpha}(\mathfrak{A},\mathfrak{A}) = I^{\aleph_0}_{\infty}(\mathfrak{A},\mathfrak{A})$. □

Exercise 2.1. Compute the Scott height of $\langle \omega, \leq \rangle$.

Theorem 2.4. *For all structures* $\mathfrak{A}$ *and* $\mathfrak{B}$*, we have*

$$\mathfrak{B} \vDash \varphi^\infty_{\mathfrak{A}} \quad \text{iff} \quad \mathfrak{B} \cong_\infty \mathfrak{A} .$$

Proof. Let α be the Scott height of $\mathfrak{A}$.

($\Rightarrow$) If $\mathfrak{B} \vDash \varphi^\infty_{\mathfrak{A}}$ then $\bar{a} \mapsto \bar{b} \in I^{\aleph_0}_\alpha(\mathfrak{A}, \mathfrak{B})$ implies $\bar{a} \mapsto \bar{b} \in I^{\aleph_0}_{\alpha+1}(\mathfrak{A}, \mathfrak{B})$. Hence,

$$I^{\aleph_0}_{\alpha+1}(\mathfrak{A}, \mathfrak{B}) = I^{\aleph_0}_\alpha(\mathfrak{A}, \mathfrak{B}) = I^{\aleph_0}_\infty(\mathfrak{A}, \mathfrak{B}) .$$

Furthermore, $I^{\aleph_0}_\infty(\mathfrak{A}, \mathfrak{B})$ is not empty since $\mathfrak{B} \vDash \varphi^\alpha_{\mathfrak{A},\langle\rangle}$ implies $\langle\rangle \mapsto \langle\rangle \in I_\alpha(\mathfrak{A}, \mathfrak{B})$.

($\Leftarrow$) Suppose that $\mathfrak{B} \cong_\infty \mathfrak{A}$. Then we have $\mathfrak{B} \vDash \varphi^\alpha_{\mathfrak{A},\langle\rangle}$. To see that $\mathfrak{B}$ also satisfies the second part of the formula $\varphi^\infty_{\mathfrak{A}}$ we have to show that

$$\bar{a} \mapsto \bar{b} \in I^{\aleph_0}_\alpha(\mathfrak{A}, \mathfrak{B}) \quad \text{implies} \quad \bar{a} \mapsto \bar{b} \in I^{\aleph_0}_{\alpha+1}(\mathfrak{A}, \mathfrak{B}) .$$

Let $\bar{a} \mapsto \bar{b} \in I^{\aleph_0}_\alpha(\mathfrak{A}, \mathfrak{B})$. We claim that $\bar{a} \mapsto \bar{b}$ has the back-and-forth property with respect to $I^{\aleph_0}_\alpha(\mathfrak{A}, \mathfrak{B})$.

For the forth property let $c \in A$. Since $\mathfrak{A} \cong_\infty \mathfrak{B}$ there exist some tuple $\bar{b}' \subseteq A$ with $\langle \mathfrak{A}, \bar{b}' \rangle \cong_\infty \langle \mathfrak{B}, \bar{b} \rangle$. Hence,

$$\langle \mathfrak{A}, \bar{b}' \rangle \cong_\alpha \langle \mathfrak{B}, \bar{b} \rangle \cong_\alpha \langle \mathfrak{A}, \bar{a} \rangle .$$

Since α is the Scott height of $\mathfrak{A}$ is follows that

$$\langle \mathfrak{A}, \bar{b}' \rangle \cong_{\alpha+1} \langle \mathfrak{A}, \bar{a} \rangle .$$

Hence, we can find some $d' \in A$ with

$$\langle \mathfrak{A}, \bar{b}' d' \rangle \cong_\alpha \langle \mathfrak{A}, \bar{a} c \rangle .$$

Since $\langle \mathfrak{A}, \bar{b}' \rangle \cong_\infty \langle \mathfrak{B}, \bar{b} \rangle$ there is some $d \in B$ such that

$$\langle \mathfrak{A}, \bar{b}' d' \rangle \cong_\infty \langle \mathfrak{B}, \bar{b} d \rangle .$$

Consequently, we have

$$\langle \mathfrak{A}, \bar{a}c \rangle \cong_\alpha \langle \mathfrak{A}, \bar{b}'d' \rangle \cong_\alpha \langle \mathfrak{B}, \bar{b}d \rangle ,$$

and $\bar{a}c \mapsto \bar{b}d \in I_\alpha(\mathfrak{A}, \mathfrak{B})$. The back property follows analogously. □

Corollary 2.5. $\mathfrak{A} \equiv_{|A|^+} \mathfrak{B}$ *implies* $\mathfrak{A} \cong_\infty \mathfrak{B}$.

Proof. If α is the Scott height of $\mathfrak{A}$ then $\mathrm{qr}(\varphi^\infty_{\mathfrak{A}}) \leq \alpha + \omega < |A|^+$. □

3. Ehrenfeucht-Fraïssé games

Ehrenfeucht-Fraïssé games provide an intuitive way of describing back-and-forth systems.

Definition 3.1. Let $\mathfrak{A}$ and $\mathfrak{B}$ Σ-structures, $\bar{a}_0 \subseteq A$, $\bar{b}_0 \subseteq B$, and let α be an ordinal.

(a) The *Ehrenfeucht-Fraïssé game* $\mathrm{EF}_\alpha(\mathfrak{A}, \bar{a}_0, \mathfrak{B}, \bar{b}_0)$ is played by two players (*spoiler* and *duplicator*) according to the following rules:

- A *position* in the game is a tuple $\langle \beta, \bar{a}, \bar{b} \rangle$ where $\beta \leq \alpha$, $\bar{a} \subseteq A$, $\bar{b} \subseteq B$, and $|\bar{a}| = |\bar{b}|$.
- The *initial position* is $\langle \alpha, \bar{a}_0, \bar{b}_0 \rangle$.
- In the position $\langle \beta, \bar{a}, \bar{b} \rangle$ spoiler chooses an ordinal $\gamma < \beta$ and either an element $c \in A$ or some $d \in B$. Duplicator responds by selecting an element of the other structure, i.e., either $d \in B$ or $c \in A$. The new position is $\langle \gamma, \bar{a}c, \bar{b}d \rangle$.
- Spoiler loses if he cannot choose γ because $\beta = 0$. Duplicator loses if a position $\langle \beta, \bar{a}, \bar{b} \rangle$ is reached where $\bar{a} \mapsto \bar{b} \notin \mathrm{pIso}(\mathfrak{A}, \mathfrak{B})$.

(b) The *infinite* version $\mathrm{EF}^\kappa_\infty(\mathfrak{A}, \bar{a}_0, \mathfrak{B}, \bar{b}_0)$ of the Ehrenfeucht-Fraïssé game is played in the same way as $\mathrm{EF}_\alpha(\mathfrak{A}, \bar{a}_0, \mathfrak{B}, \bar{b}_0)$ with the exception that the first component of all positions is omitted and every play has length κ. Hence, duplicator wins if she can continue the game for κ steps while, as before, spoiler wins if a position $\langle \bar{a}, \bar{b} \rangle$ is reached such that $\bar{a} \mapsto \bar{b}$ is not a partial isomorphism.

(c) A *winning strategy* of one of the players is a function mapping positions to moves such that, regardless of the moves of his opponent, the player wins if he always plays the moves given by the strategy. We say that a player *wins* the game $\mathrm{EF}_\alpha(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$ if he has a winning strategy.

Example. Let $\mathfrak{A} = \langle \mathbb{Z}, < \rangle$ and $\mathfrak{B} = \langle \mathbb{Q}, < \rangle$. Spoiler wins the 3 round game $\mathrm{EF}_3(\mathfrak{A}, \mathfrak{B})$. The game starts in position

$$\langle 3,\ \langle \rangle,\ \langle \rangle \rangle\,.$$

In the first round, spoiler chooses $2 < 3$ and $0 \in \mathbb{Z}$. Duplicator has to answer with some number $a \in \mathbb{Q}$. The new position is

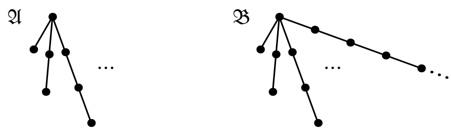
$$\langle 2,\ \langle 0 \rangle,\ \langle a \rangle \rangle\,.$$

In the second round, spoiler chooses $1 < 2$ and $1 \in \mathbb{Z}$. Duplicator replies with some $b \in \mathbb{Q}$ such that $b > a$. The new position is

$$\langle 1,\ \langle 0, 1 \rangle,\ \langle a, b \rangle \rangle\,.$$

Finally, spoiler chooses $0 < 1$ and $(a + b)/2 \in \mathbb{Q}$. Duplicator has to respond with some element $z \in \mathbb{Z}$ such that $0 < z < 1$. Since there is no such element she loses.

Exercise 3.1. Let $\mathfrak{A}$ be the tree consisting of one path of length n, for every $n < \omega$, and let $\mathfrak{B}$ be the tree consisting of one path of length α, for every $\alpha \leq \omega$.



Find the least ordinal α such that Spoiler wins $\mathrm{EF}_\alpha(\mathfrak{A}, \mathfrak{B})$.

Immediately from the definition we obtain the following connection between Ehrenfeucht-Fraïssé games and the back-and-forth property.

Lemma 3.2. *Duplicator wins* $\mathrm{EF}_\alpha(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$ *if and only if*

- *for all* $\beta < \alpha$ *and every* $c \in A$ *there is some* $d \in B$ *such that she wins* $\mathrm{EF}_\beta(\mathfrak{A}, \bar{a}c, \mathfrak{B}, \bar{b}d)$, *and*
- *for all* $\beta < \alpha$ *and every* $d \in B$ *there is some* $c \in A$ *such that she wins* $\mathrm{EF}_\beta(\mathfrak{A}, \bar{a}c, \mathfrak{B}, \bar{b}d)$.

By induction it follows that the winning positions in the game form a back-and-forth system.

Lemma 3.3. *Duplicator wins* $\mathrm{EF}_\alpha(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$ *iff* $\bar{a} \mapsto \bar{b} \in I_\alpha(\mathfrak{A}, \mathfrak{B})$.

Proof. We show the claim by induction on α.

($\alpha = 0$) By definition, duplicator wins $\mathrm{EF}_0(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$ iff $\bar{a} \mapsto \bar{b} \in \mathrm{pIso}(\mathfrak{A}, \mathfrak{B}) = I_0(\mathfrak{A}, \mathfrak{B})$.

($\alpha = \beta + 1$) Duplicator wins $\mathrm{EF}_{\beta+1}(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$

iff for all $c \in A$ there is $d \in B$ such that she wins $\mathrm{EF}_\beta(\mathfrak{A}, \bar{a}c, \mathfrak{B}, \bar{b}d)$
and for all $d \in B$ there is $c \in A$ such that she wins $\mathrm{EF}_\beta(\mathfrak{A}, \bar{a}c, \mathfrak{B}, \bar{b}d)$
iff for all $c \in A$ there is $d \in B$ such that $\bar{a}c \mapsto \bar{b}d \in I_\beta(\mathfrak{A}, \mathfrak{B})$
and for all $d \in B$ there is $c \in A$ such that $\bar{a}c \mapsto \bar{b}d \in I_\beta(\mathfrak{A}, \mathfrak{B})$
iff $\bar{a} \mapsto \bar{b}$ has the back-and-forth property w.r.t $I_\beta(\mathfrak{A}, \mathfrak{B})$
iff $\bar{a} \mapsto \bar{b} \in I_{\beta+1}(\mathfrak{A}, \mathfrak{B})$.

(α limit) Duplicator wins $\mathrm{EF}_\alpha(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$

iff she wins $\mathrm{EF}_\beta(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$ for all $\beta < \alpha$
iff $\bar{a} \mapsto \bar{b} \in I_\beta(\mathfrak{A}, \mathfrak{B})$ for all $\beta < \alpha$
iff $\bar{a} \mapsto \bar{b} \in I_\alpha(\mathfrak{A}, \mathfrak{B})$. □

We have seen that the relation $\equiv_\alpha$ refines $\cong_\alpha$. The following lemma establishes the converse.

Lemma 3.4. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be structures with elements $\bar{a} \subseteq A$ and $\bar{b} \subseteq B$. If there exists a formula $\varphi(\bar{x}) \in \mathrm{FO}_{\infty\aleph_0}[\Sigma, X]$ of quantifier rank $\mathrm{qr}(\varphi) \leq \alpha$ such that*

$$\mathfrak{A} \vDash \varphi(\bar{a}) \quad \textit{and} \quad \mathfrak{B} \nvDash \varphi(\bar{b})$$

then spoiler wins $\mathrm{EF}_\alpha(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$.

Proof. W.l.o.g. we may assume that φ is in negation normal form. We prove the claim by induction on φ.

(φ literal) As $\bar{a}$ and $\bar{b}$ are distinguished by an atomic formula the mapping $\bar{a} \mapsto \bar{b}$ cannot be a partial isomorphism. Hence, spoiler wins the game $\mathrm{EF}_0(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$ immediately.

($\varphi = \bigwedge \Phi$) There is some formula $\psi \in \Phi$ such that

$$\mathfrak{A} \vDash \psi(\bar{a}) \quad \text{and} \quad \mathfrak{B} \nvDash \psi(\bar{b})\,.$$

Since $\mathrm{qr}(\psi) \leq \alpha$ spoiler wins $\mathrm{EF}_\alpha(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$, by inductive hypothesis.

($\varphi = \bigvee \Phi$) follows in the same way.

($\varphi = \exists x \psi$) Let $\beta := \mathrm{qr}(\psi) < \alpha$. There is some element $c \in A$ such that $\mathfrak{A} \vDash \psi(\bar{a}, c)$, but $\mathfrak{B} \nvDash \psi(\bar{b}, d)$, for all $d \in B$. In the first move spoiler can choose β and the element $c \in A$. Duplicator responds with some element $d \in B$. By inductive hypothesis, spoiler can win the resulting game $\mathrm{EF}_\beta(\mathfrak{A}, \bar{a}c, \mathfrak{B}, \bar{b}d)$. Therefore, he also wins $\mathrm{EF}_\alpha(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$.

($\varphi = \forall x \psi$) analogously by choosing some $d \in B$. □

Theorem 3.5 (Karp). *Let $\mathfrak{A}$ and $\mathfrak{B}$ be structures and α an ordinal.*

(a) *The following statements are equivalent:*

(1) $\langle \mathfrak{A}, \bar{a} \rangle \equiv_\alpha \langle \mathfrak{B}, \bar{b} \rangle$.

(2) $\langle \mathfrak{A}, \bar{a} \rangle \cong_\alpha \langle \mathfrak{B}, \bar{b} \rangle$.

(3) $\langle \mathfrak{B}, \bar{b} \rangle \vDash \varphi^\alpha_{\mathfrak{A},\bar{a}}$.

(4) *Duplicator wins* $\mathrm{EF}_\alpha(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$.

(b) *The following statements are equivalent:*

(1) $\langle \mathfrak{A}, \bar{a} \rangle \equiv_\infty \langle \mathfrak{B}, \bar{b} \rangle$.
(2) $\langle \mathfrak{A}, \bar{a} \rangle \cong_\infty \langle \mathfrak{B}, \bar{b} \rangle$.
(3) $\langle \mathfrak{B}, \bar{b} \rangle \vDash \varphi^{\infty}_{\mathfrak{A},\bar{a}}$.
(4) *Duplicator wins* $\mathrm{EF}^{\aleph_0}_{\infty}(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$.

Proof. (a) We have already shown in Lemmas 2.1 and 3.3 that (2), (3), and (4) are equivalent.

(1) ⇒ (3) follows directly from the definition of $\equiv_\alpha$ since $\mathrm{qr}(\varphi^{\alpha}_{\mathfrak{A},\bar{a}}) \leq \alpha$.

(4) ⇒ (1) If $\langle \mathfrak{A}, \bar{a} \rangle \not\equiv_\alpha \langle \mathfrak{B}, \bar{b} \rangle$ then there is some formula $\varphi \in \mathrm{FO}_{\infty\aleph_0}$ of quantifier rank $\mathrm{qr}(\varphi) \leq \alpha$ such that $\mathfrak{A} \vDash \varphi(\bar{a})$ and $\mathfrak{B} \nvDash \varphi(\bar{b})$. By Lemma 3.4, spoiler wins $\mathrm{EF}_\alpha(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$.

(b) The equivalence (2) ⇔ (3) was proved in Theorem 2.4, and the implication (1) ⇒ (3) is trivial.

(2) ⇒ (4) Duplicator can win if she ensures that only positions $\langle \bar{c}, \bar{d} \rangle$ are reached where $\bar{c} \mapsto \bar{d} \in I_\infty(\mathfrak{A}, \mathfrak{B}) \subseteq \mathrm{pIso}(\mathfrak{A}, \mathfrak{B})$. But this is easily done since $I_\infty(\mathfrak{A}, \mathfrak{B})$ has the back-and-forth property with respect to itself. If spoiler chooses some element $c \in A$ then there exists an element $d \in B$ with $\bar{a}c \mapsto \bar{b}d \in I_\infty(\mathfrak{A}, \mathfrak{B})$. Similarly, if spoiler plays in $\mathfrak{B}$ then duplicator can respond in $\mathfrak{A}$.

(4) ⇒ (1) If $\langle \mathfrak{A}, \bar{a} \rangle \not\equiv_\infty \langle \mathfrak{B}, \bar{b} \rangle$ then there is some formula $\varphi \in \mathrm{FO}_{\infty\aleph_0}$ such that $\mathfrak{A} \vDash \varphi(\bar{a})$ and $\mathfrak{B} \nvDash \varphi(\bar{b})$. Let $\alpha := \mathrm{qr}(\varphi)$. By Lemma 3.4, spoiler wins $\mathrm{EF}_\alpha(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$. He can use the same strategy to win the infinite game $\mathrm{EF}^{\aleph_0}_{\infty}(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$. □

Corollary 3.6 (Ehrenfeucht, Fraïssé). *Let Σ be a relational signature and $\mathfrak{A}$ and $\mathfrak{B}$ Σ-structures. For $m < \omega$, let $\Delta_m \subseteq \mathrm{FO}[\Sigma]$ be the set of all first-order formulae of quantifier rank at most m.*

(a) $\mathfrak{A} \equiv_{\Delta_m} \mathfrak{B}$ iff $\mathfrak{A}|_{\Sigma_0} \cong_m \mathfrak{B}|_{\Sigma_0}$ *for all finite* $\Sigma_0 \subseteq \Sigma$,

(b) $\mathfrak{A} \equiv_{\mathrm{FO}} \mathfrak{B}$ iff $\mathfrak{A}|_{\Sigma_0} \cong_\omega \mathfrak{B}|_{\Sigma_0}$ *for all finite* $\Sigma_0 \subseteq \Sigma$.

Exercise 3.2. Find structures $\mathfrak{A}$ and $\mathfrak{B}$ such that $\mathfrak{A} \equiv_{\mathrm{FO}} \mathfrak{B}$ but $\mathfrak{A} \ncong_\omega \mathfrak{B}$.

Corollary 3.7. *Every formula $\psi \in \mathrm{FO}_{\infty\aleph_0}[\Sigma, X]$ of quantifier rank α is equivalent to a disjunction of Hintikka formulae of quantifier rank α. For $\psi \in \mathrm{FO}[\Sigma, X]$ and relational Σ, we can choose this disjunction in $\mathrm{FO}[\Sigma, X]$.*

Proof. We have $\psi \equiv \bigvee \Phi$ where

$$\Phi := \{ \varphi^{\alpha}_{\mathfrak{A},\bar{a}} \mid \mathfrak{A} \vDash \psi(\bar{a}) \}$$

is the set of all Hintikka formulae corresponding to models of ψ.

If $\psi \in \mathrm{FO}[\Sigma, X]$ then $\alpha < \omega$ and there exist finite subsets $\Sigma_0 \subseteq \Sigma$ and $X_0 \subseteq X$ such that $\psi \in \mathrm{FO}[\Sigma_0, X_0]$. Hence, we have $\psi \equiv \bigvee \Phi_0$ where $\Phi_0 := \Phi \cap \mathrm{FO}[\Sigma_0, X_0]$ is finite. □

We conclude this section with several applications of Ehrenfeucht-Fraïssé games.

Lemma 3.8. *There exists no first-order formula φ such that, for every finite structure $\mathfrak{A}$, we have*

$$\mathfrak{A} \vDash \varphi \quad \text{iff} \quad |A| \text{ is even.}$$

Proof. Suppose that such a formula φ exists and let $m := \mathrm{qr}(\varphi)$. By Corollary 1.14, we have

$$\langle [2^m], \leq \rangle \vDash \varphi \quad \text{iff} \quad \langle [2^m + 1], \leq \rangle \vDash \varphi\,.$$

A contradiction. □

Let us apply Ehrenfeucht-Fraïssé games to equivalence relations. Recall that we write $m =_k n$ iff $m = n$ or $m, n \geq k$. If E is an equivalence relation then we denote by $N^{=}_k(E)$ the number of E-classes $[a]_E$ of size $|[a]_E| = k$ and $N^{>}_k(E)$ denotes the number of classes of size $|[a]_E| > k$.

Lemma 3.9. *Let E and F be equivalence relations on the sets A and B, respectively. We have $\langle A, E\rangle \cong_m \langle B, F\rangle$ if and only if*

$$N^{=}_k(E) =_{m-k} N^{=}_k(F) \quad \text{and} \quad N^{>}_k(E) =_{m-k} N^{>}_k(F)\,,$$

for all $k \leq m$.

Proof. (⇒) First, suppose that $N_k^{=}(E) > N_k^{=}(F) =: s$. We claim that spoiler wins $\mathrm{EF}_{s+k+1}(\mathfrak{A}, \mathfrak{B})$. Since $N_k^{=}(E) > s$ we can find $s+1$ different E-classes $[a_0]_E, \dots, [a_s]_E$ of size $|[a_i]_E| = k$. In the first part of the game spoiler plays their representatives $a_0, \dots, a_s$. Duplicator has to answer with elements $b_0, \dots, b_s$ of different F-classes in B. Since we have $N_k^{=}(F) < s+1$ there is an index i such that $l := |[b_i]_F| \neq k$. If $l < k$ then spoiler continues by playing $k-1$ different elements

$$c_0, \dots, c_{k-2} \in [a_i]_E \setminus \{a_i\} .$$

Since $|[b_i]_F \setminus \{b_i\}| < k-1$ duplicator cannot answer all of them. Consequently, spoiler wins after at most $s+1+k-1 = s+k$ rounds. Similarly, if $l > k$ then spoiler plays k different elements

$$d_0, \dots, d_{k-1} \in [b_i]_F \setminus \{b_i\} ,$$

and again duplicator cannot answer all of them. In this case spoiler wins after at most $s+1+k$ rounds.

It remains to consider the case that $N_k^{>}(E) > N_k^{>}(F) =: s$. By a similar argument as above we show that spoiler wins $\mathrm{EF}_{s+k+1}(\mathfrak{A}, \mathfrak{B})$. Since $N_k^{>}(E) > s$ we can find $s+1$ different E-classes $[a_0]_E, \dots, [a_s]_E$ of size $|[a_i]_E| > k$. In the first part of the game spoiler plays their representatives $a_0, \dots, a_s$. Duplicator has to answer with elements $b_0, \dots, b_s$ of different F-classes in B. Since we have $N_k^{>}(F) < s+1$ there is an index i such that $|[b_i]_F| \leq k$. In the second part of the game spoiler plays k different elements

$$c_0, \dots, c_{k-1} \in [a_i]_E \setminus \{a_i\} .$$

Since $|[b_i]_F \setminus \{b_i\}| < k$ duplicator cannot answer all of them. Consequently, spoiler wins after at most $s+1+k$ rounds.

(⇐) For $k \leq m$, let I_k be the set of all partial isomorphisms $\bar{a} \mapsto \bar{b}$ with $\bar{a} \in A^{m-k}$ and $\bar{b} \in B^{m-k}$ such that

$$\big|[a_i]_E \setminus \bar{a}\big| =_k \big|[b_i]_F \setminus \bar{b}\big| , \quad \text{for all } i < m-k .$$

We claim that $(I_k)_k : \langle A, E\rangle \cong_m \langle B, F\rangle$. Clearly, we have $\langle\rangle \mapsto \langle\rangle \in I_m$. By symmetry, it is therefore sufficient to prove the forth property.

Let $\bar{a} \mapsto \bar{b} \in I_{k+1}$, and $c \in A$. We have to find some $d \in B$ such that $\bar{a}c \mapsto \bar{b}d \in I_k$. We consider several cases. If $c = a_i$, for some i, then $\bar{a}c \mapsto \bar{b}b_i \in I_k$. If $c \in [a_i]_E \setminus \bar{a}$, for some i, then

$$\big|[a_i]_E \setminus \bar{a}\big| =_{k+1} \big|[b_i]_E \setminus \bar{b}\big|$$

implies that there is some $d \in [b_i]_E \setminus \bar{b}$. It follows that $\bar{a}c \mapsto \bar{b}d \in I_k$.

It remains to consider the case that $c \notin [a_i]_E$, for all i. Set $s := |[c]_E|$. We are looking for an element $d \in B$ with $s =_{k+1} |[d]_F|$ and $[d]_F \cap \bar{b} = \varnothing$. First, consider the case that $s \leq k$. Then we have

$$|[a_i]_E| = s \quad \text{iff} \quad |[b_i]_F| = s\,.$$

Let l be the number of indices i with $|[a_i]_E| = s$. Since

$$N_s^{=}(E) =_{m-s} N_s^{=}(F) \quad \text{and} \quad l + 1 \leq m - k - 1 + 1 \leq m - s\,,$$

it follows that $N_s^{=}(E) \geq l + 1$ implies $N_s^{=}(F) \geq l + 1$. Consequently, we can choose some element $d \in B$ such that $|[d]_F| = s$ and $[d]_F \cap \bar{b} = \varnothing$.

The proof for the case that $s > k$ is analogous. Then we have

$$|[a_i]_E| > k \quad \text{iff} \quad |[b_i]_F| > k\,,$$

and we denote by l the number of indices i with $|[a_i]_E| > k$. Since

$$N_k^{>}(E) =_{m-k} N_k^{>}(F) \quad \text{and} \quad l + 1 \leq m - k\,,$$

it follows that $N_k^{>}(E) \geq l+1$ implies $N_k^{>}(F) \geq l+1$. Consequently, we can choose some element $d \in B$ such that $|[d]_F| > k$ and $[d]_F \cap \bar{b} = \varnothing$. □

We have seen in Lemma c1.1.7 that we can define every ordinal $\alpha < \kappa$ in $\mathrm{FO}_{\kappa\aleph_0}[<]$. Nevertheless there is no $\mathrm{FO}_{\infty\aleph_0}[<]$-formula that axiomatises the class of all well-orders.

Lemma 3.10. *For every ordinal α, there exists an ordinal $\delta > \alpha$ such that $\delta \equiv_\delta \delta + \delta \cdot \tau$, for each linear order τ.*

Proof. By Lemma A4.5.6, we can choose $\delta > \alpha$ such that $\omega^{(\delta)} = \delta$. Then δ is a limit ordinal such that $\delta = \omega^{(\beta)}\delta$, for all $\beta < \delta$. Hence, for each $\beta < \delta$, we can write δ as sum of δ copies of the order $\omega^{(\beta)}$. We call such a summand a $\omega^{(\beta)}$-*interval* of δ.

$$\delta: \quad \overbrace{\underbrace{\rule{1.5em}{0.4pt}}_{\omega^{(\beta)}}\ \underbrace{\rule{1.5em}{0.4pt}}_{\omega^{(\beta)}}\ \underbrace{\rule{1.5em}{0.4pt}}_{\omega^{(\beta)}}\ \underbrace{\rule{1.5em}{0.4pt}}_{\omega^{(\beta)}}\cdots}^{\delta}$$

In the same way we can write linear orders of the form $\delta + \delta\tau$ as a sum of $\omega^{(\beta)}$-intervals.

For $\beta < \delta$, let I_β be the set of all finite partial isomorphisms $\bar{a} \mapsto \bar{b} \in \mathrm{pIso}_{\aleph_0}(\delta, \delta + \delta\tau)$ satisfying the following conditions. For notational simplicity we assume that $a_0 < \dots < a_{n-1}$.

(1) a_i and a_{i+1} belong to the same $\omega^{(\beta)}$-interval iff b_i and b_{i+1} belong to the same $\omega^{(\beta)}$-interval.

(2) a_i is the α-th element of the $\omega^{(\beta)}$-interval containing a_i if and only if b_i is the α-th element of the $\omega^{(\beta)}$-interval containing b_i.

(3) a_0 is in the first $\omega^{(\beta)}$-interval if and only if b_0 is in the first $\omega^{(\beta)}$-interval.

Further, we set $I_\delta := \{\langle\rangle \mapsto \langle\rangle\}$. We claim that $(I_\beta)_{\beta<\delta} : \delta \cong_\delta \delta + \delta\tau$.

To prove the back property, suppose that $\bar{a} \mapsto \bar{b} \in I_{\beta+1}$ where $a_0 < \dots < a_{n-1}$, and let $d \in \delta + \delta\tau$.

If d belongs to the $\omega^{(\beta)}$-interval of some b_i then let c be the corresponding element in the $\omega^{(\beta)}$-interval of a_i. It follows that $\bar{a}c \mapsto \bar{b}d \in I_\beta$. If d belongs to the first $\omega^{(\beta)}$-interval or if $d > b_{n-1}$ then we can easily find a suitable element $c \in \delta$ such that $\bar{a}c \mapsto \bar{b}d \in I_\beta$.

It remains to consider the case that the $\omega^{(\beta)}$-interval of d lies strictly between those of b_i and b_{i+1}. Since a_i and a_{i+1} do not belong to the same $\omega^{(\beta)}$-interval we can choose some $\omega^{(\beta)}$-interval between those containing a_i and a_{i+1}. Let c be the α-th element of this interval, where

α is the position of d in its $\omega^{(\beta)}$-interval. Again, it follows that $\bar{a}c \mapsto \bar{b}d \in I_\beta$.

In the same way, we can prove the forth property. Since $\langle\rangle \mapsto \langle\rangle \in I_\beta$, for all $\beta < \delta$, it follows that $(I_\beta)_{\beta<\delta} : \delta \cong_\delta \delta + \delta\tau$. □

Theorem 3.11. *There is no sentence* $\varphi \in \mathrm{FO}_{\infty\aleph_0}[\leq]$ *such that*

$$\mathfrak{A} \vDash \varphi \quad \text{iff} \quad \mathfrak{A} \text{ is a well-order.}$$

Proof. Suppose there is such a formula φ. Let $\alpha := \mathrm{qr}(\varphi)$. By the preceding lemma we can find an ordinal $\delta > \alpha$ such that $\delta \equiv_\delta \delta + \delta\zeta$ where $\zeta := \langle\mathbb{Z}, \leq\rangle$. Since δ is a well-order we have $\delta \vDash \varphi$. This implies that $\delta + \delta\zeta \vDash \varphi$. Contradiction. □

4. *κ-complete back-and-forth systems*

Sometimes the partial isomorphisms of a back-and-forth systems can be used to construct a total isomorphism between two structures.

Definition 4.1. Let κ be an infinite cardinal and $I \subseteq \mathrm{pIso}(\mathfrak{A}, \mathfrak{B})$.

(a) The set I is *κ-complete* if, for every increasing chain $(p_i)_{i<\alpha} \subseteq I$ and every subset $X \subseteq \bigcup_{i<\alpha} p_i$ of size $|X| < \kappa$, there is some $q \in I$ with $\bigcup_{i<\alpha} p_i \restriction X \subseteq q$.

(b) I is *κ-bounded* if, for every $p \in I$ and each subset $X \subseteq \mathrm{dom}\, p$, there is a partial isomorphism $q \in I$ of size $|q| < |X|^+ \oplus \kappa$ such that $p \restriction X \subseteq q \subseteq p$.

(c) We call I *κ-finitary* if, for every $p \in \mathrm{pIso}_\kappa(\mathfrak{A}, \mathfrak{B})$, we have

$$p \in I \quad \text{iff} \quad p \restriction X \in I \quad \text{for all } \textit{finite } X \subseteq \mathrm{dom}\, p\,.$$

Remark. Note that every κ-finitary set is κ-complete and $\aleph_0$-bounded.

Definition 4.2. For structures $\mathfrak{A}$ and $\mathfrak{B}$ and a cardinal κ, we set

$$I^\kappa_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B}) := \left\{ \bar{a} \mapsto \bar{b} \in \mathrm{pIso}_\kappa(\mathfrak{A}, \mathfrak{B}) \;\middle|\; \langle\mathfrak{A}, \bar{a}\rangle \equiv_{\mathrm{FO}} \langle\mathfrak{B}, \bar{b}\rangle \right\}.$$

Remark. Since every first-order formula refers only to finitely many constants it follows that the sets $I^{\kappa}_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B})$ are κ-finitary and, hence, κ-complete.

Definition 4.3. Let $\mathfrak{A}$ and $\mathfrak{B}$ be structures and κ an infinite cardinal.

(a) For a set $I \subseteq \mathrm{pIso}(\mathfrak{A}, \mathfrak{B})$, we write $I : \langle \mathfrak{A}, \bar{a} \rangle \sqsubseteq^{\kappa}_{\mathrm{iso}} \langle \mathfrak{B}, \bar{b} \rangle$ if

- $\bar{a} \mapsto \bar{b} \subseteq p$, for some $p \in I$ with $|\mathrm{dom}\, p \setminus \bar{a}| < \kappa$,
- I is κ-complete and κ-bounded,
- I has the *forth property* with respect to itself. (We do not require the back property.)

Similarly, we define $I : \langle \mathfrak{A}, \bar{a} \rangle \cong^{\kappa}_{\mathrm{iso}} \langle \mathfrak{B}, \bar{b} \rangle$ if

- $\bar{a} \mapsto \bar{b} \subseteq p$, for some $p \in I$ with $|\mathrm{dom}\, p \setminus \bar{a}| < \kappa$,
- I is κ-complete and κ-bounded,
- I has the back-and-forth property with respect to itself,

that is, if

$$I : \langle \mathfrak{A}, \bar{a} \rangle \sqsubseteq^{\kappa}_{\mathrm{iso}} \langle \mathfrak{B}, \bar{b} \rangle \quad \text{and} \quad I : \langle \mathfrak{A}, \bar{a} \rangle \sqsupseteq^{\kappa}_{\mathrm{iso}} \langle \mathfrak{B}, \bar{b} \rangle .$$

We write $\mathfrak{A} \cong^{\kappa}_{\mathrm{iso}} \mathfrak{B}$ if there exists some set I with $I : \mathfrak{A} \cong^{\kappa}_{\mathrm{iso}} \mathfrak{B}$, and similarly for $\sqsubseteq^{\kappa}_{\mathrm{iso}}$.

(b) Of particular interest are the following special cases.

$$\begin{aligned}
\langle \mathfrak{A}, \bar{a} \rangle \sqsubseteq^{\kappa}_{0} \langle \mathfrak{B}, \bar{b} \rangle \quad &:\text{iff} \quad I^{\kappa}_{0}(\mathfrak{A}, \mathfrak{B}) : \langle \mathfrak{A}, \bar{a} \rangle \sqsubseteq^{\kappa}_{\mathrm{iso}} \langle \mathfrak{B}, \bar{b} \rangle , \\
\langle \mathfrak{A}, \bar{a} \rangle \cong^{\kappa}_{0} \langle \mathfrak{B}, \bar{b} \rangle \quad &:\text{iff} \quad I^{\kappa}_{0}(\mathfrak{A}, \mathfrak{B}) : \langle \mathfrak{A}, \bar{a} \rangle \cong^{\kappa}_{\mathrm{iso}} \langle \mathfrak{B}, \bar{b} \rangle , \\
\langle \mathfrak{A}, \bar{a} \rangle \sqsubseteq^{\kappa}_{\mathrm{FO}} \langle \mathfrak{B}, \bar{b} \rangle \quad &:\text{iff} \quad I^{\kappa}_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B}) : \langle \mathfrak{A}, \bar{a} \rangle \sqsubseteq^{\kappa}_{\mathrm{iso}} \langle \mathfrak{B}, \bar{b} \rangle , \\
\langle \mathfrak{A}, \bar{a} \rangle \cong^{\kappa}_{\mathrm{FO}} \langle \mathfrak{B}, \bar{b} \rangle \quad &:\text{iff} \quad I^{\kappa}_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B}) : \langle \mathfrak{A}, \bar{a} \rangle \cong^{\kappa}_{\mathrm{iso}} \langle \mathfrak{B}, \bar{b} \rangle , \\
\langle \mathfrak{A}, \bar{a} \rangle \sqsubseteq^{\kappa}_{\infty} \langle \mathfrak{B}, \bar{b} \rangle \quad &:\text{iff} \quad I^{\kappa}_{\infty}(\mathfrak{A}, \mathfrak{B}) : \langle \mathfrak{A}, \bar{a} \rangle \sqsubseteq^{\kappa}_{\mathrm{iso}} \langle \mathfrak{B}, \bar{b} \rangle , \\
\langle \mathfrak{A}, \bar{a} \rangle \cong^{\kappa}_{\infty} \langle \mathfrak{B}, \bar{b} \rangle \quad &:\text{iff} \quad I^{\kappa}_{\infty}(\mathfrak{A}, \mathfrak{B}) : \langle \mathfrak{A}, \bar{a} \rangle \cong^{\kappa}_{\mathrm{iso}} \langle \mathfrak{B}, \bar{b} \rangle .
\end{aligned}$$

Remark. (a) $I^{\aleph_0}_{\infty}(\mathfrak{A}, \mathfrak{B})$ is trivially $\aleph_0$-complete and $\aleph_0$-bounded. Hence, we have

$$\mathfrak{A} \cong^{\aleph_0}_{\infty} \mathfrak{B} \quad \text{iff} \quad \mathfrak{A} \cong_{\infty} \mathfrak{B} .$$

(b) The sets $I^\kappa_\text{o}(\mathfrak{A},\mathfrak{B})$ and $I^\kappa_\text{FO}(\mathfrak{A},\mathfrak{B})$ are κ-finitary and, hence, κ-complete and κ-bounded. Consequently, we have

$$\mathfrak{A} \sqsubseteq^\kappa_\text{o} \mathfrak{B} \quad \text{iff} \quad I^\kappa_\text{o}(\mathfrak{A},\mathfrak{B}) \text{ is nonempty and it has the forth property with respect to itself.}$$

and similarly for the relations $\cong^\kappa_\text{o}$, $\sqsubseteq^\kappa_\text{FO}$, and $\cong^\kappa_\text{FO}$.

(c) Note that we have

$$I^\kappa_\infty(\mathfrak{A},\mathfrak{B}) \subseteq I^\kappa_\text{FO}(\mathfrak{A},\mathfrak{B}) \subseteq I^\kappa_\text{o}(\mathfrak{A},\mathfrak{B}) \,.$$

Furthermore, we have shown in Lemma 1.11 that

$$I : \mathfrak{A} \cong^\kappa_\text{iso} \mathfrak{B} \quad \text{implies} \quad I \subseteq I_\infty(\mathfrak{A},\mathfrak{B}) \,.$$

Let us summarise these remarks in the following lemma.

Lemma 4.4. *Let κ be a cardinal and $x \in \{\text{o}, \text{FO}\}$.*

(a) *The following statements are equivalent:*

(1) $\mathfrak{A} \sqsubseteq^\kappa_x \mathfrak{B}$.

(2) *The set $I^\kappa_x(\mathfrak{A},\mathfrak{B})$ is nonempty and it has the forth property with respect to itself.*

(b) *The following statements are equivalent:*

(1) $\mathfrak{A} \cong^\kappa_x \mathfrak{B}$.

(2) $I^\kappa_x(\mathfrak{A},\mathfrak{B}) = I^\kappa_\infty(\mathfrak{A},\mathfrak{B}) \neq \emptyset$.

(3) *The set $I^\kappa_x(\mathfrak{A},\mathfrak{B})$ is nonempty and it has the back-and-forth property with respect to itself.*

As an example we consider dense linear orders.

Definition 4.5. Let $\mathfrak{A} = \langle A, < \rangle$ be a linear order.

(a) For $C, D \subseteq A$, we write $C < D$ if $c < d$, for all $c \in C$ and $d \in D$.

(b) $\mathfrak{A}$ is *κ-dense* if, for all sets $C, D \subseteq A$ of size $|C|, |D| < \kappa$ with $C < D$, there exists an element $a \in A$ such that $C < a < D$. Note that we allow $C = \emptyset$ or $D = \emptyset$.

Lemma 4.6. *If* $\mathfrak{B} = \langle B, < \rangle$ *is a* κ*-dense linear order then we have*

$$\mathfrak{A} \sqsubseteq^{\kappa}_{o} \mathfrak{B}, \quad \textit{for every linear order } \mathfrak{A}.$$

Proof. We have already noted that $\mathrm{pIso}_\kappa(\mathfrak{A}, \mathfrak{B})$ is κ-complete. Furthermore, since linear orders are relational structures we have

$$\langle\rangle \mapsto \langle\rangle \in \mathrm{pIso}_\kappa(\mathfrak{A}, \mathfrak{B}) \neq \varnothing.$$

Consequently, it remains to prove the forth property.

Let $p \in \mathrm{pIso}_\kappa(\mathfrak{A}, \mathfrak{B})$ and $a \in A$. If $a \in \operatorname{dom} p$ then we are done. Otherwise, we can partition the domain of p into

$$C := \{ c \in \operatorname{dom} p \mid c < a \} \quad \text{and} \quad D := \{ d \in \operatorname{dom} p \mid a < d \}.$$

Then $C < D$ which implies that $p[C] < p[D]$. Since $\mathfrak{B}$ is κ-dense and $|C|, |D| \leq |\operatorname{dom} p| < \kappa$ we can find some element $b \in B$ with

$$p[C] < b < p[D].$$

Hence, $p \cup \{\langle a, b \rangle\}$ is the desired partial isomorphism extending p. □

Corollary 4.7. *If* $\mathfrak{A}$ *and* $\mathfrak{B}$ *are* κ*-dense linear orders then* $\mathfrak{A} \cong^{\kappa}_{o} \mathfrak{B}$.

The relation $\cong^{\kappa}_{\mathrm{iso}}$ can also be characterised via Ehrenfeucht-Fraïssé games. The proof is completely analogous to that of Lemma 3.3.

Theorem 4.8. *Let* $\mathfrak{A}$ *and* $\mathfrak{B}$ *be structures and* κ *a cardinal. The following statements are equivalent:*

(1) $\langle \mathfrak{A}, \bar{a} \rangle \cong^{\kappa}_{\mathrm{iso}} \langle \mathfrak{B}, \bar{b} \rangle$.

(2) *Duplicator wins* $\mathrm{EF}^{\kappa}_{\infty}(\mathfrak{A}, \bar{a}, \mathfrak{B}, \bar{b})$.

κ-complete sets with the back-and-forth property can be used to construct embeddings and isomorphisms.

Lemma 4.9. *Let* $\mathfrak{A}$ *and* $\mathfrak{B}$ *be structures and* κ *an infinite cardinal.*

(a) *Suppose that* $I : \mathfrak{A} \cong^{\kappa}_{\text{iso}} \mathfrak{B}$. *For all sequences* $\bar{a} \in A^{\kappa}$ *and* $\bar{b} \in B^{\kappa}$, *there exist sequences* $\bar{c} \in A^{\kappa}$ *and* $\bar{d} \in B^{\kappa}$ *such that, for all* $\alpha < \kappa$,

$$I : \langle \mathfrak{A}, (a_i)_{i<\alpha}, (c_i)_{i<\alpha} \rangle \cong^{\kappa}_{\text{iso}} \langle \mathfrak{B}, (d_i)_{i<\alpha}, (b_i)_{i<\alpha} \rangle .$$

In particular, we have $\langle \mathfrak{A}, \bar{a}\bar{c} \rangle \cong_{0} \langle \mathfrak{B}, \bar{d}\bar{b} \rangle$.

(b) *Suppose that* $I : \mathfrak{A} \sqsubseteq^{\kappa}_{\text{iso}} \mathfrak{B}$. *For every sequence* $\bar{a} \in A^{\kappa}$, *there exist a sequence* $\bar{b} \in B^{\kappa}$ *such that*

$$I : \langle \mathfrak{A}, (a_i)_{i<\alpha} \rangle \sqsubseteq^{\kappa}_{\text{iso}} \langle \mathfrak{B}, (b_i)_{i<\alpha} \rangle , \quad \textit{for all } \alpha < \kappa .$$

In particular, we have $\langle \mathfrak{A}, \bar{a} \rangle \cong_{0} \langle \mathfrak{B}, \bar{b} \rangle$.

Proof. (a) We construct an increasing chain $(p_i)_{i<\kappa}$ of partial isomorphisms $p_i \in I$ with $|p_i| < \kappa$ such that $a_i \in \operatorname{dom} p_{i+1}$ and $b_i \in \operatorname{rng} p_{i+1}$, for all $i < \kappa$. Then we obtain the desired sequences $\bar{c}$ and $\bar{d}$ by setting

$$c_i := (p_{i+1})^{-1}(b_i) \quad \text{and} \quad d_i := p_{i+1}(a_i) .$$

Since $I : \mathfrak{A} \sqsubseteq^{\kappa}_{\text{iso}} \mathfrak{B}$ there is some $p_0 \in I$ with $|p_0| < \kappa$. Suppose that we have already defined $p_i \in I$, for $i < \alpha$. If α is a limit ordinal then, I being κ-complete, there is some $p_\alpha \in I$ such that

$$\bigcup_{i<\alpha} p_i \upharpoonright \big[\{ a_i \mid i < \alpha \} \cup \{ p^{-1}_{i+1}(b_i) \mid i < \alpha \} \big] \subseteq p_\alpha .$$

Finally, suppose that $\alpha = \gamma + 1$ is a successor. By the forth property we can find some $q \in I$ extending p_γ with $a_\gamma \in \operatorname{dom} q$. Analogously, there is some $p_\alpha \in I$ extending q with $b_\gamma \in \operatorname{rng} p_\alpha$.

(b) is proved in the same way as (a). We define an increasing chain $(p_i)_{i<\kappa}$ of partial isomorphisms such that $a_i \in \operatorname{dom} p_{i+1}$. For every a_i, we can use the forth property to find a suitable b_i. □

Lemma 4.10. *Let* $\mathfrak{A}$ *and* $\mathfrak{B}$ *be structures generated by* $A_0 \subseteq A$ *and* $B_0 \subseteq B$, *respectively.*

(a) *If* $\kappa \geq |A_0| \oplus |B_0|$ *and* $I : \mathfrak{A} \cong^{\kappa}_{\text{iso}} \mathfrak{B}$ *then* $\mathfrak{A} \cong \mathfrak{B}$.

(b) *If* $\kappa \geq |A_0|$ *and* $I : \mathfrak{A} \sqsubseteq^{\kappa}_{\text{iso}} \mathfrak{B}$ *then there exists an embedding* $\mathfrak{A} \to \mathfrak{B}$.

Proof. (a) Let $\bar{a}$ be an enumeration of A_0 and $\bar{b}$ one of B_0. By the preceding lemma, there are sequences $\bar{c} \subseteq A$ and $\bar{d} \subseteq B$ such that

$$\langle \mathfrak{A}, \bar{a}\bar{c} \rangle \cong_0 \langle \mathfrak{B}, \bar{d}\bar{b} \rangle .$$

In particular, the map $p : \bar{a}\bar{c} \mapsto \bar{d}\bar{b}$ is a partial isomorphism. By definition, there exists an isomorphism

$$\pi : \langle\!\langle \text{dom}\, p \rangle\!\rangle_{\mathfrak{A}} \cong \langle\!\langle \text{rng}\, p \rangle\!\rangle_{\mathfrak{B}}$$

extending p. Since $\text{dom}\, p \supseteq \bar{a} = A_0$ and $\text{rng}\, p \supseteq \bar{b} = B_0$ it follows that π is a total isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$.

(b) Given an enumeration $\bar{a}$ of A_0 we can find a sequence $\bar{b} \subseteq B$ such that $\langle \mathfrak{A}, \bar{a} \rangle \cong_0 \langle \mathfrak{B}, \bar{b} \rangle$. Hence, $\bar{a} \mapsto \bar{b}$ is a partial isomorphism that can be extended to an isomorphism

$$\pi : \langle\!\langle \bar{a} \rangle\!\rangle_{\mathfrak{A}} \cong \langle\!\langle \bar{b} \rangle\!\rangle_{\mathfrak{B}} .$$

Since $\langle\!\langle \bar{a} \rangle\!\rangle_{\mathfrak{A}} = A$ it follows that π is the desired embedding. □

Corollary 4.11. *If* $\mathfrak{A}$ *and* $\mathfrak{B}$ *are countable structures with* $\mathfrak{A} \equiv_{\omega_1} \mathfrak{B}$ *then* $\mathfrak{A} \cong \mathfrak{B}$.

Proof. Let α be the Scott height of $\mathfrak{A}$. Then $\alpha < |A|^+ \leq \aleph_1$. Hence, $\mathfrak{A} \equiv_{\omega_1} \mathfrak{B}$ implies that $\mathfrak{B} \vDash \varphi^{\infty}_{\mathfrak{A}}$ where $\varphi^{\infty}_{\mathfrak{A}}$ is the Scott sentence of $\mathfrak{A}$. By Theorem 2.4, it follows that $\mathfrak{A} \cong_{\infty} \mathfrak{B}$. This is equivalent to $\mathfrak{A} \cong^{\aleph_0}_{\infty} \mathfrak{B}$ since $I^{\aleph_0}_{\infty}(\mathfrak{A}, \mathfrak{B})$ is always $\aleph_0$-complete. Hence, Lemma 4.10 (a) implies that $\mathfrak{A} \cong \mathfrak{B}$. □

Corollary 4.12. (a) *If* $\mathfrak{A}$ *and* $\mathfrak{B}$ *are* κ*-dense linear orders of size at most* κ *then* $\mathfrak{A} \cong \mathfrak{B}$.

(b) *If* $\mathfrak{B}$ *is a* κ*-dense linear order then every linear order* $\mathfrak{A}$ *of size at most* κ *can be embedded into* $\mathfrak{B}$.

Proof. Immediately from Lemma 4.6 and Corollary 4.7. □

We can show that the relation $\cong^{\kappa}_{\text{iso}}$ is reflexive and symmetric, but it is unknown whether it is also transitive. The relations $\cong^{\kappa}_{0}$, $\cong^{\kappa}_{\text{FO}}$, and $\cong^{\kappa}_{\infty}$, on the other hand, are transitive and symmetric but not reflexive.

Lemma 4.13. *If* $\mathfrak{A} \cong \mathfrak{B}$ *then* $\mathfrak{A} \cong^{\kappa}_{\text{iso}} \mathfrak{B}$*, for all* κ*.*

Proof. Fix an isomorphism $\pi : \mathfrak{A} \cong \mathfrak{B}$. The set

$$I := \{ p \in \text{pIso}_{\kappa}(\mathfrak{A}, \mathfrak{B}) \mid p \subseteq \pi \}$$

is nonempty, κ-finitary, and it has the back-and-forth property with respect to itself. Hence, we have $I : \mathfrak{A} \cong^{\kappa}_{\text{iso}} \mathfrak{B}$. □

Remark. The above lemma fails for the relations $\cong^{\kappa}_{0}$, $\cong^{\kappa}_{\text{FO}}$, and $\cong^{\kappa}_{\infty}$. In fact, we can even find structures $\mathfrak{A}$ such that $\mathfrak{A} \not\cong^{\aleph_0}_{0} \mathfrak{A}$ or $\mathfrak{A} \not\cong^{\aleph_0}_{\text{FO}} \mathfrak{A}$. For instance, if we take $\mathfrak{A} := \langle \omega, \leq \rangle$ then $0 \mapsto 1 \in I^{\aleph_0}_{0}(\mathfrak{A}, \mathfrak{A})$ but there exists no element $a \in \omega$ with $\langle 0, a \rangle \mapsto \langle 1, 0 \rangle \in I^{\aleph_0}_{0}(\mathfrak{A}, \mathfrak{A})$. Structures such that $\mathfrak{A} \cong^{\kappa}_{\text{FO}} \mathfrak{A}$ are called *κ-homogeneous.* They are the subject of Section E1.1.

Lemma 4.14. *Let* κ *be a cardinal and* $x \in \{0, \text{FO}, \infty\}$*.*

$$\mathfrak{A} \sqsubseteq^{\kappa}_{x} \mathfrak{B} \sqsubseteq^{\kappa}_{x} \mathfrak{C} \quad \textit{implies} \quad \mathfrak{A} \sqsubseteq^{\kappa}_{x} \mathfrak{C} .$$

Proof. Let $L_x \subseteq \text{FO}_{\infty\aleph_0}$ be the logic such that

$$\bar{a} \mapsto \bar{b} \in I^{\kappa}_{x}(\mathfrak{A}, \mathfrak{B}) \quad \text{iff} \quad \langle \mathfrak{A}, \bar{a} \rangle \equiv_{L_x} \langle \mathfrak{B}, \bar{b} \rangle .$$

We start by showing that

$$I^{\kappa}_{x}(\mathfrak{A}, \mathfrak{C}) = \{ q \circ p \mid p \in I^{\kappa}_{x}(\mathfrak{A}, \mathfrak{B}) ,\ q \in I^{\kappa}_{x}(\mathfrak{B}, \mathfrak{C}) \} .$$

Clearly, if p and q preserve all L_x-formulae then so does $q \circ p$. Therefore, we only have to show that, for every $\bar{a} \mapsto \bar{c} \in I^{\kappa}_{x}(\mathfrak{A}, \mathfrak{C})$, there is some tuple $\bar{b}$ such that $\bar{a} \mapsto \bar{b} \in I^{\kappa}_{x}(\mathfrak{A}, \mathfrak{B})$ and $\bar{b} \mapsto \bar{c} \in I^{\kappa}_{x}(\mathfrak{B}, \mathfrak{C})$.

Given $\bar{a}$ of length $|\bar{a}| < \kappa$, we can find, by Lemma 4.9, some tuple $\bar{b}$ such that $\bar{a} \mapsto \bar{b} \in I^{\kappa}_{x}(\mathfrak{A}, \mathfrak{B})$. Since the maps $\bar{b} \mapsto \bar{a}$ and $\bar{a} \mapsto \bar{c}$ preserve all L_x-formulae it follows that so does $\bar{b} \mapsto \bar{c}$. Consequently, we also have $\bar{b} \mapsto \bar{c} \in I_x^\kappa(\mathfrak{B}, \mathfrak{C})$, as desired.

To prove the lemma, first note that the claim implies that $\langle\rangle \mapsto \langle\rangle \in I_x^\kappa(\mathfrak{A}, \mathfrak{C}) \neq \emptyset$. Therefore, it remains to check that $I_x^\kappa(\mathfrak{A}, \mathfrak{C})$ has the forth property with respect to itself. Let $\pi \in I_x^\kappa(\mathfrak{A}, \mathfrak{C})$ and $a \in A$. Then $\pi = q \circ p$, for some $p \in I_x^\kappa(\mathfrak{A}, \mathfrak{B})$ and $q \in I_x^\kappa(\mathfrak{B}, \mathfrak{C})$. Since these sets have the forth property, we can find elements $b \in B$ and $c \in C$ such that

$$p' := p \cup \{\langle a, b\rangle\} \in I_x^\kappa(\mathfrak{A}, \mathfrak{B})$$
$$\text{and} \quad q' := q \cup \{\langle b, c\rangle\} \in I_x^\kappa(\mathfrak{B}, \mathfrak{C})\,.$$

It follows that $\pi \cup \{\langle a, c\rangle\} = q' \circ p' \in I_x^\kappa(\mathfrak{A}, \mathfrak{C})$. □

Since the relations $\cong_x^\kappa$ are clearly symmetric we have the following corollaries.

Corollary 4.15. *Let κ be a cardinal and $x \in \{0, \mathrm{FO}, \infty\}$.*

(a) *If $\mathfrak{A} \cong_x^\kappa \mathfrak{B}$ then $\mathfrak{A} \cong_x^\kappa \mathfrak{A}$.*

(b) *The relation $\sqsubseteq_x^\kappa$ is a preorder on the class*

$$\mathcal{C} := \{\,\mathfrak{A} \mid \mathfrak{A} \cong_x^\kappa \mathfrak{A}\,\}\,.$$

5. *The theorems of Hanf and Gaifman*

In nontrivial applications the combinatorics involved in playing Ehrenfeucht-Fraïssé games quickly become unmanageable. Therefore, it is desirable to develop methods to simplify such games.

Definition 5.1. Let $\mathfrak{A}$ be a relational Σ-structure. The *Gaifman graph* of $\mathfrak{A}$ is the graph $\mathcal{G}(\mathfrak{A}) := \langle A, E\rangle$ with edge relation

$$E := \{\,\langle a, b\rangle \in A^2 \mid a \neq b \text{ and } a, b \in \bar{c} \text{ for some } \bar{c} \in R^{\mathfrak{A}},\ R \in \Sigma\,\}\,.$$

Definition 5.2. Let $\mathfrak{A}$ be a relational structure. The following definitions will only be used in this section.

(a) We denote by $d(a, b)$ the distance between a and b in $\mathcal{G}(\mathfrak{A})$.

(b) For $X, Y \subseteq A$, we set $d(X, Y) := \min \{ d(a, b) \mid a \in X,\ b \in Y \}$.

(c) The *r-neighbourhood* of $a \in A$ is the set

$$N(r, a) := \{ b \in A \mid d(a, b) < r \} .$$

For $\bar{a} \in A^n$, we set $N(r, \bar{a}) := \bigcup_i N(r, a_i)$. In particular, we have $N(r, \langle\rangle) := \varnothing$. Finally,

$$\mathfrak{N}(r, \bar{a}) := \langle \mathfrak{A}|_{N(r,\bar{a})}, \bar{a} \rangle$$

is the substructure induced by $N(r, \bar{a})$.

(d) The *$N(r)$-type* of $\bar{a} \subseteq A$ is the isomorphism type of $\mathfrak{N}(r, \bar{a})$, i.e., the $\cong$-class of this structure.

(e) For a $N(r)$-type τ, let $\#_\tau(\mathfrak{A})$ be the number of tuples $\bar{a} \subseteq A$ that have $N(r)$-type τ.

(f) Finally, for $k, m, n < \omega$, recall that

$$m =_k n \quad :\text{iff} \quad m = n \text{ or } m, n \geq k .$$

Theorem 5.3 (Hanf). *Let $m < \omega$ and let $\mathfrak{A}$ and $\mathfrak{B}$ be relational structures such that every 3^m-neighbourhood in $\mathfrak{A}$ and $\mathfrak{B}$ has at most $k < \aleph_0$ elements. If*

$$\#_\tau(\mathfrak{A}) =_{mk} \#_\tau(\mathfrak{B}), \qquad \textit{for every } N(n)\textit{-type } \tau \textit{ with } n \leq 3^m,$$

then $\mathfrak{A} \equiv_m \mathfrak{B}$.

Proof. Let I_n be the set of all partial isomorphisms $\bar{a} \mapsto \bar{b}$ with $\bar{a} \in A^{m-n}$ and $\bar{b} \in B^{m-n}$ such that $\mathfrak{N}(3^n, \bar{a}) \cong \mathfrak{N}(3^n, \bar{b})$. We claim that $(I_n)_n : \mathfrak{A} \cong_m \mathfrak{B}$.

We have $\langle\rangle \mapsto \langle\rangle \in I_m$. By symmetry, we therefore only need to prove the forth property. Suppose that $\bar{a} \mapsto \bar{b} \in I_{n+1}$. By definition, there exists an isomorphism

$$\pi : \mathfrak{N}(3^{n+1}, \bar{a}) \cong \mathfrak{N}(3^{n+1}, \bar{b}) .$$

Let $c \in A$. If $c \in N(2 \cdot 3^n, \bar{a})$ then $N(3^n, \bar{a}c) \subseteq N(3^{n+1}, \bar{a})$ and setting $d := \pi(c)$ we have

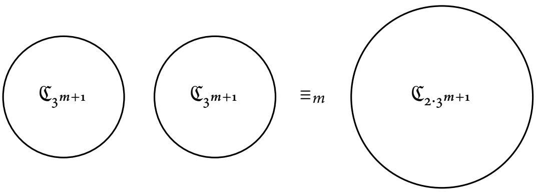
$$\pi : \mathfrak{N}(3^n, \bar{a}c) \cong \mathfrak{N}(3^n, \bar{b}d), \quad \text{that is,} \quad \bar{a}c \mapsto \bar{b}d \in I_n .$$

If, on the other hand, $c \notin N(2 \cdot 3^n, \bar{a})$ then $d(N(3^n, \bar{a}), N(3^n, c)) > 1$. Let τ be the $N(3^n)$-type of c. Since π is an isomorphism we have the same number of elements of 3^n-type τ in $N(2 \cdot 3^n, \bar{a})$ and $N(2 \cdot 3^n, \bar{b})$. This number is at most $|\bar{a}| \cdot k = (m - n - 1) \cdot k < mk$. Since $\#_\tau(\mathfrak{A}) =_{mk} \#_\tau(\mathfrak{B})$ there exists some $d \in B \setminus N(2 \cdot 3^n, \bar{b})$ of $N(3^n)$-type τ. Let $\sigma : \mathfrak{N}(3^n, c) \cong \mathfrak{N}(3^n, d)$ be the corresponding isomorphism of neighbourhoods. It follows that

$$\pi \cup \sigma : \mathfrak{N}(3^n, \bar{a}c) \cong \mathfrak{N}(3^n, \bar{b}d),$$

which implies that $\bar{a}c \mapsto \bar{b}d \in I_n$. □

Example. (a) We have already seen in the example on page 518 that there is no first-order formula expressing that a graph is connected. The Theorem of Hanf allows an easy alternate proof. For a contradiction, suppose that there is such a formula φ and let m be its quantifier rank.

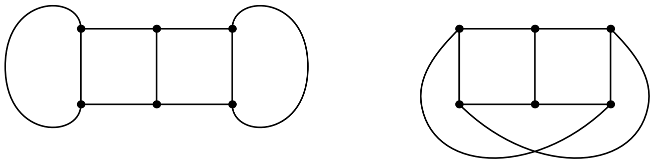


Let $\mathfrak{A} := \mathfrak{C}_{3^{m+1}} \dot\cup \mathfrak{C}_{3^{m+1}}$ be the graph consisting of two disjoint copies of the cycle of length 3^{m+1} and let $\mathfrak{B} := \mathfrak{C}_{2 \cdot 3^{m+1}}$ be the cycle of length $2 \cdot 3^{m+1}$. Then we have

$$\#_\tau(\mathfrak{A}) = \#_\tau(\mathfrak{B}), \qquad \text{for every } N(r)\text{-type } \tau \text{ with } r \leq 3^m.$$

By the Theorem of Hanf, it follows that $\mathfrak{A} \equiv_m \mathfrak{B}$. In particular, $\mathfrak{A} \vDash \varphi$ iff $\mathfrak{B} \vDash \varphi$. Contradiction.

(b) In the same way we can prove that planarity of a graph is not expressible in first-order logic. If φ is a formula of quantifier rank m then, by the Theorem of Hanf, it cannot distinguish between the graphs



where each line represents a path of length 3^{m+1}. Since one of the graphs if planar while the other one is not it follows that φ does not define the class of planar graphs.

With the help of the Theorem of Hanf we can avoid playing Ehrenfeucht-Fraïssé games, but the theorem can only be applied to structures where the r-neighbourhoods are finite. If we want to drop this restriction we have to replace the isomorphism type of a neighbourhood by its α-equivalence type. This is the idea behind the Theorem of Gaifman below.

Remark. Let Σ be a finite signature. For all $n < \omega$, there exists a formula $\varphi_n(x, y) \in \mathrm{FO}[\Sigma]$ such that

$$\mathfrak{A} \vDash \varphi_n(a, b) \quad \text{iff} \quad d(a, b) < n\,, \quad \text{for every } \Sigma\text{-structure } \mathfrak{A}\,.$$

Definition 5.4. (a) A set $X \subseteq A$ is *r-scattered* if $d(a, b) \geq r$, for all distinct elements $a, b \in X$.

(b) For $\varphi(\bar{x}) \in \mathrm{FO}[\Sigma, X]$, we denote by $\varphi^{(r)}(\bar{x})$ the relativisation of φ to the (definable) set $N(r, \bar{x})$.

(c) A sentence of the form

$$\exists x_0 \cdots \exists x_{n-1} \bigwedge_{i<k} \big(d(x_i, x_k) \geq 2r \wedge \psi^{(r)}(x_i)\big)$$

is called *basic local.* A boolean combination of basic local sentences is called *local.*

Lemma 5.5. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be Σ-structures. We have $\mathfrak{A} \equiv_{\mathrm{FO}} \mathfrak{B}$ if and only if*

$$\mathfrak{A} \vDash \varphi \quad \text{iff} \quad \mathfrak{B} \vDash \varphi \qquad \textit{for all basic local sentences } \varphi\,.$$

Proof. We have to show that $\mathfrak{A}|_{\Sigma_0} \equiv_m \mathfrak{B}|_{\Sigma_0}$, for all $m < \omega$ and all finite $\Sigma_0 \subseteq \Sigma$. Fix m and Σ_0 and let I_n be the set of all partial isomorphisms $\bar{a} \mapsto \bar{b} \in \mathrm{pIso}(\mathfrak{A}|_{\Sigma_0}, \mathfrak{B}|_{\Sigma_0})$ with $|\bar{a}| \in A^{m-n}$ and $|\bar{b}| \in B^{m-n}$ such that

$$\mathfrak{N}(7^n, \bar{a}) \equiv_{g(n)} \mathfrak{N}(7^n, \bar{b})\,,$$

where $g : \omega \to \omega$ is some function that will be specified below. We claim that $(I_n)_n : \mathfrak{A}|_{\Sigma_0} \equiv_m \mathfrak{B}|_{\Sigma_0}$.

Since $\langle\rangle \mapsto \langle\rangle \in I_m$ it remains to prove the forth property. Let $\bar{a} \mapsto \bar{b} \in I_{n+1}$ and $c \in A$. By Lemma 2.1, there exist formulae $\varphi^n_{\mathfrak{D},\bar{d}}$ such that

$$\mathfrak{C} \vDash \varphi^n_{\mathfrak{D},\bar{d}}(\bar{c}) \quad \text{iff} \quad \langle \mathfrak{C}, \bar{c} \rangle \equiv_n \langle \mathfrak{D}, \bar{d} \rangle\,.$$

If we define

$$\psi^n_{\bar{d}} := \big(\varphi^{g(n)}_{\mathfrak{N}(7^n,\bar{d})}\big)^{(7^n)}$$

then we have

$$\mathfrak{C} \vDash \psi^n_{\bar{d}}(\bar{c}) \quad \text{iff} \quad \mathfrak{N}(7^n, \bar{d}) \equiv_{g(n)} \mathfrak{N}(7^n, \bar{c})\,.$$

We distinguish two cases. If $c \in N(2 \cdot 7^n, \bar{a})$ then

$$\mathfrak{N}(7^{n+1}, \bar{a}) \vDash \exists z\big(d(\bar{a}, z) < 2 \cdot 7^n \wedge \psi^n_{\bar{a}c}(\bar{a}z)\big)\,.$$

Choose $g(n+1)$ such that it is larger than the quantifier rank of this formula. Then it follows that

$$\mathfrak{N}(7^{n+1}, \bar{b}) \vDash \exists z\big(d(\bar{b}, z) < 2 \cdot 7^n \wedge \psi^n_{\bar{a}c}(\bar{b}z)\big)\,.$$

Therefore, there is some $d \in N(7^{n+1}, \bar{b})$ such that

$$\mathfrak{N}(7^n, \bar{a}c) \equiv_{g(n)} \mathfrak{N}(7^n, \bar{b}d)\,, \quad \text{that is,} \quad \bar{a}c \mapsto \bar{b}d \in I_n\,.$$

It remains to consider the case that $c \notin N(2 \cdot 7^n, \bar{a})$. Then

$$d(N(7^n, \bar{a}), N(7^n, c)) > 1 .$$

The formula

$$\delta_s(\bar{x}) := \bigwedge_{l<k<s} d(x_l, x_k) \geq 4 \cdot 7^n \wedge \bigwedge_{l<s} \psi_c^n(x_l)$$

says that the set $\{x_0, \dots, x_{s-1}\}$ is $(4 \cdot 7^n)$-scattered and the 7^n-neighbourhood of every x_l is $g(n)$-equivalent to $\mathfrak{N}(7^n, c)$. Choose e maximal such that

$$\mathfrak{N}(7^{n+1}, \bar{a}) \vDash \chi_e := \exists x_0 \cdots \exists x_{e-1} \Big(\delta_e(\bar{x}) \wedge \bigwedge_{k<e} d(\bar{a}, x_k) < 2 \cdot 7^n \Big) .$$

Note that e is well-defined since $N(2 \cdot 7^n, \bar{a})$ does not contain a $(4 \cdot 7^n)$-scattered set of size greater than $|\bar{a}| = m - n - 1$. If we choose $g(n+1)$ large enough such that $\mathrm{qr}(\chi_e \wedge \neg\chi_{e+1}) \leq g(n+1)$ it follows that

$$\mathfrak{N}(7^{n+1}, \bar{b}) \vDash \chi_e \wedge \neg\chi_{e+1} .$$

Since the sentence $\vartheta_i := \exists x_0 \cdots \exists x_{i-1} \delta_i(\bar{x})$ is basic local we have

$$\mathfrak{B} \vDash \vartheta_i \quad \text{iff} \quad \mathfrak{A} \vDash \vartheta_i .$$

If $\mathfrak{B} \vDash \vartheta_{e+1}$ then there exists some $d \in B \setminus N(2 \cdot 7^n, \bar{b})$ such that $\mathfrak{B} \vDash \psi_c^n(d)$. It follows that $\mathfrak{N}(7^n, c) \equiv_{g(n)} \mathfrak{N}(7^n, d)$ and $\bar{a}c \mapsto \bar{b}d \in I_n$.

It remains to consider the case that $\mathfrak{B} \nvDash \vartheta_{e+1}$. Then the distance between $\bar{a}$ and every element satisfying $\psi_c^n(x)$ is less than

$$4 \cdot 7^n + 2 \cdot 7^n = 6 \cdot 7^n < 7^{n+1}.$$

Since $c \notin N(2 \cdot 7^n, \bar{a})$ we have

$$\mathfrak{N}(7^{n+1}, \bar{a}) \vDash \exists z \big[2 \cdot 7^n \leq d(\bar{a}, z) < 6 \cdot 7^n \wedge \psi_c^n(z) \wedge \psi_{\bar{a}}^n(\bar{a}) \big]$$

which implies that

$$\mathfrak{N}(7^{n+1}, \bar{b}) \vDash \exists z\big[2 \cdot 7^n \leq d(\bar{b}, z) < 6 \cdot 7^n \wedge \psi^n_c(z) \wedge \psi^n_{\bar{a}}(\bar{b})\big]$$

if we choose $g(n+1)$ larger than the quantifier rank of this formula. Therefore, there exists some element $d \in N(7^{n+1}, \bar{b})$ with

$$2 \cdot 7^n \leq d(\bar{b}, d) < 6 \cdot 7^n$$

such that

$$\mathfrak{N}(7^n, c) \equiv_{g(n)} \mathfrak{N}(7^n, d)\,.$$

It follows that $\mathfrak{N}(7^n, \bar{a}c) \equiv_{g(n)} \mathfrak{N}(7^n, \bar{b}d)$ and $\bar{a}c \mapsto \bar{b}d \in I_n$, as desired. □

The preceding lemma implies that every sentence is equivalent to a local one.

Theorem 5.6 (Gaifman). *Every sentence $\varphi \in \mathrm{FO}^0$ is equivalent to some local sentence.*

Proof. Let $\Phi := \{\, \psi \mid \psi \text{ is local and } \varphi \vDash \psi \,\}$. We claim that $\Phi \vDash \varphi$. By the Compactness Theorem, it then follows that $\Phi_0 \vDash \varphi$, for some finite subset $\Phi_0 \subseteq \Phi$. This implies that $\varphi \equiv \bigwedge \Phi_0$.

Suppose that $\mathfrak{A} \vDash \Phi$. We have to show that $\mathfrak{A} \vDash \varphi$. Set

$$\Psi := \{\, \psi \mid \psi \text{ is local and } \mathfrak{A} \vDash \psi \,\}\,.$$

If $\Psi \cup \{\varphi\}$ has some model $\mathfrak{B}$ then, since $\mathfrak{B} \vDash \Psi$ and local sentences are closed under negation, it follows by the preceding lemma that $\mathfrak{B} \equiv_{\mathrm{FO}} \mathfrak{A}$ and

$$\mathfrak{B} \vDash \varphi \quad \text{implies} \quad \mathfrak{A} \vDash \varphi\,.$$

Therefore, it is sufficient to show that $\Psi \cup \{\varphi\}$ is satisfiable. Suppose otherwise. Then, by the Compactness Theorem, there are finitely many formulae $\psi_0, \dots, \psi_n \in \Psi$ such that

$$\psi_0 \wedge \dots \wedge \psi_n \vDash \neg\varphi\,.$$

Hence, we have $\neg\psi_0 \vee \dots \vee \neg\psi_n \in \Phi$ which implies that $\mathfrak{A} \vDash \neg\psi_0 \vee \dots \vee \neg\psi_n$. It follows that there is some $i \leq n$ with $\psi_i \notin \Psi$. Contradiction. $\square$

C5. General model theory

1. Classifying logical systems

In this chapter we start with a more systematic investigation of the various extensions of first-order logic. Let us isolate some desirable properties a logic may have.

Definition 1.1. Let L and L' be logics. We write $L \leq L'$ if, for every $\varphi \in L$, there exists a formula $\varphi' \in L'$ such that $\mathrm{Mod}_{L'}(\varphi') = \mathrm{Mod}_L(\varphi)$.

Similarly, if $\mathcal{L} : \mathfrak{S} \to \mathfrak{Logic}$ and $\mathcal{L}' : \mathfrak{S}' \to \mathfrak{Logic}$ are logical systems then we write $\mathcal{L} \leq \mathcal{L}'$ if there exists a functor $F : \mathfrak{S} \to \mathfrak{S}'$ such that

$$\mathcal{L}[s] \leq \mathcal{L}'[F(s)]\,, \quad \text{for all } s \in \mathfrak{S}\,.$$

We write $L \equiv L'$ if $L \leq L'$ and $L \geq L'$. By $L < L'$ we denote the fact that $L \leq L'$ and $L \not\equiv L'$. The same notation is used for logical systems.

Definition 1.2. Let L be a logical system.

(a) L has the *finite occurrence property* if L is algebraic and, for every $\varphi \in L[\Sigma]$, there exists a finite set S of sorts and a finite S-sorted signature $\Sigma_0 \subseteq \Sigma$ such that φ is equivalent to some formula in $L[\Sigma_0]$.

(b) L is *compact* if every inconsistent set $\Phi \subseteq L[s]$ has a finite subset that is already inconsistent. Similarly, we call L *countably compact* if every countable inconsistent set $\Phi \subseteq L[s]$ has a finite inconsistent subset.

(c) L has the *Löwenheim-Skolem property* if it is algebraic and every formula $\varphi \subseteq L[\Sigma]$ that is satisfiable has a countable model.

(d) L has the *Karp property* if it is algebraic and

$$\mathfrak{A} \cong_\infty \mathfrak{B} \quad \text{implies} \quad \mathfrak{A} \equiv_L \mathfrak{B}\,.$$

(e) L is *closed under relativisations* if it is algebraic and, for all formulae $\varphi \in L[\Sigma]$ and $\chi_i \in L^{s_i}[\Sigma \cup \Gamma]$, for $i < n$, there exists a formula $\varphi^{(\bar{\chi})} \in L[\Sigma]$ such that we have

$$\mathfrak{A} \vDash \varphi^{(\bar{\chi})} \quad \text{iff} \quad (\mathfrak{A}|_\Sigma)|_{\bigcup_i \chi_i^{\mathfrak{A}}} \vDash \varphi\,,$$

whenever $\mathfrak{A}$ is a $(\Sigma \cup \Gamma)$-structure such that the set $\bigcup_i \chi_i^{\mathfrak{A}}$ induces a substructure of $\mathfrak{A}|_\Sigma$.

(f) L is *closed under substitutions* if it is algebraic and, for all formulae $\varphi \in L[\Sigma \cup \{R\}]$ and $\chi \in L^{\bar{s}}[\Sigma]$ where R is a relation symbol of type $\bar{s}$, there exists a formula $\varphi' \in L[\Sigma]$ such that

$$\mathfrak{A} \vDash \varphi' \quad \text{iff} \quad \langle \mathfrak{A}, \chi^{\mathfrak{A}} \rangle \vDash \varphi\,, \quad \text{for every } \Sigma\text{-structure } \mathfrak{A}\,.$$

(g) L has the *Tarski union property* if it is algebraic and, for every L-chain $(\mathfrak{A}_\alpha)_{\alpha<\delta}$, we have $\mathfrak{A}_\beta \preceq_L \bigcup_{\alpha<\delta} \mathfrak{A}_\alpha$, for all $\beta < \delta$.

(h) Let us define the following abbreviations:

- (A) L is algebraic.
- (B) L is boolean closed.
- (B_+) L is closed under finite conjunctions and disjunctions.
- (C) L is compact.
- (CC) L is countably compact.
- (FOP) L has the finite occurrence property.
- (KP) L has the Karp property.
- (LSP) L has the Löwenheim-Skolem property.
- (REL) L is closed under relativisations.
- (SUB) L is closed under substitutions.
- (TUP) L has the Tarski union property.

(i) L is called *weakly regular* if it satisfies (A), (B_+), and (FOP). If L satisfies (A), (B), (FOP), (REL), and (SUB) then it is called *regular.*

Example. FO° has all of the above properties but, if $\kappa > \aleph_0$ then $\text{FO}^\circ_{\kappa\aleph_0}$ satisfies only (A), (B), (B_+), (KP), (REL), and (SUB).

Exercise 1.1. Prove that SO does not have the Karp property.

Lemma 1.3. *Suppose that* $L_0 \leq L_1$. *If* L_1 *satisfies* (C), (CC), (LSP), *or* (KP) *then so does* L_0.

Exercise 1.2. (a) Suppose that L is closed under disjunction. Prove that L is compact if and only if the type space $S(L)$ is compact.

(b) Suppose that the logic L is compact and closed under negation. Let $\Phi \subseteq L$ and $\varphi \in L$. Prove that $\Phi \vDash \varphi$ if and only if $\Phi_0 \vDash \varphi$, for some finite subset $\Phi_0 \subseteq \Phi$.

The following lemmas summarise some consequences of compactness.

Lemma 1.4. *Let L be a logic with* (B) *and* (C). *If*

$$\varphi \equiv \bigvee_{i \in I} \bigwedge \Phi_i\,, \quad \textit{for } \varphi \in L \textit{ and } \Phi_i \subseteq L\,, i \in I\,,$$

then there exist finite sets $I_0 \subseteq I$ *and* $\Phi_i^0 \subseteq \Phi_i$ *such that*

$$\varphi \equiv \bigvee_{i \in I_0} \bigwedge \Phi_i^0\,.$$

Proof. For every $i \in I$, we have $\Phi_i \vDash \varphi$ which implies that $\Phi_i \cup \{\neg\varphi\}$ is inconsistent. Since L is compact it follows that there exists a finite subset $\Phi_i^0 \subseteq \Phi_i$ such that $\Phi_i^0 \cup \{\neg\varphi\}$ is inconsistent, i.e., $\Phi_i^0 \vDash \varphi$. Set $\psi_i := \bigwedge \Phi_i^0$ and let $\Psi := \{\, \psi_i \mid i \in I \,\}$. If the set

$$\Gamma := \{\varphi\} \cup \{\, \neg\psi \mid \psi \in \Psi \,\}$$

has a model $\mathfrak{J}$ then $\mathfrak{J} \vDash \varphi$ implies that $\mathfrak{J} \vDash \Phi_i$, for some i. In particular, we have $\mathfrak{J} \vDash \psi_i$ in contradiction to $\mathfrak{J} \vDash \neg\psi_i$.

Consequently, Γ is inconsistent and there exists a finite subset $\Psi_0 \subseteq \Psi$ such that

$$\{\varphi\} \cup \{\, \neg\psi \mid \psi \in \Psi_0 \,\}$$

is inconsistent. Set $\vartheta := \bigvee \Psi_0$. It follows that $\varphi \vDash \vartheta$.

Conversely, if $\mathfrak{J} \vDash \vartheta$ then $\mathfrak{J} \vDash \psi_i$, for some i, and $\psi_i \vDash \varphi$ implies that $\mathfrak{J} \vDash \varphi$. Hence, we also have $\vartheta \vDash \varphi$. Let $I_0 := \{ i \in I \mid \psi_i \in \Psi_0 \}$. Then we have

$$\varphi \equiv \vartheta \equiv \bigvee_{i \in I_0} \bigwedge \Phi_i^0 .$$

□

Lemma 1.5. *Let $L_0 \leq L_1$ be logics where L_0 satisfies* (B+) *and L_1 satisfies* (B) *and* (C). *If*

$$\mathfrak{A} \equiv_{L_0} \mathfrak{B} \quad \textit{implies} \quad \mathfrak{A} \equiv_{L_1} \mathfrak{B}$$

then $L_0 \equiv L_1$.

Proof. Let φ be an L_1-formula. Then

$$\varphi \equiv \bigvee \{ \bigwedge \mathrm{Th}_{L_0}(\mathfrak{J}) \mid \mathfrak{J} \in \mathrm{Mod}_{L_1}(\varphi) \} .$$

By Lemma 1.4, we can find finitely many interpretations $\mathfrak{J}_0, \dots, \mathfrak{J}_n$ and finite subsets $\Phi_i \subseteq \mathrm{Th}_{L_0}(\mathfrak{J}_i)$ such that

$$\varphi \equiv \bigwedge \Phi_0 \vee \dots \vee \bigwedge \Phi_n .$$

Since L_0 satisfies (B+) it follows that there is an L_0-formula $\psi \equiv \varphi$. □

Lemma 1.6. *Let L be an algebraic logic with* (B) *and $\forall \leq L$. If L has the compactness property then it has the finite occurrence property.*

Proof. Suppose that $\varphi \in L[\Sigma]$. Let $\Sigma' := \{ \xi' \mid \xi \in \Sigma \}$ be a disjoint copy of Σ and let $\mu : \Sigma \to \Sigma' : \xi \mapsto \xi'$ be the corresponding bijection. Consider the set of first-order formulae

$$\begin{aligned}\Phi := &\{ \forall \bar{x}(R\bar{x} \leftrightarrow R'\bar{x}) \mid R \in \Sigma \text{ a relation symbol} \} \\ &\cup \{ \forall \bar{x}(f\bar{x} = f'\bar{x}) \mid f \in \Sigma \text{ a function symbol} \} .\end{aligned}$$

Since $\forall \leq L$ there exists an equivalent set $\tilde{\Phi} \subseteq L[\Sigma \cup \Sigma']$ of L-formulae. If $\varphi' := L[\mu](\varphi)$ then

$$\tilde{\Phi} \cup \{\varphi\} \vDash \varphi' \quad \text{and} \quad \tilde{\Phi} \cup \{\varphi'\} \vDash \varphi .$$

By (c), we can find finite subsets $\tilde{\Phi}_0, \tilde{\Phi}_1 \subseteq \tilde{\Phi}$ such that

$$\tilde{\Phi}_0 \cup \{\varphi\} \vDash \varphi' \quad \text{and} \quad \tilde{\Phi}_1 \cup \{\varphi'\} \vDash \varphi\,.$$

Let Φ_0 and Φ_1 be the subsets of Φ corresponding to $\tilde{\Phi}_0$ and $\tilde{\Phi}_1$. Fix a finite signature Γ such that $\Phi_0, \Phi_1 \subseteq \mathrm{FO}[\Gamma \cup \Gamma']$. For a Σ-structure $\mathfrak{A}$, we denote by $\mathfrak{A}_+$ the $(\Sigma \cup \Gamma')$-expansion of $\mathfrak{A}$ where $(\xi')^{\mathfrak{A}_+} = \xi^{\mathfrak{A}}$, for all $\xi \in \Gamma$. We claim that

$$\mathfrak{A}|_\Gamma \equiv_L \mathfrak{B}|_\Gamma \quad \text{implies} \quad \mathfrak{A} \vDash \varphi \Leftrightarrow \mathfrak{B} \vDash \varphi\,,$$

for all Σ-structures $\mathfrak{A}$ and $\mathfrak{B}$. Suppose that $\mathfrak{A} \vDash \varphi$. Then $\mathfrak{A}_+ \vDash \tilde{\Phi}_0 \cup \varphi$, which implies that $\mathfrak{A}_+ \vDash \varphi'$. Note that $\mathfrak{A}|_\Gamma \equiv_L \mathfrak{B}|_\Gamma$ implies that $\mathfrak{A}_+|_{\Gamma'} \equiv_L \mathfrak{B}_+|_{\Gamma'}$. Consequently, it follows that $\mathfrak{B}_+ \vDash \varphi'$. Since $\mathfrak{B}_+ \vDash \tilde{\Phi}_1$ we obtain $\mathfrak{B} \vDash \varphi$, as desired.

For $\mathfrak{A} \in \mathrm{Str}[\Sigma]$, let $\Phi_{\mathfrak{A}} := \mathrm{Th}_{L[\Gamma]}(\mathfrak{A}|_\Gamma)$. By the above claim it follows that

$$\varphi \equiv \bigvee \{ \bigwedge \Phi_{\mathfrak{A}} \mid \mathfrak{A} \in \mathrm{Mod}_{L[\Sigma]}(\varphi) \}\,.$$

By Lemma 1.4, there are finitely many structures $\mathfrak{A}_0, \dots, \mathfrak{A}_n$ and finite subsets $\Psi_i \subseteq \Phi_{\mathfrak{A}_i}$ such that

$$\varphi \equiv \bigwedge \Psi_0 \vee \dots \vee \bigwedge \Psi_n \in L[\Gamma]\,. \qquad \square$$

2. *Hanf and Löwenheim numbers*

The Compactness Theorem and the Upward and Downward Löwenheim-Skolem Theorems are central results in first-order model theory. While the Compactness Theorem fails for many natural logics, we can generalise the Löwenheim-Skolem theorems to most of them. The *Hanf* and the *Löwenheim number* of a logic measure the extend to which a logic satisfies these theorems. For their definition we need the following notions.

Definition 2.1. Let L be an algebraic logic and $\Phi \subseteq L[\Sigma]$ a set of L-formulae.

(a) We say that Φ *pins down* a cardinal κ if there is a unary predicate $P \in \Sigma$ such that Φ has a model $\mathfrak{A}$ with $|P^{\mathfrak{A}}| = \kappa$ but Φ does not have models $\mathfrak{A}$ where $P^{\mathfrak{A}}$ has arbitrarily high cardinality.

(b) Φ *pins down* an ordinal α if there exists a binary relation $< \in \Sigma$ such that

- in every model of Φ the relation $<$ is a well-order of its field and
- there exists a model of Φ such that $<$ is of order type α.

Definition 2.2. Let L be an algebraic logic and κ a cardinal.

(a) The *Hanf number* $\mathrm{hn}_\kappa(L)$ of L is the supremum of all cardinals that can be pinned down by a set of L-formulae of size at most κ. If the supremum is undefined we set $\mathrm{hn}_\kappa(L) := \infty$.

(b) The *Löwenheim number* $\mathrm{ln}_\kappa(L)$ of L is the least cardinal λ such that every satisfiable set of L-formulae of size at most κ has a model of cardinality at most λ. If there is no such cardinal then we set $\mathrm{ln}_\kappa(L) := \infty$.

(c) The *well-ordering number* $\mathrm{wn}_\kappa(L)$ of L is the supremum of all ordinals α that can be pinned down by a set of L-formulae of size at most κ. If the supremum is undefined we set $\mathrm{wn}_\kappa(L) := \infty$. If $\mathrm{wn}_1(L) < \infty$ then L is called *bounded.*

(d) The *occurrence number* $\mathrm{occ}(L)$ of L is the least cardinal κ such that, for every signature Σ and all formulae $\varphi \in L[\Sigma]$, there exists a signature $\Sigma_0 \subseteq \Sigma$ and a formula $\psi \in L[\Sigma_0]$ such that $|\Sigma_0| \leq \kappa$ and $\psi \equiv \varphi$. Again, if there is no such cardinal then we set $\mathrm{occ}(L) := \infty$.

Remark. A logic L has (LSP) iff $\mathrm{ln}_1(L) = \aleph_0$.

Hanf and Löwenheim numbers for first-order logic were already computed in Theorems C2.4.12 and C2.3.7.

Theorem 2.3. $\mathrm{hn}_\kappa(\mathrm{FO}) = \aleph_0$ *and* $\mathrm{ln}_\kappa(\mathrm{FO}) = \kappa \oplus \aleph_0$, *for all* κ.

Theorem 2.4. $\mathrm{ln}_\kappa(\mathrm{FO}_{\kappa^+\aleph_0}) = \kappa$.

Lemma 2.5. *For every regular cardinal* κ, *we have* $\mathrm{wn}_1(\mathrm{FO}_{\kappa\aleph_0}) \geq \kappa$ *and* $\mathrm{occ}(\mathrm{FO}_{\kappa\aleph_0}) = \kappa^- := \sup\{\, \lambda \mid \lambda < \kappa \,\}$.

Proof. We have already seen in Lemma C1.1.7 that every ordinal $\alpha < \kappa$ is finitely $\mathrm{FO}_{\kappa\aleph_0}$-axiomatisable.

For the occurrence number note that $\mathrm{occ}(\mathrm{FO}_{\kappa\aleph_0}) < \kappa$ since each $\mathrm{FO}_{\kappa\aleph_0}$-formula has less than κ subformulae. Conversely, for every $\lambda < \kappa$, we have the formula

$$\bigwedge_{i<\lambda} P_i x$$

with λ different relation symbols. □

Lemma 2.6. $\mathrm{wn}_1(\mathrm{MSO}) = \infty$.

Proof. The example on page 484 shows that the class of all well-orders is finitely MSO-axiomatisable. □

In general the Hanf numbers of $\mathrm{FO}_{\kappa^+\aleph_0}$ depend on the model of set theory. In ZFC we can only prove the following bounds.

Theorem 2.7. $\beth_{\kappa^+} \leq \mathrm{hn}_1(\mathrm{FO}_{\kappa^+\aleph_0}) < \beth_{(2^\kappa)^+}$.

For the special case of $\mathrm{FO}_{\aleph_1\aleph_0}$ the exact value can be computed. (The proof is based on the study of Borel subsets of the type space and employs Corollary C4.2.5.)

Theorem 2.8 (Hanf). $\mathrm{hn}_1(\mathrm{FO}_{\aleph_1\aleph_0}) = \beth_{\omega_1}$.

(Note that $\mathrm{hn}_1(\mathrm{FO}_{\kappa^+\aleph_0}) = \mathrm{hn}_\kappa(\mathrm{FO}_{\kappa^+\aleph_0})$ since we can take conjunctions over sets of size κ.) We will prove the lower bound in Corollary 2.12 below. The computation of the upper bound is deferred to Corollary E7.1.13 (where we only prove the weaker statement that $\mathrm{hn}_1(\mathrm{FO}_{\kappa^+\aleph_0}) \leq \beth_{(2^\kappa)^+}$).

Lemma 2.9. *Let L be a logical system with $\forall\exists \leq L$.*

(a) *If* $\mathrm{hn}_\kappa(L) < \infty$ *then* $\mathrm{hn}_\kappa(L)$ *is a limit cardinal and a cardinal λ can be pinned down by a set $\Phi \subseteq L$ of size κ if and only if $\lambda < \mathrm{hn}_\kappa(L)$.*

(b) *If* $\mathrm{wn}_\kappa(L) < \infty$ *then* $\mathrm{wn}_\kappa(L)$ *is a limit ordinal and an ordinal α can be pinned down by a set $\Phi \subseteq L$ of size κ if and only if $\alpha < \mathrm{wn}_\kappa(L)$.*

Proof. (a) Let Φ be a set of size at most κ pinning down the cardinal μ via the relation symbol P. We construct a set Ψ of the same size pinning down μ^+. Let S be the set of sorts appearing in Φ. Choose new binary relation symbols $<$ and R_s, for $s \in S$, a new unary relation symbol Q, and a new binary function symbol f. Ψ consists of formulae expressing the following properties.

- $<$ is a linear order of the set Q.
- For every $u \in Q$, the set $R(u) := \{ x \mid \langle u, x \rangle \in R_s \text{ for some } s \in S \}$ induces a substructure satisfying Φ.
- For every $u \in Q$, the function $x \mapsto f(u, x)$ is an injective map from $\Downarrow u$ into $R(u) \cap P$.

It follows that Ψ has a model where $<$ has the order type μ^+. To see that $|Q|$ cannot become arbitrarily large let λ be some cardinal such that Φ has no models with $|P| = \lambda$. Given any model of Ψ fix a strictly increasing cofinal map $f : \alpha \to Q$. By the third condition above we have $|\Downarrow f(i)| < \lambda$, for all $i < \alpha$. Consequently,

$$Q = \bigcup_{i<\alpha} \Downarrow f(i)$$

implies that $|Q| \leq \lambda$.

(b) The statement that $\mathfrak{A} = \langle A, \leq \rangle$ is a linear order with exactly $n < \omega$ elements can be expressed in $\forall$. Since $\forall\exists \leq L$ it follows that $\mathrm{wn}_\kappa(L) \geq \omega$.

To prove the claim we show that if α is pinned down by some $\Phi \subseteq L$ of size $|\Phi| \leq \kappa$ then so is $\alpha + 1$ and every ordinal $\beta \leq \alpha$.

Suppose that $\Phi \subseteq L$ pins down $\alpha \geq \omega$ via the relation symbol $<$. Let P be a new unary relation symbol and $\sqsubset$ a new binary one.

We can construct a set $\Phi \cup \{\psi\}$ pinning down every ordinal $\beta \leq \alpha$ via $\sqsubset$ by defining

$$\psi := \forall x \forall y (x \sqsubset y \leftrightarrow (Px \wedge Py \wedge x < y)),$$

which expresses that $\sqsubset \, = \, <|_P$.

Similarly, we can define a set $\Phi \cup \{\psi\}$ pinning down $\alpha + 1$ via $\sqsubset$ by defining

$$\psi := \forall x \forall y \big[x \sqsubset y \leftrightarrow \big[(x < y \wedge \exists z (z < x)) \\ \vee (y < x \wedge \neg \exists z (z < y)) \big] \big] ,$$

which states that $\sqsubset$ is the order obtained from $<$ by moving the least element to the end. □

Under very general conditions, we can show that a logical system L has a Hanf number and a Löwenheim number.

Proposition 2.10. *Let L be an algebraic logic such that $L[\Sigma]$ is a set, for all Σ. If $\mathrm{occ}(L) < \infty$ then we have $\mathrm{hn}_\kappa(L) < \infty$ and $\mathrm{ln}_\kappa(L) < \infty$, for all κ.*

Proof. Set $\mu := \kappa \otimes \mathrm{occ}(L)$ and fix an *universal* signature Σ of size μ, that is, Σ is S-sorted, for some set of sorts with $|S| = \mu$, and Σ contains, for all sorts $\bar{s}$ and t, μ relation symbols of type $\bar{s}$ and μ function symbols of type $\bar{s} \to t$. It is sufficient to consider sets $\Phi \subseteq L[\Sigma]$ since every signature of size μ can be embedded into Σ and, by definition of a logical system, L-formulae are invariant under such changes of the signature.

For every set $\Phi \subseteq L[\Sigma]$ of size $|\Phi| \leq \kappa$ and every unary predicate $P \in \Sigma$, we define two cardinals $\nu_{\Phi,P}$ and λ_Φ as follows. If Φ has models $\mathfrak{A}$ where $P^{\mathfrak{A}}$ can be arbitrarily large then we set $\nu_{\Phi,P} := 0$. Otherwise, let $\nu_{\Phi,P}$ be the least cardinal such that Φ has only models $\mathfrak{A}$ with $|P^{\mathfrak{A}}| \leq \nu_\Phi$. Similarly, if Φ is satisfiable then we set

$$\lambda_\Phi := \min \{ |A| \mid \mathfrak{A} \vDash \Phi \} .$$

Otherwise, we let λ_Φ undefined. It follows that

$$\mathrm{hn}_\kappa(L) = \sup \{ \nu_{\Phi,P} \mid P \in \Sigma,\ \Phi \subseteq L[\Sigma] \text{ of size } |\Phi| \leq \kappa \} ,$$

and $$\mathrm{ln}_\kappa(L) = \sup \{ \lambda_\Phi \mid \Phi \subseteq L[\Sigma] \text{ satisfiable and of size } |\Phi| \leq \kappa \} .$$

Note that the supremum on the right-hand side exists since, by the Axiom of Replacement, it is taken over a set of cardinals. □

Theorem 2.11. *Let L be a regular logical system with* FO $\leq L$ *such that, for every ordinal* $\alpha < \mathrm{wn}_\kappa(L)$, *there exists a set* $\Phi_\alpha \subseteq L[\Sigma_\alpha]$ *of size* $|\Phi_\alpha| < \kappa$ *pinning down* α *in a model of size at most* $\mathrm{hn}_\kappa(L)$. *Then we have*

$$\mathrm{hn}_\kappa(L) \geq \beth_{\mathrm{wn}_\kappa(L)}(\lambda)\,, \quad \text{for all } \lambda < \mathrm{hn}_\kappa(L)\,.$$

Proof. Let X be a set of size λ. We define inductively a variant of the cumulative hierarchy by

$$\begin{aligned} P_0(X) &:= X\,, \\ P_{\alpha+1}(X) &:= \mathcal{P}(P_\alpha(X))\,, \\ P_\delta(X) &:= \bigcup_{\alpha<\delta} P_\alpha(X)\,, \quad \text{for limit ordinals } \delta\,. \end{aligned}$$

Then $|P_\alpha(X)| = \beth_\alpha(\lambda)$.

Since $\lambda < \mathrm{hn}_\kappa(L)$ we can find a set $\Psi \subseteq L[\Gamma]$ of size $|\Psi| \leq \kappa$ pinning down λ via a predicate Q. Suppose that Σ_α is S-sorted and Γ is T-sorted with $S \cap T = \varnothing$ and let $p \notin S \cup T$ be a new sort. Choose new unary predicates O, U, a binary relation symbol E, unary functions ρ, ζ, and a constant 0. We define a set Θ_α of formulae that is meant to describe a structure $\mathfrak{A}$ of the following form. We have $\mathfrak{A}|_S \vDash \Phi_\alpha$ and $\mathfrak{A}|_T \vDash \Psi$. Furthermore, $U \subseteq A_p \subseteq P_\alpha(U)$ for the ordinal α encoded in $\mathfrak{A}|_S$. The relation E is the membership relation of sets, $\rho : A_p \to O$ maps every set in $P_\beta(U)$ to the ordinal β, and $\zeta : Q \to U$ is a bijection. Formally, Θ_α consists of the union $\Psi \cup \Phi_\alpha$ together with the following formulae.

- The domains with sort T form a model of Φ_α and 0 is the least element of <.

  $$\forall x(Ox \leftrightarrow x \leq x)$$
  $$(\forall x.Ox)(0 \leq x)$$

- $\zeta : Q \to U$ is a bijection and ρ maps A_p to the field of <.

  $$\forall x(Qx \leftrightarrow U\zeta x)$$
  $$\forall x \forall y(\zeta x = \zeta y \to x = y)$$
  $$\forall x O\rho x$$

(In the last formula x is of sort p.)

- $\rho^{-1}(\alpha) \subseteq P_\alpha(U)$ and E is the element relation.

$$\begin{aligned}
&\forall x(Ux \leftrightarrow \rho x = 0) \\
&(\forall x.\neg Ux)[\forall y.\neg Uy)(\forall z(Ezx \leftrightarrow Ezy) \to x = y] \\
&\forall x(\forall u.Ou)\big[\rho x = u \leftrightarrow [(\forall y.Eyx)(\rho y < \rho x) \\
&\qquad\qquad\qquad \wedge (\forall v.v < u)(\exists y.Eyx)(\rho y \geq v)]\big]
\end{aligned}$$

If $\mathfrak{A}$ is a model of Θ_α then $<^{\mathfrak{A}}$ is a well-order of type $\beta < \mathrm{wn}_\kappa(L)$ and there exists an injective function $A \to P_\beta(U^{\mathfrak{A}})$. Consequently,

$$|A_p| \leq \beth_\beta(|U^{\mathfrak{A}}|)\,.$$

Since Ψ pins down a cardinal we further have

$$|U^{\mathfrak{A}}| = |Q^{\mathfrak{A}}| \leq \mathrm{hn}_\kappa(L)\,.$$

Therefore, Θ_α does not have models where A_p is arbitrarily large, but it does have a model $\mathfrak{A}$ with $|A_p| = \beth_\alpha(\lambda)$. □

We have shown in Lemma C1.1.7 that every ordinal $\alpha < \kappa^+$ can be defined in $\mathrm{FO}_{\kappa^+\aleph_0}$. Consequently, we obtain the following lower bound on the Hanf number.

Corollary 2.12. $\mathrm{hn}_\kappa(\mathrm{FO}_{\kappa^+\aleph_0}) \geq \beth_{\kappa^+}$

Lemma 2.13. *Suppose that L is a regular logical system with* FO $\leq L$. *Then L is countably compact if and only if* $\mathrm{wn}_{\aleph_0}(L) = \omega$.

Proof. A standard compactness argument shows that if L is countably compact and $\Phi \subseteq L$ has a model such that $<$ is of order type ω then there also is a model where $<$ contains an infinite descending chain. Consequently, (CC) implies $\mathrm{wn}_{\aleph_0}(L) \leq \omega$.

For the converse, assume that there exists a countable inconsistent set $\{\, \varphi_n \mid n < \omega \,\} \subseteq L$ every finite subset of which is satisfiable. By

Lemma 2.9 (b) we can prove that $\mathrm{wn}_{\aleph_0}(L) > \omega$ by constructing a countable set $\Phi \subseteq L$ pinning down ω.

Let S be the set of sorts appearing in some φ_n and choose new binary relation symbols $<$ and R_s, for $s \in S$. The set Φ consists of the following statements all of which can be expressed in first-order logic:

- $<$ is a linear ordering of its field.
- For all elements a of the field of $<$ there is some element b with $\langle a, b\rangle \in R$.
- If there are at least n elements $<$-less than a then the set $\{ b \mid \langle a, b\rangle \in R_s$ for some $s \in S \}$ induces a substructure satisfying φ_n.

It follows that if $\mathfrak{A}$ is a model of Φ then every element in the field of $<^{\mathfrak{A}}$ has only finitely many elements below. Consequently, Φ pins down all ordinals $\alpha \leq \omega$. □

3. The Theorem of Lindström

We have seen that first-order logic has many pleasant properties like compactness and the Löwenheim-Skolem property. On the other hand, its expressive power is rather restricted as far as certain aspects like counting and recursion are concerned. The question naturally arises of whether there is a stronger logic that shares the good properties of first-order logic. Surprisingly, it turns out that one can prove that such a logic does not exist.

In many of the following proofs we consider a structure containing two other structures, say, specified by unary predicates P and Q. We use additional relations to encode a back-and-forth systems between these substructures.

Definition 3.1. Suppose that Σ and Γ are signatures and $\mu : \Sigma \to \Gamma$ is an isomorphism of $\mathfrak{Sig}$. Let $\mathfrak{A}$ be a $(\Sigma \cup \Gamma)$-structure and $P, Q \subseteq A$ subsets of A.

(a) A *partial isomorphism modulo μ from P to Q* is a function $p : \bar{a} \mapsto \bar{b}$ with $\bar{a} \subseteq P$ and $\bar{b} \subseteq Q$ such that, for all term-reduced atomic first-order formulae $\varphi(\bar{x})$, we have

$$\mathfrak{A} \models \varphi(\bar{a}) \quad \text{iff} \quad \mathfrak{A} \models \mathrm{FO}[\mu](\varphi)(\bar{b})\,.$$

(b) A *pseudo back-and-forth system* (*modulo μ from P to Q*) is a sequence $(I_\alpha)_{\alpha \in U}$ where

- each I_α is a set of partial isomorphisms modulo μ from P to Q,
- U is a nonempty linear order such that every element $\alpha \in U$ has an immediate successor $\alpha + 1$, except possibly for the last element,
- we have $I_\delta := \bigcap_{\alpha<\delta} I_\alpha$, for elements $\delta \in U$ without immediate predecessor, and
- every $I_{\alpha+1}$ has the back-and-forth property restricted to P and Q with respect to I_α, that is,
  - if $\bar{a} \mapsto \bar{b} \in I_{\alpha+1}$ and $c \in P$ then there is some $d \in Q$ with $\bar{a}c \mapsto \bar{b}d \in I_\alpha$, and
  - if $\bar{a} \mapsto \bar{b} \in I_{\alpha+1}$ and $d \in Q$ then there is some $c \in P$ with $\bar{a}c \mapsto \bar{b}d \in I_\alpha$.

(c) We say that a tuple $\langle U, <, \bar{P}, \bar{Q}, I, F, \bar{G} \rangle$ *encodes* a pseudo back-and-forth system $(I_\alpha)_{\alpha \in X}$ modulo μ from P to Q if there exist a finite set of sorts S and sorts u and f such that

- $\bar{P} = (P_s)_{s \in S}$, $\bar{Q} = (Q_s)_{s \in S}$, and $\bar{G} = (G_s)_{s \in S}$,
- $P = \bigcup_s P_s$ and $Q = \bigcup_s Q_s$,
- $U \subseteq A_u$, $\quad F \subseteq A_f$, $\quad P_s \subseteq A_s$, $\quad Q_s \subseteq A_{\mu(s)}$,
  $I \subseteq U \times F$, $\quad < \subseteq U \times U$, $\quad G_s \subseteq F \times P_s \times Q_s$,
- there exists an isomorphism $\iota : \langle U, < \rangle \cong \langle X, < \rangle$,
- there exists a bijection $\pi : F \to \bigcup_\alpha I_\alpha$,
- $I = \{ \langle u, p \rangle \in U \times F \mid \pi p \in I_{\iota u} \}$,
- $G_s = \{ \langle p, a, b \rangle \in F \times P_s \times Q_s \mid (\pi p)(a) = b \}$.

Lemma 3.2. *Suppose that Σ and Γ are finite signatures and $\mu : \Sigma \to \Gamma$ an isomorphism of $\mathfrak{Sig}$. There exists a first-order formula*

$$\beta_\mu(U, <, \bar{P}, \bar{Q}, I, F, \bar{G})$$

that holds if and only if $\langle U, <, \bar{P}, \bar{Q}, I, F, \bar{G}\rangle$ encodes a pseudo back-and-forth system modulo μ from P to Q.

Proof. We have to express the following properties:

(a) $\langle U, <\rangle$ is a nonempty linear order and every element has an immediate successor, except for the last one.

$$\begin{aligned}
&\exists u Uu \\
&\forall u \forall v (u < v \to Uu \wedge Uv) \\
&\forall u (\neg u < u) \\
&\forall u \forall v \forall w (u < v \wedge v < w \to u < w) \\
&(\forall u.Uu)(\forall v.Uv)(u < v \vee u = v \vee u > v) \\
&\forall u [\exists v (u < v) \to \exists v (u < v \wedge \neg \exists w (u < w \wedge w < v))]
\end{aligned}$$

(b) $G_s \subseteq F \times P_s \times Q_s$ encodes a set of partial isomorphisms modulo μ.

$$\begin{aligned}
&\forall p \forall a \forall b (G_s pab \to Fp \wedge P_s a \wedge Q_s b) \\
&\forall p \forall a_0 \forall a_1 \forall b_0 \forall b_1 [G_s p a_0 b_0 \wedge G_s p a_1 b_1 \to (a_0 = a_1 \leftrightarrow b_0 = b_1)]
\end{aligned}$$

For all n-ary relation symbols $R \in \Sigma$,

$$\forall p \forall \bar{a} \forall \bar{b} [G_{s_0} p a_0 b_0 \wedge \cdots \wedge G_{s_{n-1}} p a_{n-1} b_{n-1} \to (R\bar{a} \leftrightarrow \mu(R)\bar{b})] .$$

For all n-ary function symbols $f \in \Sigma$,

$$\begin{aligned}
\forall p \forall \bar{a} \forall c \forall \bar{b} \forall d [G_{s_0} p a_0 b_0 \wedge \cdots \wedge G_{s_{n-1}} p a_{n-1} b_{n-1} \wedge G_t pcd \to \\
(f\bar{a} = c \leftrightarrow \mu(f)\bar{b} = d)] .
\end{aligned}$$

(c) $I \subseteq U \times F$ encodes a sequence of nonempty sets with the back-and-forth property.

$$\forall u \forall p(Iup \to Uu \wedge Fp)$$
$$\forall u \exists p Iup$$
$$\forall u \forall v \forall p[Iup \wedge v < u \to (\forall c.P_s c)\exists d \exists q \eta_s)]$$
$$\forall u \forall v \forall p[Iup \wedge v < u \to (\forall d.Q_s d)\exists c \exists q \eta_s)]$$

where $\eta_s := Ivq \wedge G_s qcd \wedge \bigwedge_t \forall a \forall b(G_t pab \to G_t qab)$. □

Lemma 3.3. *Let Σ and Γ be finite signatures and $\mu : \Sigma \to \Gamma$ an isomorphism of $\mathfrak{Sig}$. Let $\mathfrak{A}$ be a $(\Sigma \cup \Gamma)$-structure and $P, Q \subseteq A$. Suppose that P and Q induce substructures of, respectively, $\mathfrak{A}|_\Sigma$ and $\mathfrak{A}|_\Gamma$.*

If there exists a pseudo back-and-forth system $(I_\alpha)_{\alpha \in U}$ modulo μ from P to Q where U is not well-ordered then

$$\mathfrak{A}|_\Sigma|_P \cong_\infty \mathfrak{A}|_\Gamma|_Q|_\mu .$$

Proof. Fix an infinite descending sequence $\alpha_0 > \alpha_1 > \dots$ in U. We claim that $J = \bigcup_n I_{\alpha_n}$ has the back-and-forth property with respect to itself. If $p \in J$ then $p \in I_{\alpha_n}$, for some n. Hence, for every $c \in P$ or $d \in Q$, we can find a suitable extension $q \in I_{\alpha_{n+1}} \subseteq J$ with, respectively, $c \in \operatorname{dom} q$ or $d \in \operatorname{rng} q$. Consequently,

$$J : \mathfrak{A}|_\Sigma|_P \cong_\infty \mathfrak{A}|_\Gamma|_Q|_\mu .$$ □

Definition 3.4. Let L and L' be logical systems and $\varphi, \psi \in L[s]$.

(a) φ and ψ are *contradictory* if

$$\operatorname{Mod}_L(\varphi) \cap \operatorname{Mod}_L(\psi) = \varnothing .$$

(b) A formula $\chi \in L'[t]$ *separates* φ from ψ if

$$\operatorname{Mod}_{L'}(\chi) \supseteq \operatorname{Mod}_L(\varphi) \quad \text{and} \quad \operatorname{Mod}_{L'}(\chi) \cap \operatorname{Mod}_L(\psi) = \varnothing .$$

We start by investigating logical systems containing first-order logic that have the Löwenheim-Skolem property. First, we show that if the logic is strictly more expressive than first-order logic then it can express finiteness.

Lemma 3.5. *Let L be a weakly regular logical system with* $\mathrm{FO}^{\circ} \leq L$ *and* (LSP).

If there are contradictory formulae $\varphi, \psi \in L[\Sigma]$ that are not separated by any first-order formula $\chi \in \mathrm{FO}[\Sigma]$ then there exists a signature Γ, a unary predicate $U \in \Gamma$, and a formula $\vartheta \in L[\Gamma]$ satisfying the following conditions:

(1) *If $\mathfrak{A} \vDash \vartheta$ then $U^{\mathfrak{A}}$ is finite and nonempty.*

(2) *For every $0 < n < \omega$, there exists a model $\mathfrak{A} \vDash \vartheta$ with $|U^{\mathfrak{A}}| = n$.*

Proof. For a contradiction, suppose that $\varphi, \psi \in L[\Sigma]$ are not separated by any first-order formula but there is no formula ϑ satisfying (1) and (2). By (FOP), we may assume that Σ is finite. We proceed in several steps.

(a) First, we prove that every formula $\chi \in L[\Gamma]$ that is not equivalent to a first-order formula has a model of cardinality $\aleph_0$. Let χ be such a formula. If χ has infinite models then choose a new unary function symbol $f \notin \Gamma$ and consider the formula

$$\chi' := \chi \wedge \text{“}f \text{ is injective but not surjective”}\,.$$

Since χ has infinite models it follows that χ' is satisfiable. By (LSP), there exists a countable model of χ'. Since there are no finite models of χ' it follows that this model is countably infinite.

It remains to consider the case that χ has only finite models. By (FOP), we may assume that Γ is finite. Thus, for every $n < \omega$, there are only finitely many non-isomorphic Γ-structures $\mathfrak{A}$ of cardinality n and each of them can be axiomatised by a first-order formula $\eta_{\mathfrak{A}}$. Consequently, χ must have models of arbitrarily large finite cardinality since, otherwise, χ would be equivalent to a finite disjunction of first-order formulae $\eta_{\mathfrak{A}}$. If $U \notin \Gamma$ is a new unary relation symbol then the formula

$$\vartheta := \chi \wedge \exists x U x$$

satisfies (1) and (2). A contradiction.

(b) Second, we prove that, for every $n < \omega$, there are countably infinite structures $\mathfrak{A}_n$ and $\mathfrak{B}_n$ such that

$$\mathfrak{A}_n \vDash \varphi\,, \quad \mathfrak{B}_n \vDash \psi\,, \quad \text{and} \quad \mathfrak{A}_n \equiv_n \mathfrak{B}_n\,.$$

Let $\eta^n_{\mathfrak{A}}$ be the Hintikka-formula of $\mathfrak{A}$ of quantifier-rank n and set

$$\chi_n := \bigvee \{\, \eta^n_{\mathfrak{A}} \mid \mathfrak{A} \vDash \varphi \,\}\,.$$

Since Σ is finite we have $\eta^n_{\mathfrak{A}} \in \mathrm{FO}[\Sigma]$ and there are only finitely many different Hintikka-formulae of quantifier rank n. Consequently, $\chi_n \in \mathrm{FO}[\Sigma]$.

Since $\varphi \vDash \chi_n$ we have $\mathrm{Mod}(\varphi) \subseteq \mathrm{Mod}(\chi_n)$. As φ and ψ cannot be separated it follows that

$$\mathrm{Mod}(\chi_n) \cap \mathrm{Mod}(\psi) \neq \varnothing\,.$$

Hence, $\psi \wedge \chi_n$ is satisfiable and it is not equivalent to any first-order formula. By (a), there exists a countably infinite model $\mathfrak{B}_n \vDash \psi \wedge \chi_n$. In particular, we have $\mathfrak{B}_n \vDash \eta_{\mathfrak{A}}$, for some $\mathfrak{A} \vDash \varphi$. Moreover, $\varphi \wedge \eta_{\mathfrak{A}}$ is satisfiable and not equivalent to any first-order formula. Thus, by (a), we can find a countably infinite model $\mathfrak{A}_n \vDash \varphi \wedge \eta_{\mathfrak{A}}$. Note that $\mathfrak{A}_n \equiv_n \mathfrak{B}_n$ because both $\mathfrak{A}_n$ and $\mathfrak{B}_n$ satisfy $\eta_{\mathfrak{A}}$. Hence, $\mathfrak{A}_n$ and $\mathfrak{B}_n$ have the desired properties.

(c) Finally, we derive a contradiction as follows. Let Σ' be a disjoint copy of Σ and let $\mu : \Sigma \to \Sigma'$ be the corresponding bijection. If $\mathfrak{C}$ is a model of the L-formula

$$\begin{aligned}
\vartheta := {} & \varphi \wedge L[\mu](\psi) \\
& \wedge \beta_\mu(U, <, \bar{P}, \bar{Q}, I, F, \bar{G}) \\
& \wedge \textstyle\bigwedge_s \forall x(P_s x \wedge Q_s x) \\
& \wedge \exists x(\forall y.Uy)(y = x \vee x < y) \\
& \wedge \exists x(\forall y.Uy)(y = x \vee y < x) \\
& \wedge \forall x[\exists y(y < x) \to (\exists y.y < x)\neg\exists z(y < z \wedge z < x)]
\end{aligned}$$

then $\langle U, <, \bar{P}, \bar{Q}, I, F, \bar{G} \rangle$ encodes a pseudo back-and-forth system modulo μ from C to C where $\langle U, < \rangle$ is a discrete linear order with a least and greatest element. Furthermore, the Σ-reduct of $\mathfrak{C}$ satisfies φ and its Σ'-reduct satisfies $\mu(\psi)$.

For every $n < \omega$, we can find a model $\mathfrak{C}_n$ of ϑ with $|U^{\mathfrak{C}_n}| = n + 1$ as follows. By (b), there are countably infinite structures $\mathfrak{A}_n$ and $\mathfrak{B}_n$ with $\mathfrak{A}_n \equiv_n \mathfrak{B}_n$ such that $\mathfrak{A}_n \vDash \varphi$ and $\mathfrak{B}_n \vDash \psi$. Since $|A_n| = |B_n|$ we may w.l.o.g. assume that $A_n = B_n$. We form the structure $\mathfrak{C}_n$ with universe $A_n = B_n$ where, for every $\xi \in \Sigma$, we have two relations or functions

$$\xi^{\mathfrak{C}_n} := \xi^{\mathfrak{A}_n} \quad \text{and} \quad \mu(\xi)^{\mathfrak{C}_n} := \xi^{\mathfrak{B}_n}.$$

Hence, the Σ-reduct of $\mathfrak{C}_n$ equals $\mathfrak{A}_n$ and its Σ'-reduct equals $\mu(\mathfrak{B}_n)$. Furthermore, since $\mathfrak{A}_n \equiv_n \mathfrak{B}_n$ we can add relations $U, <, I, \bar{P}, \bar{Q}, F, \bar{G}$ encoding some back-and-forth system modulo μ where $|U| = n + 1$.

Consequently, the formula $\vartheta \wedge |U| = n$ is satisfiable, for all $0 < n < \omega$. This concludes the proof of (2). For (1), assume that ϑ has models where U is infinite. If f is a new unary function symbol then it follows that the formula

$$\vartheta' := \vartheta \wedge \text{“} f \text{ is injective but not surjective”}$$

is satisfiable. By (LSP), ϑ' has a countable model $\mathfrak{C}$. Let u_0 be the greatest element of $U^{\mathfrak{C}}$. Since every element of $U^{\mathfrak{C}}$ has an immediate predecessors we obtain an infinite descending sequence $u_0 > u_1 > \dots$. Hence, Lemma 3.3 implies that

$$\mathfrak{A} \cong \mathfrak{C}|_{\Sigma} \cong_{\infty} \operatorname{Str}[\mu](\mathfrak{C}|_{\Sigma'}) \cong \mathfrak{B}.$$

Since $\mathfrak{A}$ and $\mathfrak{B}$ are countable structures it follows by Corollary C4.4.11 that $\mathfrak{A} \cong \mathfrak{B}$. But $\mathfrak{A} \vDash \varphi$ and $\mathfrak{B} \vDash \psi$. This contradicts the fact that φ and ψ are contradictory. □

Lemma 3.6. *If L is a regular logic with* $\mathrm{FO}^{\circ} \leq L$ *then* (LSP) *implies* (KP).

Proof. For a contradiction, suppose that L is a regular logic with the Löwenheim-Skolem property but there are structures $\mathfrak{A} \cong_\infty \mathfrak{B}$ such that $\mathfrak{A} \vDash \varphi$ and $\mathfrak{B} \nvDash \varphi$, for some L-formula φ. By (FOP) we may assume that the signature of φ is finite.

Let $U, <, \bar{P}, \bar{Q}, I, F, \bar{G}$ be new relation symbols. By Lemma 3.2, there exists a formula $\beta_{\mathrm{id}}(U, <, \bar{P}, \bar{Q}, I, F, \bar{G})$ saying that $\langle U, <, \bar{P}, \bar{Q}, I, F, \bar{G} \rangle$ encodes a pseudo back-and-forth system from $P := \bigcup_s P_s$ to $Q := \bigcup_s Q_s$. The formula

$$\chi := \beta_{\mathrm{id}} \wedge \varphi^{(\bar{P})} \wedge (\neg\varphi)^{(\bar{Q})} \wedge (\forall x.Ux)\exists y(y < x)$$

has a model $\mathfrak{C}$ where $\langle U, < \rangle$ is an arbitrary discrete order without least element, $\mathfrak{C}|_P \cong \mathfrak{A}$, and $\mathfrak{C}|_Q \cong \mathfrak{B}$. (Note that, if there exists a pseudo back-and-forth system $(I_u)_{u \in U}$ from P to Q and the ordering U has arbitrarily large finite increasing chains then P and Q are closed under the functions of Σ. Hence, the formula implies that the sets P and Q induce substructures of $\mathfrak{C}|_\Sigma$.)

By (LSP), it follows that χ has a countable model $\mathfrak{C}$. Since $\langle U^{\mathfrak{C}}, <^{\mathfrak{C}} \rangle$ is not well-ordered we have $\mathfrak{C}|_P \cong_\infty \mathfrak{C}|_Q$, by Lemma 3.3. Because these substructures are countable it follows that $\mathfrak{C}|_P \cong \mathfrak{C}|_Q$. But $\mathfrak{C}|_P \vDash \varphi$ and $\mathfrak{C}|_Q \nvDash \varphi$. Contradiction. □

Lemma 3.7. *Let L be a weakly regular logical system with* $\mathrm{FO}^\circ \leq L$.

If L is countably compact and L has the Löwenheim-Skolem property then every pair of contradictory L-formulae can be separated by some FO°*-formula.*

Proof. Suppose that L satisfies (LSP) but there exists a pair of contradictory L-formulae that cannot be separated by any first-order formula. By Lemma 3.5, there exists a formula $\vartheta \in L[\Gamma]$ and a unary predicate $U \in \Gamma$ such that in models $\mathfrak{A}$ of ϑ the set $U^{\mathfrak{A}}$ can have any finite cardinality, but no infinite one. Let $\varphi_n \in L[\Gamma]$ be the L-formula equivalent to the first-order formula

$$\exists x_0 \cdots \exists x_{n-1} \Big(\bigwedge_i U x_i \wedge \bigwedge_{i \neq k} x_i \neq x_k \Big)$$

which expresses that $|U| \geq n$. By construction, the set

$$\{\vartheta\} \cup \{ \varphi_n \mid n < \omega \}$$

is inconsistent, but each of its finite subsets is satisfiable. Consequently, L is not countably compact. □

Combining the preceding technical lemmas we can prove that there does not exist a proper extension of first-order logic that has the Löwenheim-Skolem property and that is countably compact.

Theorem 3.8 (Lindström). *Let L be a weakly regular logical system with* (B) *and* $\mathrm{FO}^\circ \leq L$. *If L has the Löwenheim-Skolem property and L is countably compact then* $L \equiv \mathrm{FO}^\circ$.

Proof. Let $\varphi \in L[\Sigma]$. By Lemma 3.7, there exists a first-order formula χ separating φ from $\neg\varphi$. It follows that $\mathrm{Mod}(\chi) = \mathrm{Mod}(\varphi)$. □

We conclude this section with several variants of the Theorem of Lindström where (LSP) and (CC) are replaced by other properties.

Lemma 3.9. *Let L be a regular logical system with* $\mathrm{FO}^\circ < L$. *If L has the Karp property then there exists a satisfiable formula $\varphi(U, <) \in L$ such that, for all models $\mathfrak{A} \vDash \varphi$, we have*

$$\langle U^{\mathfrak{A}}, <^{\mathfrak{A}} \rangle \cong \langle \omega, < \rangle .$$

Proof. Fix a formula $\varphi \in L[\Sigma]$ that is not equivalent to any first-order formula. By (FOP), we may assume that Σ is finite. For every $n < \omega$, there are structures $\mathfrak{A}_n \equiv_n \mathfrak{B}_n$ such that

$$\mathfrak{A}_n \vDash \varphi \quad \text{and} \quad \mathfrak{B}_n \nvDash \varphi .$$

Let $U, <, \bar{P}, \bar{Q}, I, F, \bar{G} \notin \Sigma$ be new relation symbols where U is unary, $<, \bar{P}, \bar{Q}, F$ are binary, I is ternary, and $\bar{G}$ are of arity four. We modify the formula $\beta_{\mathrm{id}}(U, <, \bar{P}, \bar{Q}, I, F, \bar{G})$ of Lemma 3.2 as follows. Let x be a variable not occurring in β_{id} and set

$$\begin{aligned}&\alpha(x, U, <, \bar{P}, \bar{Q}, I, F, \bar{G}) :=\\&\quad \beta_{\mathrm{id}}(U, <, (P_s x_-)_s, (Q_s x_-)_s, Ix_{--}, Fx_-, (G_s x_{---})_s),\end{aligned}$$

that is, we add x as new argument to every atom containing P_s, Q_s, I, F, G_s. The formula α states that these relations encode a sequence of pseudo back-and-forth systems indexed by x. Define

$$\chi := \exists x.Ux \wedge (\forall x.Ux)\big[\vartheta(x) \wedge \alpha(x) \wedge \varphi^{(\bar{P}x_-)} \wedge (\neg\varphi)^{(\bar{Q}x_-)}\big],$$

where

$$\begin{aligned}\vartheta(x) := {}& \exists y(x < y)\\&\wedge (\exists y(y < x) \rightarrow (\exists y.y < x)\neg\exists z(y < z \wedge z < x))\end{aligned}$$

says that x has a successor and, if it is not the first element then it also has an immediate predecessor. The formula χ says that

- U is a nonempty discrete linear order without last element,
- for every $u \in U$, there is a pseudo back-and-forth system $(I_\alpha)_{\alpha<u}$ from $A_u := \{ a \mid \langle u, a\rangle \in \bigcup_s P_s \}$ to $B_u := \{ b \mid \langle u, b\rangle \in \bigcup_s Q_s \}$ of length $\downarrow u$,
- A_u induces a substructure that satisfies φ while B_u induces a substructure that does not satisfy φ.

Consequently, χ has a model where $\langle U, <\rangle \cong \langle \omega, <\rangle$ and the substructures induced by A_n and B_n, for $n < \omega$, are isomorphic to $\mathfrak{A}_n$ and $\mathfrak{B}_n$, respectively. Let $\mathfrak{C}$ be an arbitrary model of χ. We have to show that the order type of $\langle U^{\mathfrak{C}}, <^{\mathfrak{C}}\rangle$ is ω. Suppose otherwise. Then there exists some element $u \in U$ such that $\downarrow u$ is infinite. Since every element except for the first one has an immediate predecessor it follows that $<$ is not a well-order. By Lemma 3.3, we can conclude that $\mathfrak{C}|_{A_u} \cong_\infty \mathfrak{C}|_{B_u}$. Hence, $\mathfrak{C}|_{A_u} \vDash \varphi$ and $\mathfrak{C}|_{B_u} \nvDash \varphi$ contradicts (KP). $\square$

Theorem 3.10. *Let L be a regular logical system with $\mathrm{FO}^\circ \leq L$. If L has the Karp property and L is countably compact then $L \equiv \mathrm{FO}^\circ$.*

Proof. The claim follows immediately from the Lemma 3.9 since $\langle \omega, < \rangle$ cannot be axiomatised in a countably compact logic. □

For the next theorem we need the following variant of the Diagram Lemma.

Lemma 3.11. *Suppose that L is a regular logical system such that L is compact and $\mathrm{FO} \leq L$. Let $\mathfrak{A}$ be a structure and $\Phi \subseteq L$.*

There exists an elementary extension $\mathfrak{B} \succeq_{\mathrm{FO}} \mathfrak{A}$ with $\mathfrak{B} \vDash \Phi$ if and only if $\mathrm{Th}_{\mathrm{FO}}(\mathfrak{A}) \cup \Phi$ is satisfiable.

Proof. $(\Rightarrow)$ Clearly, $\mathfrak{B} \succeq_{\mathrm{FO}} \mathfrak{A}$ and $\mathfrak{B} \vDash \Phi$ implies that $\mathfrak{B} \vDash \mathrm{Th}_{\mathrm{FO}}(\mathfrak{A}) \cup \Phi$.

$(\Leftarrow)$ Let $\Gamma := \mathrm{Th}_{\mathrm{FO}}(\mathfrak{A}_A)$. If $\mathfrak{B} \vDash \Gamma \cup \Phi$ then $\mathfrak{B}$ is the desired elementary extension of $\mathfrak{A}$. Hence, it is sufficient to show that $\Gamma \cup \Phi$ is satisfiable. For a contradiction, suppose otherwise. Since L is compact there exist finite subsets $\Gamma_0 \subseteq \Gamma$ and $\Phi_0 \subseteq \Phi$ such that $\Gamma_0 \cup \Phi_0$ is inconsistent. Let $\gamma(\bar{a}) := \bigwedge \Gamma_0$ where $\bar{a}$ are the constant symbols appearing in Γ_0. Then $\mathfrak{A} \vDash \exists \bar{x} \gamma(\bar{x})$. Hence, $\Phi_0 \cup \{\exists \bar{x} \gamma(\bar{x})\} \subseteq \mathrm{Th}_{\mathrm{FO}}(\mathfrak{A}) \cup \Phi$. This contradicts the assumption that the latter set is satisfiable. □

Theorem 3.12. *Let L be a regular logical system with $\mathrm{FO}^\circ \leq L$. If L has the Tarski union property and L is compact then $L \equiv \mathrm{FO}^\circ$.*

Proof. Suppose that $\mathrm{FO}^\circ < L$. By Lemma 1.5, there are structures $\mathfrak{A} \equiv \mathfrak{B}$ such that $\mathfrak{A} \vDash \varphi$ and $\mathfrak{B} \vDash \neg\varphi$, for some L-formula φ. We construct an elementary chain $(\mathfrak{A}^n)_{n<\omega}$ such that

- $\mathfrak{A}^n \leq_L \mathfrak{A}^{n+2}$, for all n, and
- $\mathfrak{A}^n \vDash \varphi$ iff n is even.

Then, $\mathfrak{C} := \bigcup_n \mathfrak{A}^n = \bigcup_n \mathfrak{A}^{2n} = \bigcup_n \mathfrak{A}^{2n+1}$. By (TUP) it follows that $\mathfrak{A}^0 \leq_L \mathfrak{C}$ and $\mathfrak{A}^1 \leq_L \mathfrak{C}$. Consequently, we have

$$\mathfrak{A}^0 \vDash \varphi \quad \text{iff} \quad \mathfrak{C} \vDash \varphi \quad \text{iff} \quad \mathfrak{A}^1 \vDash \varphi\,.$$

A contradiction.

It remains to define the chain $(\mathfrak{A}^n)_n$. Let $\mathfrak{A}^0 := \mathfrak{A}$. Since

$$\mathrm{Th}_{\mathrm{FO}}(\mathfrak{A}) \cup \{\neg\varphi\} = \mathrm{Th}_{\mathrm{FO}}(\mathfrak{B}) \cup \{\neg\varphi\}$$

is satisfiable we can use Lemma 3.11 to find an elementary extension $\mathfrak{A}^1 \succeq \mathfrak{A}^0$ with $\mathfrak{A}^1 \vDash \neg\varphi$. Suppose that $\mathfrak{A}^n$ has already been defined. Since

$$\mathrm{Th}_{\mathrm{FO}}(\mathfrak{A}^n_{A_{n-1}}) = \mathrm{Th}_{\mathrm{FO}}(\mathfrak{A}^{n-1}_{A_{n-1}}) \subseteq \mathrm{Th}_L(\mathfrak{A}^{n-1}_{A_{n-1}})$$

it follows that

$$\mathrm{Th}_{\mathrm{FO}}(\mathfrak{A}^n_{A_{n-1}}) \cup \mathrm{Th}_L(\mathfrak{A}^{n-1}_{A_{n-1}})$$

is a satisfiable set of L-formulae. By Lemma 3.11, there exists an elementary extension $\mathfrak{A}^{n+1} \succeq \mathfrak{A}^n$ with $\mathfrak{A}^{n+1} \succeq_L \mathfrak{A}^{n-1}$, as desired. □

Theorem 3.13. *Let L be a regular logical system with* $\mathrm{FO}^0 \leq L$. *If L has the Karp property and L is bounded then* $L \leq \mathrm{FO}_{\infty\aleph_0}$.

Proof. For a contradiction, suppose that there exists an L-formula φ that is not equivalent to any $\mathrm{FO}_{\infty\aleph_0}$-formula.

First, we show that there are structures $\mathfrak{A}_\alpha \equiv_\alpha \mathfrak{B}_\alpha$, for $\alpha \in \mathrm{On}$, such that $\mathfrak{A}_\alpha \vDash \varphi$ and $\mathfrak{B}_\alpha \nvDash \varphi$. Set

$$\psi_\alpha := \bigvee \{ \eta^\alpha_{\mathfrak{A}} \mid \mathfrak{A} \vDash \varphi \},$$

where $\eta^\alpha_{\mathfrak{A}}$ is the Hintikka-formula of $\mathfrak{A}$ of quantifier rank α. Then $\varphi \vDash \psi_\alpha$ and, by assumption, $\psi_\alpha \nvDash \varphi$. Hence, there exist structures $\mathfrak{B}_\alpha \vDash \psi_\alpha$ with $\mathfrak{B}_\alpha \nvDash \varphi$. By definition of ψ_α, it follows that $\mathfrak{B}_\alpha \equiv_\alpha \mathfrak{A}_\alpha$, for some $\mathfrak{A}_\alpha \vDash \varphi$.

As in Lemma 3.9 we can define a formula χ stating that,

- U is a discrete linear order without last element,
- for every $u \in U$, there exists a pseudo back-and-forth system from $A_u := \{ a \mid \langle u, a\rangle \in \bigcup_s P_s \}$ to $B_u := \{ b \mid \langle u, b\rangle \in \bigcup_s Q_s \}$ of length $\downarrow u$,

- A_u induces a substructure that satisfies φ while B_u induces a substructure that does not satisfy φ.

For every ordinal α, we can define a model $\mathfrak{C}_\alpha$ of χ where $\langle U, <\rangle$ is of order type α, $\mathfrak{C}_\alpha|_{A_\beta} \cong \mathfrak{A}_\beta$, and $\mathfrak{C}_\alpha|_{B_\beta} \cong \mathfrak{B}_\beta$. Since L is bounded it follows that χ has a model $\mathfrak{C}$ where $\langle U, <\rangle$ is not well-founded. By Lemma 3.3, it follows that $\mathfrak{C}|_{A_u} \cong_\infty \mathfrak{C}|_{B_u}$, for some $u \in U$. But $\mathfrak{C}|_{A_u} \vDash \varphi$ and $\mathfrak{C}|_{B_u} \nvDash \varphi$ contradicts (KP). □

4. Projective classes

The common idea behind Skolemisation and Chang's Reduction consists in constructing a theory T such that every structure in a given class has an expansion to a model of T. This section contains a more systematic investigation of such reductions.

Definition 4.1. (a) Let $\mathcal{K}$ be a class of Σ-structures and let $\Gamma \subseteq \Sigma$ be a subsignature. The *Γ-projection* of $\mathcal{K}$ is the class

$$\mathrm{pr}_\Gamma(\mathcal{K}) := \{\, \mathfrak{A}|_\Gamma \mid \mathfrak{A} \in \mathcal{K} \,\}$$

of all Γ-reducts of structures in $\mathcal{K}$.

(b) Let L be an algebraic logic and κ either a cardinal or ∞. A class $\mathcal{K}$ of Σ-structures is a *κ-projective L-class* if there exists a signature $\Sigma_+ \supseteq \Sigma$ and a set $\Psi \subseteq L[\Sigma_+]$ of size $|\Psi| \leq \kappa$ such that

$$\mathcal{K} = \mathrm{pr}_\Sigma(\mathrm{Mod}_{L[\Sigma_+]}(\Psi))\,.$$

The class of all such classes $\mathcal{K}$ is denoted by $\mathrm{PC}_\kappa(L, \Sigma)$. Furthermore, we set

$$\mathrm{PC}_{<\kappa}(L, \Sigma) := \bigcup_{\lambda<\kappa} \mathrm{PC}_\lambda(L, \Sigma)\,.$$

Projective FO-classes are also called *pseudo-elementary.*

(c) Let L_0 and L_1 be algebraic logics and κ a cardinal or ∞. We say that L_0 is *κ-projectively reducible* to L_1 and we write $L_0 \leq^{\kappa}_{\mathrm{pc}} L_1$ if

$$\mathrm{Mod}_{L_0[\Sigma]}(\varphi) \in \mathrm{PC}_\kappa(L_1, \Sigma)\,, \quad \text{for all } \Sigma \text{ and every } \varphi \in L_0[\Sigma]\,.$$

Example. The class of all ordered abelian groups is first-order axiomatisable. It follows that the class of all abelian groups that can be ordered is pseudo-elementary.

Exercise 4.1. Prove that $L \leq^{1}_{\mathrm{pc}} \mathrm{SO}_{\kappa\aleph_0}$ implies $L \leq \mathrm{SO}_{\kappa\aleph_0}$.

The results of Section C2.3 can be restated in the following form.

Lemma 4.2. $\mathrm{FO}_{\kappa\aleph_0} \leq^{1}_{\mathrm{pc}} \forall_{\kappa\aleph_0}$.

Proof. For every formula $\varphi \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma, X]$ we can use Lemma C2.3.3 to find a formula $\varphi^* \in \forall_{\kappa\aleph_0}[\Sigma^*, X]$ with $\varphi^* \vDash \varphi$ such that we can expand every model $\mathfrak{A}$ of φ to a model $\mathfrak{A}^*$ of φ^*. Consequently,

$$\mathrm{Mod}(\varphi) = \mathrm{pr}_\Sigma(\mathrm{Mod}(\varphi^*))\,. \qquad \square$$

Lemma 4.3. *If* $L_0 \leq^{\kappa}_{\mathrm{pc}} L_1$ *then*

$$\mathrm{Mod}_{L_0[\Sigma]}(\Phi) \in \mathrm{PC}_{\kappa\oplus|\Phi|}(L_1, \Sigma)\,, \quad \textit{for all } \Phi \subseteq L_0[\Sigma]\,.$$

Proof. For every $\varphi \in \Phi$, there exists a signature $\Sigma(\varphi) \supseteq \Sigma$ and a set $\Psi(\varphi) \subseteq L_1[\Sigma(\varphi)]$ of size at most κ such that

$$\mathrm{Mod}(\varphi) = \mathrm{pr}_\Sigma(\mathrm{Mod}_{L[\Sigma(\varphi)]}(\Psi(\varphi)))\,.$$

We can choose these signatures such that $\Sigma(\varphi) \cap \Sigma(\psi) = \Sigma$, for $\varphi \neq \psi$. Setting $\Psi := \bigcup_{\varphi\in\Phi} \Psi(\varphi)$ it follows that

$$\mathrm{Mod}(\Phi) = \mathrm{pr}_\Sigma(\mathrm{Mod}(\Psi))\,. \qquad \square$$

Lemma 4.4. $L_0 \leq^{\kappa}_{\mathrm{pc}} L_1$ *implies that*

(a) $\mathrm{hn}_\kappa(L_0) \leq \mathrm{hn}_\kappa(L_1)$,

(b) $\mathrm{wn}_\kappa(L_0) \leq \mathrm{wn}_\kappa(L_1)$,

(c) $\ln_\kappa(L_0) \leq \ln_\kappa(L_1)$.

Proof. For (a) and (b), note that if there is a set $\Phi \subseteq L_0[\Sigma]$ of size $|\Phi| \leq \kappa$ that pins down a cardinal λ or an ordinal α then we can find a signature $\Sigma_+ \supseteq \Sigma$ and a set $\Phi_+ \subseteq L_1[\Sigma_+]$ of size $|\Phi_+| \leq |\Phi| \oplus \kappa = \kappa$ that does the same.

(c) Let λ be a cardinal such that every set Φ of L_1-formulae of size $|\Phi| \leq \kappa$ has a model of size at most λ. We claim that $\ln_\kappa(L_0) \leq \lambda$. For each $\Psi \subseteq L_0[\Sigma]$ of size at most κ we can find a set $\Psi_+ \subseteq L_1[\Sigma_+]$ of size $|\Psi_+| \leq |\Psi| \oplus \kappa = \kappa$ such that $\mathrm{Mod}(\Psi) = \mathrm{pr}_\Sigma(\mathrm{Mod}(\Psi_+))$. Consequently, $\mathrm{Mod}(\Phi)$ contains a structure of size at most λ. □

Lemma 4.5. *Let L_0 and L_1 be algebraic logics.*

(a) *If $L_0 \leq_{\mathrm{pc}}^{\infty} L_1$ and L_1 is compact then so is L_0.*

(b) *If $L_0 \leq_{\mathrm{pc}}^{\aleph_0} L_1$ and L_1 is countably compact then so is L_0.*

Proof. Both claims can be proved in the same way. Suppose that every finite subset of $\Phi \subseteq L_0[\Sigma]$ is satisfiable. For every finite $\Phi_0 \subseteq \Phi$, fix a signature $\Sigma(\Phi_0) \supseteq \Sigma$ and a set $\Phi_0^+ \subseteq L_1[\Sigma(\Phi_0)]$ such that

$$\mathrm{Mod}(\Phi_0) = \mathrm{pr}_\Sigma(\mathrm{Mod}(\Phi_0^+)).$$

For (b), we can choose Φ_0^+ to be countable. By replacing

$$\Sigma(\Phi_0) \quad \text{by} \quad \bigcup\{\,\Sigma(\Psi) \mid \Psi \subseteq \Phi_0\,\}$$
$$\text{and} \qquad \Phi_0^+ \quad \text{by} \quad \bigcup\{\,\Psi^+ \mid \Psi \subseteq \Phi_0\,\}$$

we may assume that $\Phi_0 \subseteq \Phi_1$ implies $\Sigma(\Phi_0) \subseteq \Sigma(\Phi_1)$ and $\Phi_0^+ \subseteq \Phi_1^+$.

We claim that the set

$$\Phi^+ := \bigcup\{\,\Phi_0^+ \mid \Phi_0 \subseteq \Phi \text{ finite}\,\}$$

is satisfiable. Note that, in case (b), Φ^+ is a countable union of countable sets. Since L_1 is, respectively, compact and countably compact it is sufficient to prove that every finite subset of Φ^+ is satisfiable.

For every finite subset $\Psi \subseteq \Phi^+$ we can find finitely many finite subsets $\Phi_0, \dots, \Phi_n \subseteq \Phi$ such that $\Psi \subseteq \Phi_0^+ \cup \dots \cup \Phi_n^+$. Setting $\Gamma := \Phi_0 \cup \dots \cup \Phi_n$ it follows that $\Psi \subseteq \Gamma^+$. Hence, $\mathrm{Mod}(\Gamma) \neq \emptyset$ implies that $\mathrm{Mod}(\Psi) \neq \emptyset$, as desired.

Consequently, there exists a model $\mathfrak{A}^+ \vDash \Phi^+$. Let $\mathfrak{A} := \mathfrak{A}^+|_\Sigma$. Then we have $\mathfrak{A} \vDash \Phi_0$, for all finite subsets $\Phi_0 \subseteq \Phi$. This implies that $\mathfrak{A} \vDash \Phi$. □

Lemma 4.6. *Let L_0, L_1 be algebraic logics and $\langle \alpha, \beta \rangle : L_0[\Sigma_0] \to L_1[\Sigma_1]$ a comorphism such that, for every signature $\Gamma_0 \supseteq \Sigma_0$, there exist a signature $\Gamma_1 \supseteq \Sigma_1$ an epimorphism $\langle \alpha_+, \beta_+ \rangle : L_1[\Gamma_1] \to L_0[\Gamma_0]$, and a set $\Psi \subseteq L_1[\Gamma_1]$ such that*

$$\beta_+(\mathfrak{A})|_{\Sigma_1} = \beta(\mathfrak{A}|_{\Sigma_1})\,, \quad \text{for all } \Gamma_0\text{-sructures } \mathfrak{A}\,,$$

and $\operatorname{rng} \beta_+ = \mathrm{Mod}_{L_1[\Gamma_1]}(\Psi)$.

$$\begin{array}{ccc} \mathrm{Str}[\Gamma_0] & \xrightarrow{\ \beta_+\ } & \mathrm{Str}[\Gamma_1] \\ \downarrow{\scriptstyle \mathrm{pr}_{\Sigma_0}} & & \downarrow{\scriptstyle \mathrm{pr}_{\Sigma_1}} \\ \mathrm{Str}[\Sigma_0] & \xrightarrow[\ \beta\ ]{} & \mathrm{Str}[\Sigma_1] \end{array}$$

Then $\mathcal{K} \in \mathrm{PC}_\kappa(L_0, \Sigma_0)$ implies $\beta[\mathcal{K}] \in \mathrm{PC}_\kappa(L_1, \Sigma_1)$.

Proof. Suppose that $\mathcal{K} = \mathrm{pr}_{\Sigma_0}(\mathrm{Mod}(\Phi_0))$, for some $\Phi_0 \subseteq L_0[\Gamma_0]$. Let $\langle \alpha_+, \beta_+ \rangle : L_1[\Gamma_1] \to L_0[\Gamma_0]$ be the corresponding epimorphism of the expansion and $\langle \gamma, \delta \rangle : L_0[\Gamma_0] \to L_1[\Gamma_1]$ its right inverse. We set

$$\Phi_1 := \gamma[\Phi_0] \cup \Psi\,.$$

Then we have

$$\begin{aligned} \mathfrak{B} \vDash \Phi_1 \quad & \text{iff} \quad \mathfrak{B} \vDash \gamma[\Phi_0] \text{ and } \mathfrak{B} = \beta_+(\mathfrak{A}) \text{ for some } \mathfrak{A} \\ & \text{iff} \quad \mathfrak{B} = \beta_+(\mathfrak{A}) \text{ for some } \mathfrak{A} \text{ with } \beta_+(\mathfrak{A}) \vDash \gamma[\Phi_0] \\ & \text{iff} \quad \mathfrak{B} = \beta_+(\mathfrak{A}) \text{ for some } \mathfrak{A} \text{ with } \mathfrak{A} \vDash (\alpha \circ \gamma)[\Phi_0] \\ & \text{iff} \quad \mathfrak{B} = \beta_+(\mathfrak{A}) \text{ for some } \mathfrak{A} \text{ with } \mathfrak{A} \vDash \Phi_0\,. \end{aligned}$$

Hence, $\mathrm{Mod}_{L_1}(\Phi_1) = \beta_+[\mathrm{Mod}_{L_0}(\Phi_0)]$ and it follows that

$$\begin{aligned} \mathfrak{A} \in \beta[\mathcal{K}] \quad &\text{iff} \quad \mathfrak{A} = \beta(\mathfrak{A}'|_{\Sigma_0}) \quad &&\text{for some } \mathfrak{A}' \vDash \Phi_0 \\ &\text{iff} \quad \mathfrak{A} = \beta_+(\mathfrak{A}')|_{\Sigma_1} \quad &&\text{for some } \mathfrak{A}' \vDash \Phi_0 \\ &\text{iff} \quad \mathfrak{A} = \mathfrak{A}'|_{\Sigma_1} \quad &&\text{for some } \mathfrak{A}' \vDash \Phi_1 . \end{aligned}$$

Consequently, we have $\beta[\mathcal{K}] = \mathrm{pr}_{\Sigma_1}(\mathrm{Mod}(\Phi_1))$. □

Corollary 4.7. *Suppose that $\Sigma_0 \subseteq \Sigma_1$ are signatures and $(\varphi_\xi)_{\xi \in \Sigma_1 \setminus \Sigma_0}$ is a sequence of $\mathrm{FO}_{\kappa\aleph_0}[\Sigma_0]$-formulae. Let $\langle \alpha, \beta \rangle : \mathrm{FO}_{\kappa\aleph_0}[\Sigma_0] \to \mathrm{FO}_{\kappa\aleph_0}[\Sigma_1]$ be the comorphism where β maps a structure $\mathfrak{A}$ to its expansion defined by $(\varphi_\xi)_\xi$. If $\mathcal{K} \in \mathrm{PC}_\kappa(\mathrm{FO}_{\kappa\aleph_0}, \Sigma_0)$ then $\beta[\mathcal{K}] \in \mathrm{PC}_\kappa(\mathrm{FO}_{\kappa\aleph_0}, \Sigma_1)$.*

Proof. We have to show that $\langle \alpha, \beta \rangle$ satisfies the condition of the preceding lemma. Given Γ_0 set $\Gamma_1 := \Sigma_1 \cup \Gamma_0$. We define $\langle \alpha_+, \beta_+ \rangle$ as follows. The function β_+ maps a Γ_0-structure $\mathfrak{A}$ to the Γ_1-structure $\mathfrak{B}$ such that $\mathfrak{B}|_{\Gamma_0} = \mathfrak{A}$ and $\mathfrak{B}|_{\Sigma_1} = \beta(\mathfrak{A}|_{\Sigma_0})$. Then $\langle \alpha_+, \beta_+ \rangle$ is an epimorphism whose right inverse is given by the reduct operation. By definition, it satisfies $\beta_+(\mathfrak{A})|_{\Sigma_1} = \beta(\mathfrak{A}|_{\Sigma_0})$. Furthermore, we can define the range of β_+ by formulae of the form

$$\forall \bar{x}[R\bar{x} \leftrightarrow \varphi_R(\bar{x})] \quad \text{and} \quad \forall \bar{x} \forall y[f\bar{x} = y \leftrightarrow \varphi_f(\bar{x}, y)] .$$ □

We can generalise the notion of a projective class by replacing the reduct operation by a combination of a reduct and a domain restriction.

Definition 4.8. Let Σ be an S-sorted signature.

(a) Let $\mathfrak{A}$ be a Σ-structure. A *relativised reduct* of $\mathfrak{A}$ is a structure of the form $\mathfrak{A}|_{\Sigma_0}|_P$ where $\Sigma_0 \subseteq \Sigma$ and $P \subseteq A$ induces a substructure of $\mathfrak{A}|_{\Sigma_0}$.

(b) Let L be an algebraic logic and κ either a cardinal or ∞. A class $\mathcal{K}$ of Σ-structures is a *relativised κ-projective L-class* if there exists a signature $\Sigma_+ \supseteq \Sigma$, a set $\Psi \subseteq L[\Sigma_+]$ of size $|\Psi| \leq \kappa$, and unary predicates $P_s \in \Sigma_+$, for $s \in S$, such that

$$\mathcal{K} = \{ \mathfrak{A}|_\Sigma|_{\bigcup_s P_s^{\mathfrak{A}}} \mid \mathfrak{A} \in \mathrm{Mod}_{L[\Sigma_+]}(\Psi) \} .$$

The class of all such classes is denoted by $\mathrm{RPC}_\kappa(L, \Sigma)$.

(c) Let L_0 and L_1 be algebraic logics and κ a cardinal or ∞. We say that L_0 is *relativised κ-projectively reducible* to L_1 and we write $L_0 \leq^\kappa_{\mathrm{rpc}} L_1$, if

$$\mathrm{Mod}_{L_0[\Sigma]}(\varphi) \in \mathrm{RPC}_\kappa(L_1, \Sigma)\,, \quad \text{for all } \Sigma \text{ and every } \varphi \in L_0[\Sigma]\,.$$

Lemma 4.9. *$L_0 \leq^\kappa_{\mathrm{rpc}} L_1$ implies that $\ln_\kappa(L_0) \leq \ln_\kappa(L_1)$.*

Proof. Let λ be a cardinal such that each satisfiable set Φ of L_1-formulae of size $|\Phi| \leq \kappa$ has a model of size at most λ. We claim that $\ln_\kappa(L_0) \leq \lambda$. For each $\Phi \subseteq L_0[\Sigma]$ of size at most κ we can find a set $\Phi_+ \subseteq L_1[\Sigma_+]$ of size $|\Phi_+| \leq |\Phi| \oplus \kappa = \kappa$ such that

$$\mathrm{Mod}(\Phi) = \big\{\, \mathfrak{A}|_\Sigma|_{\bigcup_s P^{\mathfrak{A}}_s} \bigm| \mathfrak{A} \in \mathrm{Mod}_{L[\Sigma_+]}(\Phi_+) \,\big\}\,.$$

Consequently, if Φ is satisfiable then $\mathrm{Mod}(\Phi)$ contains a structure of size at most λ. □

Example. Let us show that $\mathrm{SO} \leq^1_{\mathrm{rpc}} \mathrm{MSO}$. Suppose that $\varphi \in \mathrm{SO}[\Sigma, X]$ where Σ is S-sorted for a finite set S. W.l.o.g. we may assume that φ contains no quantifiers over functions. Fix a number $n < \omega$ such that every second-order quantifier in φ ranges over a relation of arity at most n. For every sequence $\bar{s} \in S^{\leq n}$ of sorts of length at least 2, we add to Σ a new sort $p_{\bar{s}}$ and a function $g_{\bar{s}}$ of type $\bar{s} \to p_{\bar{s}}$. Let $\chi_{\bar{s}}$ be the formula stating that $g_{\bar{s}} : A_{s_0} \times \dots \times A_{s_{k-1}} \to A_{p_{\bar{s}}}$ is bijective. We construct a formula φ' by replacing in φ

- every second-order quantifier over a relation R of type $\bar{s}$ by a quantifier over a set X_R of sort $p_{\bar{s}}$,
- every atom $R\bar{t}$ where R is such a relation by the formula $X_R g_{\bar{s}} \bar{t}$.

Setting $\psi := \varphi' \wedge \bigwedge_{\bar{s} \in S^{\leq n}} \chi_{\bar{s}}$ it follows that

$$\mathrm{Mod}(\varphi) = \{\, \mathfrak{A}|_\Sigma|_S \mid \mathfrak{A} \in \mathrm{Mod}(\psi) \,\}\,.$$

Exercise 4.2. State and prove a version of Lemma 4.6 for relativised projective classes and use it to show that the image of a relativised projective class $\mathcal{K}$ under an interpretation is again a relativised projective class.

Below we will show that for first-order logic there is no difference between projective and relativised projective classes. To do so we need some technical results about recovering a structure from a substructure.

Definition 4.10. Let $\mathfrak{A}$ be a Σ-structure and $C \subseteq A$.

(a) Let $\Gamma_0(\Sigma)$ be the signature consisting of n-ary relation symbols R_φ, for every atomic formula $\varphi \in \mathrm{FO}^{<\omega}[\Sigma]$ with $\mathrm{free}(\varphi) = \{x_0, \dots, x_{n-1}\}$. We assume that $\Gamma_0(\Sigma) \cap \Sigma = \emptyset$ and we set $\Gamma(\Sigma) := \Sigma \cup \Gamma_0(\Sigma)$.

(b) By $\langle\!\langle C \rangle\!\rangle^+_{\mathfrak{A}}$ we denote the $\Gamma(\Sigma)$-expansion of $\langle\!\langle C \rangle\!\rangle_{\mathfrak{A}}$ by the relations

$$R_\varphi := \{\, \bar{a} \in C^n \mid \mathfrak{A} \vDash \varphi(\bar{a}) \,\},$$

and we define

$$\langle\!\langle C \rangle\!\rangle^\circ_{\mathfrak{A}} := \langle\!\langle C \rangle\!\rangle^+_{\mathfrak{A}}|_{\Gamma_0(\Sigma)}|_C .$$

(c) Let $\Xi(\Sigma)$ be the first-order theory of the class

$$\mathcal{K}(\Sigma) := \big\{\, \langle\!\langle C \rangle\!\rangle^\circ_{\mathfrak{A}} \;\big|\; \mathfrak{A} \text{ a } \Sigma\text{-structure with } C \subseteq A \,\big\}.$$

Remark. Note that

$$\langle\!\langle C \rangle\!\rangle^\circ_{\mathfrak{A}} \cong \langle\!\langle D \rangle\!\rangle^\circ_{\mathfrak{B}} \quad \text{implies} \quad \langle\!\langle C \rangle\!\rangle^+_{\mathfrak{A}} \cong \langle\!\langle D \rangle\!\rangle^+_{\mathfrak{B}} .$$

Lemma 4.11. *If $\mathfrak{C} \vDash \Xi(\Sigma)^{\vDash}_{\forall}$ then there exists a Σ-structure $\mathfrak{A}$ with $A \supseteq C$ such that C generates $\mathfrak{A}$ and $\mathfrak{C} = \langle\!\langle C \rangle\!\rangle^\circ_{\mathfrak{A}}$.*

Proof. We define an equivalence relation $\sim$ on the set

$$Z := \big\{\, t(\bar{c}) \;\big|\; t \text{ a } \Sigma\text{-term and } \bar{c} \subseteq C \,\big\}$$

by $\quad s(\bar{a}) \sim t(\bar{b}) \quad :\text{iff} \quad \bar{a}\bar{b} \in R^{\mathfrak{C}}_\varphi$ where $\varphi := s(\bar{x}) = t(\bar{y})$.

Note that $C \subseteq Z$ since we can choose $t = x$. Set $A := Z/{\sim}$. If $a, b \in C$ are elements with $a \neq b$ then $\langle a, b\rangle \notin R^{\mathfrak{C}}_{x=y}$ since

$$\forall x \forall y (R_{x=y}xy \rightarrow x = y) \in \Xi(\Sigma)_\forall .$$

This implies that $[a]_\sim \neq [b]_\sim$. Hence, the function $e : C \to A : a \mapsto [a]_\sim$ is an embedding. Let D be the range of this function. We construct a $\Gamma(\Sigma)$-structure $\mathfrak{A}$ with universe A such that $\langle\!\langle D \rangle\!\rangle^{\circ}_{\mathfrak{A}} \cong \mathfrak{C}$.

For $R_\varphi \in \Gamma(\Sigma)_0$, we define

$$R^{\mathfrak{A}}_\varphi := \{ e(\bar{a}) \mid \bar{a} \in R^{\mathfrak{C}}_\varphi \} .$$

For atomic formulae $\psi \in \mathrm{FO}^{<\omega}[\Sigma]$, we define

$$\mathfrak{A} \vDash \psi\big([t_0(\bar{a}_0)]_\sim, \dots, [t_{n-1}(\bar{a}_{n-1})]_\sim\big) \quad :\text{iff} \quad \bar{a}_0 \dots \bar{a}_{n-1} \in R^{\mathfrak{C}}_\varphi ,$$

where $\varphi(\bar{x}_0, \dots, \bar{x}_{n-1}) := \psi(t_0(\bar{x}_0), \dots, t_{n-1}(\bar{x}_{n-1}))$.

It remains to show that D generates $\mathfrak{A}$ and that $\langle\!\langle D \rangle\!\rangle^{\circ}_{\mathfrak{A}} \cong \mathfrak{C}$. Let $t(\bar{x})$ be a Σ-term and $\bar{a} \in C^n$. Then

$$t^{\mathfrak{A}}(e(\bar{a})) = [t(\bar{a})]_\sim$$

since setting $\psi(\bar{x}, y) := t(\bar{x}) = y$ and $\varphi := t(\bar{x}) = t(\bar{y})$ we have

$$\begin{aligned} & \mathfrak{A} \vDash t(e(\bar{a})) = [t(\bar{a})]_\sim \\ \text{iff}\quad & \mathfrak{A} \vDash \psi\big([a_0]_\sim, \dots, [a_{n-1}]_\sim, [t(\bar{a})]_\sim\big) \\ \text{iff}\quad & \bar{a}\bar{a} \in R^{\mathfrak{C}}_\varphi , \end{aligned}$$

and $\forall \bar{x} R_\varphi \bar{x}\bar{x} \in \Xi(\Sigma)_\forall$. In particular, D generates $\mathfrak{A}$.

If $\varphi(\bar{x}) \in \mathrm{FO}[\Sigma]$ is an atomic formula and $\bar{a} \in C^n$ then

$$\langle\!\langle D \rangle\!\rangle^{\circ}_{\mathfrak{A}} \vDash R_\varphi e(\bar{a}) \quad \text{iff} \quad \mathfrak{A} \vDash R_\varphi e(\bar{a}) \quad \text{iff} \quad \mathfrak{C} \vDash R_\varphi \bar{a}$$

implies that $e : \mathfrak{C} \cong \langle\!\langle D \rangle\!\rangle^{\circ}_{\mathfrak{A}}$. By taking isomorphic copies we may assume that $D = C \subseteq A$. □

Definition 4.12. For every model $\mathfrak{C} \vDash \Xi(\Sigma)^{\vDash}_{\forall}$, we denote by $\widehat{\mathfrak{C}}$ some structure as in the preceding lemma. Note that, up to isomorphism, $\widehat{\mathfrak{C}}$ is unique.

Lemma 4.13. *For every theory $T \subseteq \forall[\Sigma]$, there exists a theory $\widehat{T} \subseteq$ $\mathrm{FO}[\Gamma(\Sigma)_0]$ such that*

$$\widehat{\mathfrak{A}} \vDash T \quad \text{iff} \quad \mathfrak{A} \vDash \widehat{T}\,.$$

Proof. For every universal sentence φ we will construct a set $\Phi(\varphi)$ of $\mathrm{FO}^{<\omega}[\Gamma(\Sigma)_0]$-sentences such that

$$\widehat{\mathfrak{A}} \vDash \varphi \quad \text{iff} \quad \mathfrak{A} \vDash \Phi(\varphi)\,.$$

Then we can set $\widehat{T} := \bigcup \{ \Phi(\varphi) \mid \varphi \in T \}$.

W.l.o.g. assume that $\varphi = \forall \bar{x} \bigwedge_i \bigvee_k \vartheta_{ik}$ where the quantifier-free part is in conjunctive normal form. For an atomic formula $\vartheta(x_0, \dots, x_{n-1})$, we have

$$\begin{aligned} & \widehat{\mathfrak{A}} \vDash \vartheta\big([t_0(\bar{a}_0)]_\sim, \dots, [t_{n-1}(\bar{a}_{n-1})]_\sim\big) \\ \text{iff} \quad & \mathfrak{A} \vDash R_\psi(\bar{a}_0, \dots, \bar{a}_{n-1}) \end{aligned}$$

where $\psi := \vartheta(t_0(\bar{x}_0), \dots, t_{n-1}(\bar{x}_{n-1}))$. Consequently, if, for each tuple $\bar{t}$ of Σ-terms, we define

$$\hat{\vartheta}_{ik}[\bar{t}] := \begin{cases} R_{\psi_{ik}} \bar{x}_0 \dots \bar{x}_{n-1} & \text{if } \vartheta_{ik} \text{ is an atom}\,, \\ \neg R_{\psi_{ik}} \bar{x}_0 \dots \bar{x}_{n-1} & \text{if } \vartheta_{ik} \text{ is a negated atom}\,, \end{cases}$$

where $\psi_{ik} := \vartheta_{ik}(t_0(\bar{x}_0), \dots, t_{n-1}(\bar{x}_{n-1}))$, then it follows that

$$\begin{aligned} & \widehat{\mathfrak{A}} \vDash \bigwedge_i \bigvee_k \vartheta_{ik}\big([t_0(\bar{a}_0)]_\sim, \dots, [t_{n-1}(\bar{a}_{n-1})]_\sim\big) \\ \text{iff} \quad & \mathfrak{A} \vDash \bigwedge_i \bigvee_k \hat{\vartheta}_{ik}[\bar{t}](\bar{a}_0, \dots, \bar{a}_{n-1})\,. \end{aligned}$$

Hence, we can set

$$\Phi(\varphi) := \big\{ \forall \bar{x} \textstyle\bigwedge_i \bigvee_k \hat{\vartheta}[\bar{t}] \;\big|\; \bar{t} \text{ a tuple of } \Sigma\text{-terms} \big\}\,.$$

(Note that every element of $\widehat{\mathfrak{A}}$ is denoted by a term with parameters from A.) □

Exercise 4.3. Let κ be an infinite cardinal with $\kappa > |\Sigma|$. Prove that, for every $\forall_{\kappa\aleph_0}[\Sigma]$-theory T, there exists an $\mathrm{FO}_{\kappa\aleph_0}[\Gamma(\Sigma)_0]$-theory $\widehat{T}$ as above.

Theorem 4.14. *Let $\Sigma_- \subseteq \Sigma_+$ be signatures and $P \in \Sigma_+$ a unary predicate. Suppose that*

$$\mathcal{K} = \big\{ \mathfrak{A}|_{\Sigma_-}|_{\bigcup_s P_s^{\mathfrak{A}}} \bigm| \mathfrak{A} \in \mathrm{Mod}(\Phi) \big\}, \quad \textit{for some } \Phi \subseteq \mathrm{FO}^0[\Sigma_+]\,.$$

(a) *There exists a signature $\Gamma \supseteq \Sigma_-$ of size $|\Gamma| \leq |\Sigma_+| \oplus \aleph_0$ and a theory $\Psi \subseteq \mathrm{FO}^0[\Gamma]$ such that*

$$\mathcal{K} = \mathrm{pr}_{\Sigma_-}(\mathrm{Mod}(\Psi))\,.$$

(b) *If Φ is finite and every structure in $\mathcal{K}$ is infinite then we can choose a finite set Ψ as above.*

(c) *$\mathcal{K}$ is a pseudo-elementary class.*

Proof. W.l.o.g. we may assume that $\Sigma_- = \Sigma_+$. Hence, we drop the subscripts and just write Σ.

(b) Since Φ is finite we may assume that the signature Σ is finite. By the Theorem of Löwenheim and Skolem, it follows that, for every structure $\mathfrak{A} \in \mathcal{K}$, we can find a structure $\mathfrak{B} \in \mathrm{Mod}(\Phi)$ of cardinality $|B| = |A|$ such that $\mathfrak{A} = \mathfrak{B}|_{\bigcup_s P_s^{\mathfrak{B}}}$. Let $\Sigma' = \{ \xi' \mid \xi \in \Sigma \}$ be a disjoint copy of Σ, and set $\Gamma := \Sigma \cup \Sigma' \cup \{f\}$, where f is a new unary function symbol. Since Φ is finite there exists a sentence $\psi \in \mathrm{FO}[\Gamma]$ expressing that

- the Σ'-reduct of the given structure is a model of Φ,
- f is a bijection between the whole universe and P.

It follows that $\mathcal{K} = \{ \mathfrak{A}|_\Sigma \mid \mathfrak{A} \vDash \psi \}$.

(c) follows immediately from (a).

(a) By Skolemising we may assume that $\Phi \subseteq \forall$. Let $\Psi \subseteq \mathrm{FO}^0[\Gamma(\Sigma)]$ consist of $\widehat{\Phi} \cup \Xi(\Sigma)^{\vDash}_{\forall}$ together with with the sentences

$$\begin{aligned} &\forall \bar{x}[\varphi(\bar{x}) \leftrightarrow R_\varphi(\bar{x})] && \text{for atomic } \varphi \in \mathrm{FO}^{<\omega}[\Sigma]\,,\\ &\forall \bar{x}[R_{Pt\bar{x}}\bar{x} \to \exists y R_{t\bar{x}=y}\bar{x}y]\,, && \text{for every } \Sigma\text{-term } t\,. \end{aligned}$$

We claim that

$$\mathcal{K} = \left\{ \mathfrak{A}|_\Sigma \;\middle|\; \mathfrak{A} \in \mathrm{Mod}(\Psi) \right\}.$$

($\subseteq$) If $\mathfrak{C} \in \mathcal{K}$ then $\mathfrak{C} = \mathfrak{A}|_{\bigcup_s P_s^{\mathfrak{A}}}$, for some $\mathfrak{A} \vDash \Phi$. Since Φ is a Skolem theory we can assume that $P^{\mathfrak{A}}$ generates $\mathfrak{A}$. Hence, $\langle\!\langle X \rangle\!\rangle^+_{\mathfrak{A}}$ is defined and $\mathfrak{C} \cong \langle X \rangle^\circ_{\mathfrak{A}}|_\Sigma$. Furthermore, $\langle\!\langle X \rangle\!\rangle^\circ_{\mathfrak{A}} \vDash \Psi$, as desired.

($\supseteq$) Let $\mathfrak{A} \vDash \Psi$. Since $\mathfrak{A} \vDash \Phi \cup \Xi(\Sigma)^{\vDash}_{\forall}$ it follows that $\widehat{\mathfrak{A}}$ exists and $\langle\!\langle X \rangle\!\rangle^\circ_{\widehat{\mathfrak{A}}} = \mathfrak{A}$. Since $\mathfrak{A} \vDash \widehat{\Phi}$ we have $\widehat{\mathfrak{A}} \vDash \Phi$. Consequently, $\widehat{\mathfrak{A}}|_{P^{\widehat{\mathfrak{A}}}} \in \mathcal{K}$. We claim that $\widehat{\mathfrak{A}}|_{P^{\widehat{\mathfrak{A}}}} = \mathfrak{A}$. On the one hand, $\Xi(\Sigma)^{\vDash}_{\forall} \vDash \forall x R_{Px} x$ implies that $\mathfrak{A} \vDash Pa$, for every $a \in A$. Hence, $A \subseteq P^{\widehat{\mathfrak{A}}}$. Conversely, suppose that $a \in P^{\widehat{\mathfrak{A}}}$. Then $a = t(\bar{b})$, for some term t and parameters $\bar{b} \subseteq A$. Then

$$\mathfrak{A} \vDash R_{Pt\bar{x}}\bar{b} \wedge \exists y R_{t\bar{x}=y}\bar{b}y$$

which implies that $a \in A$. □

Corollary 4.15. $L \leq^\infty_{\mathrm{rpc}} \mathrm{FO}$ *iff* $L \leq^\infty_{\mathrm{pc}} \mathrm{FO}$.

5. *Interpolation*

For most logics L there are projective L-classes that are not L-axiomatisable. In this section we study how this additional power affects the entailment relation. Surprisingly we can find many logics where it has no effect at all.

Definition 5.1. Let L be an algebraic logic.

(a) L has the *interpolation property* if, for all finite sets $\Phi_i \subseteq L[\Sigma_i]$, $i < 2$, with $\Phi_0 \vDash \Phi_1$, there exists a finite set $\Psi \subseteq L[\Sigma_0 \cap \Sigma_1]$ such that

$$\Phi_0 \vDash \Psi \quad \text{and} \quad \Psi \vDash \Phi_1 .$$

(b) L has the *Δ-interpolation property* if every class $\mathcal{K} \in \mathrm{PC}_{<\aleph_0}(L, \Sigma)$ with $\mathrm{Str}[\Sigma] \setminus \mathcal{K} \in \mathrm{PC}_{<\aleph_0}(L, \Sigma)$ is finitely L-axiomatisable.

Remark. If L is boolean closed then the interpolation property implies the Δ-interpolation property since, if

$$\mathcal{K} = \mathrm{pr}_\Sigma(\mathrm{Mod}(\Phi_+)) \quad \text{and} \quad \mathrm{Str}[\Sigma] \setminus \mathcal{K} = \mathrm{pr}_\Sigma(\mathrm{Mod}(\Phi_-))$$

then we have

$$\Phi_+ \vDash \neg \bigwedge \Phi_-$$

and any set $\Psi \subseteq L[\Sigma]$ with

$$\Phi_+ \vDash \Psi \quad \text{and} \quad \Psi \vDash \neg \bigwedge \Phi_-$$

is an axiom system for $\mathcal{K}$.

Theorem 5.2. FO *has the interpolation property.*

Proof. Since FO is closed under conjunctions it is sufficient to consider single formulae. Hence, suppose that $\varphi_0 \vDash \varphi_1$ where $\varphi_i \in \mathrm{FO}^\circ[\Gamma_i]$, for $i < 2$. Let

$$\Psi := (\varphi_0)^\vDash \cap \mathrm{FO}^\circ[\Sigma], \quad \text{where} \quad \Sigma := \Gamma_0 \cap \Gamma_1 .$$

We claim that $\Psi \cup \{\neg\varphi_1\}$ is inconsistent. By compactness, it then follows that there is a finite subset $\Psi_0 \subseteq \Psi$ such that $\Psi_0 \cup \{\neg\varphi_1\}$ is inconsistent. Setting $\psi := \bigwedge \Psi_0$ we have $\varphi_0 \vDash \psi$ and $\psi \vDash \varphi_1$, as desired.

It remains to prove the claim. For a contradiction, suppose that the set $\Psi \cup \{\neg\varphi_1\}$ has a model $\mathfrak{A}$. By Corollary C2.5.9, there exists a model $\mathfrak{B}$ of φ_0 such that $\mathfrak{A}|_\Sigma \preceq \mathfrak{B}|_\Sigma$. Since $\mathfrak{A}|_\Sigma \equiv \mathfrak{B}|_\Sigma$ we can apply Theorem C2.5.8 to obtain a $(\Gamma_0 \cup \Gamma_1)$-structure $\mathfrak{C}$ with $\mathfrak{B} \preceq \mathfrak{C}|_{\Gamma_0}$ and an elementary embedding $g : \mathfrak{A} \to \mathfrak{C}|_{\Gamma_1}$. In particular, we have $\mathfrak{C}|_{\Gamma_0} \vDash \varphi_0$ and $\mathfrak{C}|_{\Gamma_1} \vDash \neg\varphi_1$. Hence, $\mathfrak{C} \vDash \varphi_0 \wedge \neg\varphi_1$ and $\varphi_0 \nvDash \varphi_1$. Contradiction. □

Definition 5.3. Let L be an algebraic logic, Σ a signature, $R \notin \Sigma$ an n-ary relation symbol, and $\Phi(R) \subseteq L^\circ[\Sigma \cup \{R\}]$ a set of formulae.

(a) We say that R is *implicitly defined* by Φ if, for all models $\mathfrak{A}$ and $\mathfrak{B}$ of Φ with $\mathfrak{A}|_\Sigma = \mathfrak{B}|_\Sigma$, we have $R^{\mathfrak{A}} = R^{\mathfrak{B}}$.

(b) We say that R is *explicitly defined* by a set $\Psi \subseteq L^n[\Sigma]$ with respect to Φ if $R^{\mathfrak{A}} = \Psi^{\mathfrak{A}}$, for every model $\mathfrak{A}$ of Φ.

(c) L has the *Beth property* if, for all finite sets $\Phi \subseteq L^0[\Sigma \cup \{R\}]$ that define R implicitly, there exists a finite set $\Psi \subseteq L^n[\Sigma]$ explicitly defining R with respect to Φ.

Lemma 5.4. *Every boolean closed logic L with the interpolation property has the Beth property.*

Proof. Suppose that R is implicitly defined by $\Phi(R) \subseteq L^0[\Sigma \cup \{R\}]$. Let R' be a new relation symbol. It follows that

$$\bigwedge \Phi(R) \to R\bar{x} \vDash \bigwedge \Phi(R') \to R'\bar{x}\,.$$

By the interpolation property we can find a finite set $\Psi(\bar{x})$ such that

$$\bigwedge \Phi(R) \to R\bar{x} \vDash \bigwedge \Psi(\bar{x}) \quad \text{and} \quad \bigwedge \Psi(\bar{x}) \vDash \bigwedge \Phi(R') \to R'\bar{x}\,.$$

It follows that

$$\Phi(R) \vDash R\bar{x} \leftrightarrow \bigwedge \Psi(\bar{x})\,,$$

that is, Ψ explicitly defines R with respect to Φ. □

There is a general way to extend a given logic to one that has the Δ-interpolation property.

Definition 5.5. Let L be an algebraic logic. The *interpolation closure* $\Delta(L)$ of L is the logic where $\Delta(L)[\Sigma]$ consists of all pairs

$$\langle \varphi_0, \varphi_1 \rangle \in L[\Sigma_0] \times L[\Sigma_1]$$

with $\Sigma_i \supseteq \Sigma$ and

$$\mathrm{pr}_\Sigma(\mathrm{Mod}(\varphi_1)) = \mathrm{Str}[\Sigma] \setminus \mathrm{pr}_\Sigma(\mathrm{Mod}(\varphi_0))\,.$$

The semantics of such a formula is defined by

$$\mathfrak{A} \vDash \langle \varphi_0, \varphi_1 \rangle \quad :\text{iff} \quad \mathfrak{A} \in \mathrm{pr}_\Sigma(\mathrm{Mod}(\varphi_0))\,.$$

Lemma 5.6. *Let L be an algebraic logic.*

(a) *If L satisfies* (B_+) *then* $\Delta(L)$ *is boolean closed.*

(b) *If* L_0 *is closed under negation and* $L_0 \leq^1_{\text{pc}} L_1$*, then* $L_0 \leq \Delta(L_1)$.

(c) *If L is closed under negation then* $L \leq \Delta(L)$.

(d) $L_0 \leq^1_{\text{pr}} L_1$ *implies* $\Delta(L_0) \leq \Delta(L_1)$.

(e) $\Delta(\Delta(L)) \leq \Delta(L)$.

(f) $\Delta(L)$ *has the* Δ*-interpolation property.*

(g) *If* L_1 *has the* Δ*-interpolation property then*

$$L_0 \leq^1_{\text{pr}} L_1 \quad \textit{implies} \quad \Delta(L_0) \leq L_1 \,.$$

(h)
$$\begin{aligned} \text{occ}(\Delta(L)) &= \text{occ}(L)\,, \\ \text{ln}_\kappa(\Delta(L)) &= \text{ln}_\kappa(L)\,, \\ \text{wn}_\kappa(\Delta(L)) &= \text{wn}_\kappa(L)\,. \end{aligned}$$

Proof. (a) We have

$$\begin{aligned} \langle\varphi,\psi\rangle \wedge \langle\varphi',\psi'\rangle &\equiv \langle\varphi \wedge \varphi',\ \psi \vee \psi'\rangle\,, \\ \langle\varphi,\psi\rangle \vee \langle\varphi',\psi'\rangle &\equiv \langle\varphi \vee \varphi',\ \psi \wedge \psi'\rangle\,, \end{aligned}$$

and
$$\neg\langle\varphi,\psi\rangle \equiv \langle\psi,\varphi\rangle\,.$$

(b) For every $\varphi \in L_0[\Sigma]$, there exist a signature $\Sigma_0 \supseteq \Sigma$ and a formula $\psi_0 \in L_1[\Sigma_0]$ such that

$$\text{Mod}(\varphi) = \text{pr}_\Sigma(\text{Mod}(\psi_0))\,.$$

Similarly, there exist a signature $\Sigma_1 \supseteq \Sigma$ and a formula $\psi_1 \in L_1[\Sigma_1]$ such that

$$\text{Mod}(\neg\varphi) = \text{pr}_\Sigma(\text{Mod}(\psi_1))\,.$$

It follows that $\varphi \equiv \langle\psi_0,\psi_1\rangle \in \Delta(L_1)$.

(c) follows immediately from (b).

(d) Let $\langle \varphi_0, \psi_0 \rangle \in \Delta(L_0)$ where $\varphi_0 \in L_0[\Sigma_0]$ and $\psi_0 \in L_0[\Gamma_0]$. Since $L_0 \leq^1_{\mathrm{pr}} L_1$ we can find formulae $\varphi_1 \in L_1[\Sigma_1]$ and $\psi_1 \in L_1[\Gamma_1]$ such that

$$\begin{aligned} &\mathrm{Mod}(\varphi_0) = \mathrm{pr}_{\Sigma_0}(\mathrm{Mod}(\varphi_1)) \\ \text{and}\quad &\mathrm{Mod}(\psi_0) = \mathrm{pr}_{\Gamma_0}(\mathrm{Mod}(\psi_1)) \,. \end{aligned}$$

Hence, $\langle \varphi_0, \psi_0 \rangle \equiv \langle \varphi_1, \psi_1 \rangle \in \Delta(L_1)$.

(e) Let $\langle \langle \varphi_0, \psi_0 \rangle, \langle \varphi_1, \psi_1 \rangle \rangle \in \Delta(\Delta(L))[\Sigma]$. Then

$$\begin{aligned} \mathrm{pr}_\Sigma(\mathrm{Mod}(\varphi_0)) &= \mathrm{pr}_\Sigma(\mathrm{pr}_{\Sigma_0}(\mathrm{Mod}(\varphi_0))) \\ &= \mathrm{pr}_\Sigma(\mathrm{Mod}(\langle \varphi_0, \psi_0 \rangle)) \\ &= \mathrm{Str}[\Sigma] \setminus \mathrm{pr}_\Sigma(\mathrm{Mod}(\langle \varphi_1, \psi_1 \rangle)) \\ &= \mathrm{Str}[\Sigma] \setminus \mathrm{pr}_\Sigma(\mathrm{pr}_{\Sigma_1}(\mathrm{Mod}(\varphi_1))) \\ &= \mathrm{Str}[\Sigma] \setminus \mathrm{pr}_\Sigma(\mathrm{Mod}(\varphi_1)) \,. \end{aligned}$$

Consequently, $\langle \langle \varphi_0, \psi_0 \rangle, \langle \varphi_1, \psi_1 \rangle \rangle \equiv \langle \varphi_0, \varphi_1 \rangle \in \Delta(L)[\Sigma]$.

(f) Let $\langle \varphi_0, \psi_0 \rangle, \langle \varphi_1, \psi_1 \rangle \in \Delta(L)$ be formulae such that

$$\mathrm{pr}_\Sigma(\mathrm{Mod}(\langle \varphi_0, \psi_0 \rangle)) = \mathrm{Str}[\Sigma] \setminus \mathrm{pr}_\Sigma(\mathrm{Mod}(\langle \varphi_1, \psi_1 \rangle)) \,.$$

Then $\langle \langle \varphi_0, \psi_0 \rangle, \langle \varphi_1, \psi_1 \rangle \rangle \in \Delta(\Delta(L))$ and, by (e), there is a formula $\langle \vartheta, \chi \rangle \in \Delta(L)$ such that

$$\langle \vartheta, \chi \rangle \equiv \langle \langle \varphi_0, \psi_0 \rangle, \langle \varphi_1, \psi_1 \rangle \rangle \,.$$

(g) Let $\langle \varphi_0, \psi_0 \rangle \in \Delta(L_0)[\Sigma]$ where $\varphi_0 \in L_0[\Gamma_0]$ and $\psi_0 \in L_0[\Gamma'_0]$. Since $L_0 \leq^1_{\mathrm{pr}} L_1$ we can find formulae $\varphi_1 \in L_1[\Gamma_1]$ and $\psi_1 \in L_1[\Gamma'_1]$ such that

$$\begin{aligned} &\mathrm{Mod}(\varphi_0) = \mathrm{pr}_{\Gamma_0}(\mathrm{Mod}(\varphi_1)) \\ \text{and}\quad &\mathrm{Mod}(\psi_0) = \mathrm{pr}_{\Gamma'_0}(\mathrm{Mod}(\psi_1)) \,. \end{aligned}$$

Since

$$\begin{aligned}\mathrm{pr}_\Sigma(\mathrm{Mod}(\varphi_1)) &= \mathrm{pr}_\Sigma(\mathrm{Mod}(\varphi_0)) \\ &= \mathrm{Str}[\Sigma] \setminus \mathrm{pr}_\Sigma(\mathrm{Mod}(\psi_0)) \\ &= \mathrm{Str}[\Sigma] \setminus \mathrm{pr}_\Sigma(\mathrm{Mod}(\psi_1))\end{aligned}$$

and L_1 has the Δ-interpolation property we can find a formula $\vartheta \in L_1[\Sigma]$ such that

$$\mathrm{Mod}(\vartheta) = \mathrm{pr}_\Sigma(\mathrm{Mod}(\varphi_1)) .$$

It follows that $\vartheta \equiv \langle \varphi_0, \psi_0 \rangle$.

(h) We only prove the first equation. The other ones are left as an exercise. Let $\langle \varphi, \psi \rangle \in \Delta(L)[\Sigma]$ where $\varphi \in L[\Sigma_0]$ and $\psi \in L[\Sigma_1]$. Then there exist formulae $\varphi' \in L[\Gamma_0]$ and $\psi' \in L[\Gamma_1]$ where $\Gamma_i \subseteq \Sigma_i$ are subsignatures of size $|\Gamma_i| \leq \mathrm{occ}(L)$ such that $\varphi' \equiv \varphi$ and $\psi' \equiv \psi$. Let $\Gamma := \Sigma \cap (\Gamma_0 \cup \Gamma_1)$. It follows that $\langle \varphi, \psi \rangle \equiv \langle \varphi', \psi' \rangle \in L[\Gamma]$ where $|\Gamma| \leq \mathrm{occ}(L)$. ☐

Proposition 5.7. $\mathrm{FO}_{\kappa^+\aleph_0}(\exists^\kappa) \leq^\kappa_{\mathrm{pr}} \mathrm{FO}_{\kappa^+\aleph_0}$.

Proof. Let $\varphi \in \mathrm{FO}_{\kappa^+\aleph_0}(\exists^\kappa)$. Following Chang's Reduction we introduce a new relation symbol R_ψ, for every subformula $\psi(\bar{x})$ of φ, and we write down formulae ensuring that R_ψ is the set of all tuples satisfying ψ. For the operations of $\mathrm{FO}_{\kappa^+\aleph_0}$ this can be done in the same way as in Chang's reduction. For a subformula $\exists^\kappa y \psi(\bar{x}, y)$, we introduce a new relation symbol $<_\psi$ and κ new function symbols f^α_ψ, $\alpha < \kappa$, and we add the formulae

$$\forall \bar{x}\Big(R_{\exists^\kappa y\psi}\bar{x} \leftrightarrow \bigwedge_{\alpha\neq\beta} f^\alpha_\psi \bar{x} \neq f^\beta_\psi \bar{x}\Big) \wedge \forall \bar{x} \bigwedge_{\alpha<\kappa} R_\psi \bar{x} f^\alpha_\psi \bar{x} ,$$

$$\forall \bar{x}\Big(\neg R_{\exists^\kappa y\psi}\bar{x} \leftrightarrow \bigvee_{\alpha<\kappa} \chi_\alpha(\bar{x})\Big) ,$$

where $\chi_\alpha(\bar{c})$ is the formula of Lemma C1.1.7 stating that the relation $\{ \langle a, b \rangle \mid \bar{c}a <_\psi \bar{c}b \}$ is a well-order of order type α on the set defined by $\psi(\bar{c}, y)$. Note that the first formula ensures that $R_{\exists^\kappa y\psi}$ contains only tuples $\bar{c}$ such that there are at least κ elements satisfying $\psi(\bar{c}, y)$, while the second formula ensures that all such tuples $\bar{c}$ are contained in $R_{\exists^\kappa y\psi}$. Finally, note that we have introduced at most κ formulae since φ has at most that many subformulae. □

Proposition 5.8. $\mathrm{FO}_{\aleph_2\aleph_0}(\exists^{\aleph_1})$ *does not have the Karp property.*

Proof. We consider the structures $\mathfrak{A} := \langle A \rangle$ and $\mathfrak{B} := \langle B \rangle$ over the empty signature with $|A| = \aleph_0$ and $|B| = \aleph_1$. Then we have

$$\mathrm{pIso}_{\aleph_0}(\mathfrak{A}, \mathfrak{B}) : \mathfrak{A} \cong_\infty \mathfrak{B}\,.$$

But $\quad \mathfrak{A} \nvDash \exists^{\aleph_1} x(x = x) \quad$ and $\quad \mathfrak{B} \vDash \exists^{\aleph_1} x(x = x)$

implies that $\mathfrak{A} \not\equiv_{\mathrm{FO}_{\aleph_2\aleph_0}(\exists^{\aleph_1})} \mathfrak{B}$. □

Corollary 5.9. $\Delta(\mathrm{FO}_{\aleph_2\aleph_0})$ *does not have the Karp property.*

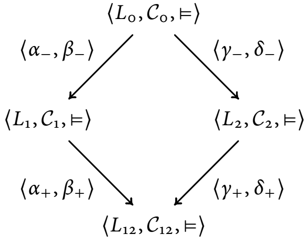
Proof. Note that

$$\mathrm{FO}_{\kappa^+\aleph_0}(\exists^\kappa) \leq^\kappa_{\mathrm{pr}} \mathrm{FO}_{\kappa^+\aleph_0} \quad \text{implies} \quad \mathrm{FO}_{\kappa^+\aleph_0}(\exists^\kappa) \leq^1_{\mathrm{pr}} \mathrm{FO}_{\kappa^+\aleph_0}$$

since $\mathrm{FO}_{\kappa^+\aleph_0}$ is closed under conjunctions of size κ. By Lemma 5.6 (b), it follows that $\mathrm{FO}_{\aleph_2\aleph_0}(\exists^{\aleph_1}) \leq \Delta(\mathrm{FO}_{\aleph_2\aleph_0})$. Since the former does not have the Karp property it follows that the latter does not have it either. □

For many logics that can be characterised via a preservation theorem we can derive the interpolation property from a general theorem which we will present below. Instead of considering the entailment relation $\Phi_0 \vDash \Phi_1$ for a single logic, we allow Φ_0 and Φ_1 to belong to different logics L_1 and L_2, and we look for an interpolant $\Phi_0 \vDash \Psi \vDash \Phi_1$ in a third logic L_0.

Definition 5.10. (a) A *weak amalgamation square* is a commuting diagram



in the category $\mathfrak{Logic}$ such that, for every pair $\mathfrak{I}_1 \in \mathcal{C}_1$ and $\mathfrak{I}_2 \in \mathcal{C}_2$ of interpretations with $\beta_-(\mathfrak{I}_1) = \delta_-(\mathfrak{I}_2)$, there exists an L_{12}-interpretation $\mathfrak{I}_{12}$ with

$$\beta_+(\mathfrak{I}_{12}) = \mathfrak{I}_1 \quad \text{and} \quad \delta_+(\mathfrak{I}_{12}) = \mathfrak{I}_2 \,.$$

(b) Given a weak amalgamation square as in (a) and sets $\Phi_1 \subseteq L_1$ and $\Phi_2 \subseteq L_2$ of formulae with $\alpha_+[\Phi_1] \vDash \gamma_+[\Phi_2]$, we call a set $\Phi_0 \subseteq L_0$ an *interpolant* of Φ_1 and Φ_2 if

$$\Phi_1 \vDash \alpha_-[\Phi_0] \quad \text{and} \quad \gamma_-[\Phi_0] \vDash \Phi_2 \,.$$

(c) Similarly, given a weak amalgamation square and classes $\mathcal{K}_1 \subseteq \mathcal{C}_1$ and $\mathcal{K}_2 \subseteq \mathcal{C}_2$ of interpretations with $\beta_+^{-1}[\mathcal{K}_1] \subseteq \delta_+^{-1}[\mathcal{K}_2]$ we call a class $\mathcal{K}_0 \subseteq \mathcal{C}_0$ an *interpolant* of $\mathcal{K}_1$ and $\mathcal{K}_2$ if

$$\mathcal{K}_1 \subseteq \beta_-^{-1}[\mathcal{K}_0] \quad \text{and} \quad \delta_-^{-1}[\mathcal{K}_0] \subseteq \mathcal{K}_2 \,.$$

Lemma 5.11. *Φ_0 is an interpolant of Φ_1 and Φ_2 if and only if* $\mathrm{Mod}(\Phi_0)$ *is an interpolant of* $\mathrm{Mod}(\Phi_1)$ *and* $\mathrm{Mod}(\Phi_2)$.

The next lemma shows that each pair of classes in a weak amalgamation square has an interpolant. For the interpolation property to hold we have to strengthen this result by proving that a pair of *axiomatisable* classes has an *axiomatisable* interpolant.

Lemma 5.12. *Consider a weak amalgamation square as above.*

(a) *$\mathcal{K}_0$ is an interpolant of $\mathcal{K}_1$ and $\mathcal{K}_2$ if and only if*

$$\beta_-[\mathcal{K}_1] \subseteq \mathcal{K}_0 \subseteq \mathcal{C}_0 \setminus \gamma_-[\mathcal{C}_2 \setminus \mathcal{K}_2]\,.$$

(b) *$\mathcal{K}_1$ and $\mathcal{K}_2$ have an interpolant.*

Proof. (a) We have

$$\mathcal{K}_1 \subseteq \beta_-^{-1}[\mathcal{K}_0] \quad \text{iff} \quad \beta_-[\mathcal{K}_1] \subseteq \mathcal{K}_0\,,$$

and $\mathcal{C}_2 \setminus \mathcal{K}_2 \subseteq \gamma_-^{-1}[\mathcal{C}_0 \setminus \mathcal{K}_0]$ iff $\gamma_-[\mathcal{C}_2 \setminus \mathcal{K}_2] \subseteq \mathcal{C}_0 \setminus \mathcal{K}_0\,.$

(b) By (a) it is sufficient to show that

$$\beta_-[\mathcal{K}_1] \subseteq \mathcal{C}_0 \setminus \gamma_-[\mathcal{C}_2 \setminus \mathcal{K}_2]\,.$$

For a contradiction, suppose that there is some interpretation

$$\mathfrak{J}_0 \in \beta_-[\mathcal{K}_1] \setminus (\mathcal{C}_0 \setminus \gamma_-[\mathcal{C}_2 \setminus \mathcal{K}_2]) = \beta_-[\mathcal{K}_1] \cap \gamma_-[\mathcal{C}_2 \setminus \mathcal{K}_2]\,.$$

Choose interpretations $\mathfrak{J}_1 \in \mathcal{K}_1$ and $\mathfrak{J}_2 \in \mathcal{C}_2 \setminus \mathcal{K}_2$ with

$$\beta_-(\mathfrak{J}_1) = \mathfrak{J}_0 = \gamma_-(\mathfrak{J}_2)\,.$$

Since the diagram is a weak amalgamation square we can find an interpretation $\mathfrak{J}_{12} \in \mathcal{C}_{12}$ with $\beta_+(\mathfrak{J}_{12}) = \mathfrak{J}_1$ and $\gamma_+(\mathfrak{J}_{12}) = \mathfrak{J}_2$. It follows that

$$\mathfrak{J}_{12} \in \beta_+^{-1}(\mathfrak{J}_1) \subseteq \beta_+^{-1}[\mathcal{K}_1] \subseteq \gamma_+^{-1}[\mathcal{K}_2]\,.$$

Consequently, we have $\mathfrak{J}_2 = \gamma_+(\mathfrak{J}_{12}) \in \mathcal{K}_2$. Contradiction. □

If a logic can be characterised by a preservation theorem then a class of interpretation is axiomatisable if and only if it is a fixed point for the operations the logic is preserved under. Hence, to prove our interpolation theorem we consider fixed points of operations.

Definition 5.13. Let A and B be sets and $\alpha : \wp(A) \to \wp(A)$ and $\beta : \wp(B) \to \wp(B)$ functions on their power sets.

(a) A function $f : A \to B$ *preserves fixed points* of α and β if

$$C \in \operatorname{fix} \alpha \quad \text{implies} \quad f[C] \in \operatorname{fix} \beta\,.$$

(b) A function $f : A \to B$ *lifts α to β* if

$$(f^{-1} \circ \beta)[X] \subseteq (\alpha \circ f^{-1})[X]\,, \quad \text{for all } X \subseteq B\,.$$

Theorem 5.14 (Popescu, Roşu, Şerbănuţă)**.** *Consider a weak amalgamation square*

$$\begin{array}{ccccc} & & \langle L_0, \mathcal{C}_0, \vDash \rangle & & \\ & \langle \alpha_-, \beta_- \rangle \swarrow & & \searrow \langle \gamma_-, \delta_- \rangle & \\ \langle L_1, \mathcal{C}_1, \vDash \rangle & & & & \langle L_2, \mathcal{C}_2, \vDash \rangle \\ & \langle \alpha_+, \beta_+ \rangle \searrow & & \swarrow \langle \gamma_+, \delta_+ \rangle & \\ & & \langle L_{12}, \mathcal{C}_{12}, \vDash \rangle & & \end{array}$$

Suppose that there are functions

$$\mu_i : \wp(\mathcal{C}_i) \to \wp(\mathcal{C}_i)\,, \quad \text{for } i \in \{0,1\}\,,$$
and $$\nu_j : \wp(\mathcal{C}_j) \to \wp(\mathcal{C}_j)\,, \quad \text{for } j \in \{0,2\}\,,$$

satisfying the following conditions:

(1) $\mu_0 \circ \nu_0 \circ \mu_0 = \nu_0 \circ \mu_0\,.$

(2) ν_0 *and* ν_2 *are closure operators.*

(3) β_- *preserves fixed points of* μ_1 *and* μ_0.

(4) γ_- *lifts* ν_2 *to* ν_0.

Every pair of fixed points $\mathcal{K}_1 \in \operatorname{fix} \mu_1$ *and* $\mathcal{K}_2 \in \operatorname{fix} \nu_2$ *with*

$$\beta_+^{-1}[\mathcal{K}_1] \subseteq \gamma_+^{-1}[\mathcal{K}_2]$$

has an interpolant $\mathcal{K}_0 \in \operatorname{fix} \mu_0 \cap \operatorname{fix} \nu_0$.

Proof. We claim that $\mathcal{K}_0 := (\nu_0 \circ \beta_-)[\mathcal{K}_1]$ is the desired interpolant. We have $\mathcal{K}_0 \in \operatorname{fix} \nu_0$ since $\nu_0 \circ \nu_0 = \nu_0$. Furthermore, $\beta_-[\mathcal{K}_1] \in \operatorname{fix} \mu_0$ as $\mathcal{K}_1 \in \operatorname{fix} \mu_1$ and β_- preserves fixed points. It follows that

$$\begin{aligned}\mu_0[\mathcal{K}_0] &= (\mu_0 \circ \nu_0 \circ \beta_-)[\mathcal{K}_1] \\ &= (\mu_0 \circ \nu_0 \circ \mu_0 \circ \beta_-)[\mathcal{K}_1] \\ &= (\nu_0 \circ \mu_0 \circ \beta_-)[\mathcal{K}_1] \\ &= (\nu_0 \circ \beta_-)[\mathcal{K}_1] = \mathcal{K}_0\,,\end{aligned}$$

and, therefore, $\mathcal{K}_0 \in \operatorname{fix} \mu_0$.

It remains to prove that $\mathcal{K}_0$ is an interpolant. Since ν_0 is a closure operator we have $\beta_-[\mathcal{K}_1] \subseteq (\nu_0 \circ \beta_-)[\mathcal{K}_1] = \mathcal{K}_0$ which implies that $\mathcal{K}_1 \subseteq \beta_-^{-1}[\mathcal{K}_0]$. For the other inclusion, note that we have

$$\gamma_-^{-1}[\mathcal{K}_0] = (\gamma_-^{-1} \circ \nu_0 \circ \beta_-)[\mathcal{K}_1] \subseteq (\nu_2 \circ \gamma_-^{-1} \circ \beta_-)[\mathcal{K}_1]$$

since γ_- lifts ν_2 to ν_0. Furthermore, $(\gamma_-^{-1} \circ \beta_-)[\mathcal{K}_1] \subseteq \mathcal{K}_2$ since we have shown in Lemma 5.12 that $\beta_-[\mathcal{K}_1]$ is an interpolant of $\mathcal{K}_1$ and $\mathcal{K}_2$. As ν_2 is a closure operator it follows that

$$(\nu_2 \circ \gamma_-^{-1} \circ \beta_-)[\mathcal{K}_1] \subseteq \nu_2[\mathcal{K}_2] = \mathcal{K}_2\,.$$

Consequently, we have

$$\gamma_-^{-1}[\mathcal{K}_0] \subseteq (\nu_2 \circ \gamma_-^{-1} \circ \beta_-)[\mathcal{K}_1] \subseteq \mathcal{K}_2\,,$$

as desired. □

We can use this theorem to obtain interpolation results for logics that can be characterised via preservation theorems.

Corollary 5.15. *Consider a weak amalgamation square as above and functions*

$$\begin{aligned}&\mu_i : \wp(\mathcal{C}_i) \to \wp(\mathcal{C}_i)\,, \quad \textit{for } i \in \{0,1\}\,,\\ \textit{and} \quad &\nu_j : \wp(\mathcal{C}_j) \to \wp(\mathcal{C}_j)\,, \quad \textit{for } j \in \{0,2\}\,,\end{aligned}$$

satisfying the conditions of the preceding theorem. Furthermore, suppose that

(1) *a class* $\mathcal{K}_1 \subseteq \mathcal{C}_1$ *is* L_1*-axiomatisable if and only if* $\mathcal{K}_1 \in \operatorname{fix} \mu_1$*;*

(2) *a class* $\mathcal{K}_2 \subseteq \mathcal{C}_2$ *is* L_2*-axiomatisable if and only if* $\mathcal{K}_2 \in \operatorname{fix} \nu_2$*;*

(3) *a class* $\mathcal{K}_0 \subseteq \mathcal{C}_0$ *is* L_0*-axiomatisable iff* $\mathcal{K}_0 \in \operatorname{fix} \mu_0 \cap \operatorname{fix} \nu_0$.

Then every pair of sets $\Phi_1 \subseteq L_1$ *and* $\Phi_2 \subseteq L_2$ *with*

$$\alpha_+[\Phi_1] \vDash \delta_+[\Phi_2]$$

has an interpolant $\Phi_0 \subseteq L_0$.

Unfortunately applications of this theorem will have to wait till Chapter D2 since at the moment we still lack the required preservation theorems.

6. Fixed-point logics

As an example we investigate extensions of first-order logic by fixed-point operators. Let $\mathfrak{A}$ be a structure and $f : \wp(A^n) \to \wp(A^n)$ a function. A fixed point of f is an n-ary relation on A. We are interested in operators that compute such fixed points for definable functions f.

Note that the partial order $\wp(A^n)$ is complete. Hence, if f is increasing then, by Theorem A2.4.3, it has a least fixed point $\operatorname{lfp} f$ and a greatest fixed point $\operatorname{gfp} f$. Similarly, if f is inflationary then we can use Theorem A3.3.14 to obtain the inductive fixed point $\operatorname{ifp} f$ of f over $\varnothing$.

If f is neither increasing nor inflationary then none of these fixed points need to exist. But we still would like to define a fixed point operator for such functions. Instead of asking for a real fixed point we will present two ways to compute an approximate one.

Firstly, we can artificially make f inflationary by replacing it with the function $x \mapsto x \cup f(x)$. Secondly, we can compute the 'fixed-point induction' $\varnothing, f(\varnothing), f^2(\varnothing), \dots$ (which generally is not increasing) and take some kind of limit.

Definition 6.1. Let X be a set and $f : \wp(X) \to \wp(X)$ an arbitrary function.

(a) The *inductive fixed point* $\operatorname{ifp} f$ of f is the inductive fixed point of the function $f' : x \mapsto x \cup f(x)$ over $\varnothing$. Correspondingly, by the *inductive fixed-point induction* of f we mean the fixed-point induction $F : \mathrm{On} \to \wp(X)$ of f' over $\varnothing$.

(b) The *lower fixed-point induction* of f is the map $F_- : \mathrm{On} \to \wp(X)$ defined by

$$\begin{aligned} F_-(0) &:= \varnothing\,, \\ F_-(\alpha+1) &:= f(F_-(\alpha))\,, \\ F_-(\delta) &:= \bigcup_{\alpha<\delta} \bigcap_{\alpha\leq\beta<\delta} F_-(\beta)\,, \quad \text{for limits } \delta\,. \end{aligned}$$

Analogously, we define the *upper fixed-point induction* F_+ by

$$\begin{aligned} F_+(0) &:= \varnothing\,, \\ F_+(\alpha+1) &:= f(F_+(\alpha))\,, \\ F_+(\delta) &:= \bigcap_{\alpha<\delta} \bigcup_{\alpha\leq\beta<\delta} F_+(\beta)\,, \quad \text{for limits } \delta\,. \end{aligned}$$

(c) The *least partial fixed point* $\liminf f$ of f is the set

$$F_-(\infty) := \bigcup_{\alpha} \bigcap_{\alpha\leq\beta} F_-(\beta)\,.$$

and its *greatest partial fixed point* $\limsup f$ is

$$F_+(\infty) := \bigcap_{\alpha} \bigcup_{\alpha\leq\beta} F_+(\beta)\,.$$

Remark. Note that, in general, $\operatorname{ifp} f$, $\liminf f$, and $\limsup f$ are *no* fixed points of f. But, if f is increasing then $\operatorname{ifp} f = \liminf f = \limsup f = \operatorname{lfp} f$.

Before defining logics with these fixed-point operators let us compute their closure ordinals.

Definition 6.2. Let $f : \wp(X) \to \wp(X)$ be a function.

(a) The *closure ordinal* for the inductive fixed-point induction F of f is the least ordinal α such that $F(\alpha) = F(\alpha + 1)$.

(b) The *closure ordinal* for the lower fixed-point induction F_- of f is the least ordinal α such that

$$F_-(\alpha) = F_-(\infty) \quad \text{and} \quad F_-(\beta) \supseteq F_-(\alpha)\,, \quad \text{for all } \beta \geq \alpha\,.$$

Similarly, we define the closure ordinal for the upper fixed-point induction F_+ as the least ordinal α such that

$$F_+(\alpha) = F_+(\infty) \quad \text{and} \quad F_+(\beta) \subseteq F_+(\alpha)\,, \quad \text{for all } \beta \geq \alpha\,.$$

Since the inductive fixed-point induction of a function is increasing we obtain the same bound on the closure ordinal as for least fixed points.

Lemma 6.3. *Let $f : \wp(X) \to \wp(X)$. The closure ordinal of* ifp f *is less than $|X|^+$.*

For partial fixed points the situation is different. The following sequence of lemmas shows that in this case the bound is $(2^{|X|})^+$. We will only consider the case of upper fixed-point inductions. The closure ordinal of a least partial fixed point can be computed in exactly the same way.

Lemma 6.4. *Let F_+ be the upper fixed-point induction of the function $f : \wp(X) \to \wp(X)$.*

(a) *If $F_+(\alpha) = F_+(\beta)$ then $F_+(\alpha + \gamma) = F_+(\beta + \gamma)$, for all γ.*

(b) *If $F_+(\alpha) = F_+(\alpha + \beta)$ then*

$$F_+(\alpha + \beta n) = F_+(\alpha)\,, \qquad \text{for all } n < \omega\,,$$
$$\text{and} \quad F_+(\alpha + \beta\omega) = \bigcup_{\gamma<\beta} F_+(\alpha + \gamma)\,.$$

(c) *If $F_+(\alpha) = F_+(\alpha + \beta) = F_+(\alpha + \gamma)$ and $\beta \leq \gamma$ then*

$$F_+(\alpha + \beta\omega) \subseteq F_+(\alpha + \gamma\omega)\,.$$

(d) *If* $F_+(\alpha) = F_+(\alpha+\beta) = F_+(\alpha+\beta\omega)$ *then*

$$F_+(\alpha+\beta\gamma) = F_+(\alpha), \qquad \text{for all } \gamma,$$
$$\text{and} \qquad F_+(\infty) = F_+(\alpha).$$

Proof. (a) is proved by a straightforward induction on γ. For $\gamma = 0$, there is nothing to do. If $\gamma = \eta + 1$ then

$$F_+(\alpha+\eta+1) = f(F_+(\alpha+\eta)) = f(F_+(\beta+\eta)) = F_+(\beta+\eta+1).$$

Finally, for limit ordinals γ we have

$$\begin{aligned} F_+(\alpha+\gamma) &= \bigcap_{i<\alpha+\gamma} \bigcup_{i\le k<\alpha+\gamma} F_+(k) \\ &= \bigcap_{\alpha\le i<\alpha+\gamma} \bigcup_{i\le k<\alpha+\gamma} F_+(k) \\ &= \bigcap_{i<\gamma} \bigcup_{i\le k<\gamma} F_+(\alpha+k) \\ &= \bigcap_{i<\gamma} \bigcup_{i\le k<\gamma} F_+(\beta+k) \\ &= \bigcap_{i<\beta+\gamma} \bigcup_{i\le k<\beta+\gamma} F_+(k) = F_+(\beta+\gamma). \end{aligned}$$

(b) The first equation follows by induction on n. For $n = 0$ there is nothing to do. For $n > 0$, it follows from (a) and the inductive hypothesis that

$$F_+(\alpha+\beta n) = F_+(\alpha+\beta(n-1)+\beta) = F_+(\alpha+\beta) = F_+(\alpha).$$

For the second equation, we have

$$\begin{aligned} F_+(\alpha+\beta\omega) &= \bigcap_{n<\omega} \bigcup_{n\le k<\omega} \bigcup_{\gamma<\beta} F_+(\alpha+\beta k+\gamma) \\ &= \bigcap_{n<\omega} \bigcup_{n\le k<\omega} \bigcup_{\gamma<\beta} F_+(\alpha+\gamma) = \bigcup_{\gamma<\beta} F_+(\alpha+\gamma). \end{aligned}$$

(c) By (b), we have

$$F_+(\alpha + \beta\omega) = \bigcup_{i<\beta} F_+(\alpha + i) \subseteq \bigcup_{i<\gamma} F_+(\alpha + i) = F_+(\alpha + \gamma\omega)\,.$$

(d) Again, we use induction on γ. For $\gamma = 0$ there is nothing to do and the inductive step follows as in (b). If γ is a limit ordinal then we have

$$\begin{aligned}
F_+(\alpha + \beta\gamma) &= \bigcap_{i<\gamma} \bigcup_{i\leq k<\gamma} \bigcup_{l<\beta} F_+(\alpha + \beta k + l) \\
&= \bigcap_{i<\gamma} \bigcup_{i\leq k<\gamma} \bigcup_{l<\beta} F_+(\alpha + l) \\
&= \bigcup_{l<\beta} F_+(\alpha + l) = F_+(\alpha + \beta\omega) = F_+(\alpha)\,,
\end{aligned}$$

by inductive hypothesis and (b).

The second claim follows from (b) and the first claim. For one direction, note that we have

$$F_+(\alpha) = F_+(\alpha + \beta\eta\omega) = \bigcup_{\gamma<\beta\eta} F_+(\alpha + \gamma) \supseteq F_+(\eta)\,,$$

which implies that

$$F_+(\infty) = \bigcap_{\gamma\geq\alpha} \bigcup_{\eta\geq\gamma} F_+(\eta) \subseteq \bigcap_{\gamma\geq\alpha} \bigcup_{\eta\geq\gamma} F_+(\alpha) = F_+(\alpha)\,.$$

Conversely, $F_+(\alpha + \beta\gamma) \subseteq \bigcup_{\eta\geq\gamma} F_+(\eta)$ implies that

$$F_+(\alpha) = \bigcap_{\gamma} F_+(\alpha + \beta\gamma) \subseteq \bigcap_{\gamma} \bigcup_{\eta\geq\gamma} F_+(\eta) = F_+(\infty)\,.$$

□

In order to prove that there exist ordinals α and β with $F_+(\alpha) = F_+(\alpha + \beta) = F_+(\alpha + \beta\omega)$ we need some results about closed unbounded sets.

Lemma 6.5. *Let F_+ be the upper fixed-point induction of the function $f : \wp(X) \to \wp(X)$. Set $\kappa := |X|$ and $\lambda := (2^\kappa)^+ \oplus \aleph_1$.*

(a) *Suppose that $\alpha < \lambda$ and $S \subseteq \lambda$ is a cofinal set such that*

$$F_+(\alpha + \beta) = F_+(\alpha), \quad \text{for all } \beta \in S.$$

If there is no $\beta \in S$ such that $F_+(\alpha + \beta\omega) = F_+(\alpha)$ then there exists an ordinal α' and a cofinal set $S' \subseteq \lambda$ such that

$$F_+(\alpha') \supset F_+(\alpha) \quad \text{and} \quad F_+(\alpha' + \beta') = F_+(\alpha') \text{ for all } \beta' \in S'.$$

(b) *There exist ordinals $\alpha, \beta < \lambda$ such that*

$$F_+(\alpha) = F_+(\alpha + \beta) = F_+(\alpha + \beta\omega).$$

Proof. (a) Since λ is regular we have $|S| = \lambda$. Let $(\beta_i)_{i<\lambda}$ be an increasing enumeration of $S \setminus \{0\}$. By Lemma 6.4 (c) it follows that $F_+(\alpha + \beta_i\omega) \subseteq F_+(\alpha + \beta_k\omega)$, for all $i \leq k$. Consequently, there is some index $m < \lambda$ such that $F_+(\alpha + \beta_i\omega) = F_+(\alpha + \beta_m\omega)$, for all $i \geq m$. Set $\alpha' := \alpha + \beta_m\omega$ and

$$S' := \{ \beta_i\omega \mid i \geq m \}.$$

By assumption, we have

$$F_+(\alpha) \neq F_+(\alpha + \beta_m\omega) = \bigcup_{\gamma<\beta_m} F_+(\alpha + \gamma) \supseteq F_+(\alpha),$$

which implies that $F_+(\alpha') \supset F_+(\alpha)$.

(b) For $Z \subseteq X$, let

$$S(Z) := \{ \alpha < \lambda \mid Z \subseteq F_+(\alpha) \}.$$

We construct a strictly increasing sequence of sets $(Z_i)_{i<\eta}$ such that each set $S(Z_i)$ is closed unbounded in λ. Let $Z_0 := \emptyset$. Then $S(Z_0) = \lambda$. For limit ordinals δ, set $Z_\delta := \bigcup_{i<\delta} Z_i$. By Proposition A4.6.4, it follows that $S(Z_\delta) = \bigcap_{i<\delta} S(Z_i)$ is closed unbounded.

For the successor step, suppose that we have already defined Z_i. Since $|\wp(X)| < \lambda$ we can find a set $Y \supseteq Z_i$ such that the set

$$P := \{ \alpha < \lambda \mid F_+(\alpha) = Y \}$$

is cofinal. Let α be the minimal element of P and set

$$Q := \{ \beta \mid \alpha + \beta \in P,\ \beta > 0 \}.$$

If there is some $\beta \in Q$ with $F_+(\alpha + \beta\omega) = F_+(\alpha)$ then we are done. Otherwise, we can use (a) to find an ordinal α' and a cofinal subset $Q' \subseteq \lambda$ such that $F_+(\alpha') \supset F_+(\alpha)$ and $F_+(\alpha' + \beta') = F_+(\alpha')$, for all $\beta' \in Q'$. We set $Z_{i+1} := F_+(\alpha')$. It remains to show that $S(Z_{i+1})$ is closed unbounded.

By construction the set $S(Z_{i+1}) \supseteq \{ \alpha + \beta \mid \beta \in Q' \}$ is cofinal. Let $X \subseteq S(Z_{i+1})$ be a subset with $\sup X < \lambda$. If $\sup X \in X$ then we are done. Otherwise, $\delta := \sup X$ is a limit ordinal and $F_+(\delta) = \bigcap_{\alpha<\delta} \bigcup_{\alpha \leq \beta < \delta} F_+(\beta)$. Since, for every $\beta < \delta$, there is some ordinal $\beta \leq \gamma < \delta$ with $F_+(\gamma) \supseteq Z_{i+1}$ it follows that $F_+(\delta) \supseteq Z_{i+1}$. Hence, $\delta \in S(Z_{i+1})$.

We continue this construction until we either find indices α and β such that $F_+(\alpha) = F_+(\alpha + \beta) = F_+(\alpha + \beta\omega)$ or we have defined Z_i, for all $i < \lambda$. In the former case we are done. The latter case cannot happen since $Z_i \subset Z_k$, for $i < k$ and there are less than λ subsets of X. □

Corollary 6.6. *Let X be a set of size $\kappa := |X|$ and let F_+ be the upper fixed-point induction of $f : \wp(X) \to \wp(X)$. Set $\lambda := (2^\kappa)^+ \oplus \aleph_1$.*

(a) *There exists some ordinal $\alpha < \lambda$ such that $F_+(\infty) = F(\alpha)$.*

(b) $F_+(\infty) = F(\lambda)$.

Proof. By the preceding lemma, we can find ordinals $\alpha, \beta < (2^\kappa)^+$ such that $F_+(\alpha) = F_+(\alpha + \beta) = F_+(\alpha + \beta\omega)$. It follows by Lemma 6.4 (d) that $F_+(\infty) = F_+(\alpha)$. Furthermore, since $\lambda := (2^\kappa)^+$ is regular we have $F_+(\lambda) = F_+(\alpha + \beta\lambda) = F_+(\alpha) = F_+(\infty)$. □

In order to add these fixed-point operators to first-order logic we start by looking at definable functions $\wp(A^n) \to \wp(A^n)$.

Definition 6.7. Let $\varphi(R, \bar{x}, \bar{y}) \in L[\Sigma \cup \{R\}, X]$ be a formula where R is an n-ary relation symbol and $\bar{x}$ a tuple of n variables.

(a) Given a Σ-structure $\mathfrak{A}$ and parameters $\bar{c} \subseteq A$, φ defines the function

$$f_\varphi : \wp(A^n) \to \wp(A^n) : R \mapsto \{\, \bar{a} \in A^n \mid \mathfrak{A} \models \varphi(R, \bar{a}, \bar{c}) \,\}\,.$$

(b) We say that the relation symbol R occurs *positively* in φ if every occurrence of R is in the scope of an even number of negation symbols. If R only appears in the scope of odd numbers of negation symbols, we say that it occurs *negatively* in φ.

Depending on which fixed-point operators we add we obtain several extensions of first-order logic.

Definition 6.8. (a) *Least fixed-point logic* $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{LFP})$ is the extension of $\mathrm{FO}_{\kappa\aleph_0}$ by formulae of the form

$$[\mathrm{lfp}\, R\bar{x} : \varphi(R, \bar{x}, \bar{y})](\bar{z}) \quad \text{and} \quad [\mathrm{gfp}\, R\bar{x} : \varphi(R, \bar{x}, \bar{y})](\bar{z})$$

where we require that the relation R appears *positively* in φ. The semantics is defined by

$$\begin{aligned} \mathfrak{A} &\models [\mathrm{lfp}\, R\bar{x} : \varphi(R, \bar{x}, \bar{c})](\bar{a}) \quad &&:\text{iff} \quad \bar{a} \in \mathrm{lfp}\, f_\varphi\,, \\ \mathfrak{A} &\models [\mathrm{gfp}\, R\bar{x} : \varphi(R, \bar{x}, \bar{c})](\bar{a}) \quad &&:\text{iff} \quad \bar{a} \in \mathrm{gfp}\, f_\varphi\,. \end{aligned}$$

(b) *Inflationary fixed-point logic* $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{IFP})$ is the extension of $\mathrm{FO}_{\kappa\aleph_0}$ by formulae of the form

$$[\mathrm{ifp}\, R\bar{x} : \varphi(R, \bar{x}, \bar{y})](\bar{z})$$

with the semantics

$$\mathfrak{A} \models [\mathrm{ifp}\, R\bar{x} : \varphi(R, \bar{x}, \bar{c})](\bar{a}) \quad :\text{iff} \quad \bar{a} \in \mathrm{ifp}\, f_\varphi\,.$$

(c) *Partial fixed-point logic* $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{PFP})$ is the extension of $\mathrm{FO}_{\kappa\aleph_0}$ by formulae of the form

$$[\liminf R\bar{x} : \varphi(R, \bar{x}, \bar{y})](\bar{z})$$

and $[\limsup R\bar{x} : \varphi(R, \bar{x}, \bar{y})](\bar{z})$

with the semantics

$$\mathfrak{A} \models [\liminf R\bar{x} : \varphi(R, \bar{x}, \bar{c})](\bar{a}) \quad :\text{iff} \quad \bar{a} \in \liminf f_\varphi\,,$$
$$\mathfrak{A} \models [\limsup R\bar{x} : \varphi(R, \bar{x}, \bar{c})](\bar{a}) \quad :\text{iff} \quad \bar{a} \in \limsup f_\varphi\,.$$

The requirement on φ in the definition of $[\text{lfp}\, \bar{R}x : \varphi]$ ensures that the least fixed point of f_φ does exist.

Lemma 6.9. *If $\varphi(R, \bar{x}, \bar{y}) \in \text{FO}_{\kappa\aleph_0}(\text{LFP})$ is a formula where the relation symbol R appears only positively then f_φ is increasing.*

Proof. One can show by a trivial induction on φ that, if $R \subseteq R'$ then

$$\mathfrak{A} \models \varphi(R, \bar{a}, \bar{c}) \quad \text{implies} \quad \mathfrak{A} \models \varphi(R', \bar{a}, \bar{c})\,. \qquad \square$$

Example. We can express in FO(LFP) that a relation $<$ is well-founded by the formula

$$\varphi_{\text{wf}} := \forall x[\text{lfp}\, Px : (\forall y. y < x)Py](x)\,.$$

The α-th stage of the fixed-point induction of this formula contains all elements of foundation rank less than α.

Remark. Note that, by duality, we have

$$[\text{gfp}\, R\bar{x} : \varphi(R, \bar{x})](\bar{z}) \equiv \neg[\text{lfp}\, R\bar{x} : \neg\varphi(\neg R, \bar{x})](\bar{z})\,,$$

where $\varphi(\neg R)$ is the formula obtained form φ by negating every atom of the form $R\bar{t}$.

Lemma 6.10. $\text{FO}_{\kappa\aleph_0} \leq \text{FO}_{\kappa\aleph_0}(\text{LFP}) \leq \text{FO}_{\kappa\aleph_0}(\text{IFP}) \leq \text{FO}_{\kappa\aleph_0}(\text{PFP})$

Proof. Clearly, $\text{FO}_{\kappa\aleph_0}(\text{LFP})$ is as least as expressive as $\text{FO}_{\kappa\aleph_0}$. For the second inclusion note that, if f is an increasing function then $\text{ifp}\, f = \text{lfp}\, f$. Hence, we can simulate each least fixed point $[\text{lfp}\, R\bar{x} : \varphi]$ by the formula $[\text{ifp}\, R\bar{x} : \varphi]$. Similarly, we have

$$[\text{ifp}\, R\bar{x} : \varphi](\bar{z}) \equiv [\liminf R\bar{x} : R\bar{x} \vee \varphi](\bar{z})\,,$$

since the fixed-point inductions of both fixed points coincide. $\square$

In order to compare fixed-point logics with infinitary first-order logic we construct formulae defining the various stages of a fixed point.

Definition 6.11. Let $\varphi(R, \bar{x})$ be a formula and α an ordinal.

(a) The α-th lfp-*approximation* of φ is defined by induction on α as

$$\begin{aligned} \varphi^0(\bar{x}) &:= \text{false}\,, \\ \varphi^{\alpha+1}(\bar{x}) &:= \varphi[R/\varphi^\alpha]\,, \\ \varphi^\delta(\bar{x}) &:= \bigvee_{\alpha<\delta} \varphi[R/\varphi^\alpha]\,, \quad \text{for limits } \delta\,, \end{aligned}$$

where $\varphi[R/\psi]$ denotes the formula obtained from φ by replacing every atom $R\bar{t}$ by the formula $\psi(\bar{t})$.

(b) The α-th ifp-*approximation* of φ is the same as the α-th lfp-approximation of the formula $R\bar{x} \vee \varphi$.

(c) The α-th lim inf-*approximation* of φ is the formula defined by

$$\begin{aligned} \varphi^0(\bar{x}) &:= \text{false}\,, \\ \varphi^{\alpha+1}(\bar{x}) &:= \varphi[R/\varphi^\alpha]\,, \\ \varphi^\delta(\bar{x}) &:= \bigvee_{\alpha<\delta} \bigwedge_{i<\alpha} \varphi[R/\varphi^i]\,, \quad \text{for limits } \delta\,. \end{aligned}$$

(d) The α-th lim sup-*approximation* of φ is the formula defined by

$$\begin{aligned} \varphi^0(\bar{x}) &:= \text{false}\,, \\ \varphi^{\alpha+1}(\bar{x}) &:= \varphi[R/\varphi^\alpha]\,, \\ \varphi^\delta(\bar{x}) &:= \bigwedge_{\alpha<\delta} \bigvee_{i<\alpha} \varphi[R/\varphi^i]\,, \quad \text{for limits } \delta\,. \end{aligned}$$

Lemma 6.12. *Let φ^α be the α-th* fp-*approximation of a formula φ where* fp *is one of* lfp, ifp, lim inf, *or* lim sup. *Let $\mathfrak{A}$ be a structure and F the fixed-point induction of $[\mathrm{fp} R\bar{x} : \varphi]$ on $\mathfrak{A}$. Then we have*

$$(\varphi^\alpha)^{\mathfrak{A}} = F(\alpha)\,.$$

Lemma 6.13. *Let $\varphi \in \mathrm{FO}_{\kappa\aleph_0}(\mathrm{PFP})$. For every regular cardinal μ, there exists a formula $\psi \in \mathrm{FO}_{\lambda\aleph_0}$ where $\lambda := (2^{\mu})^{++} \oplus \kappa \oplus \aleph_2$ such that*

$$\mathfrak{A} \vDash \varphi \leftrightarrow \psi\,, \quad \text{for every structure } \mathfrak{A} \text{ of size } |A| \leq \mu\,.$$

Proof. We prove the claim by induction on φ. Hence, we may assume that $\varphi = [\limsup R\bar{x} : \chi](\bar{x})$ with $\chi \in \mathrm{FO}_{\lambda\aleph_0}$. Let χ^{α} be the α-th lim sup-approximation of χ. Let $\lambda_0 := (2^{\mu})^{+} \oplus \aleph_1$. By Corollary 6.6 and the preceding lemma it follows that the formula χ^{λ_0} defines the partial fixed point of χ on all structures of cardinality $|A| < \mu$. Finally, note that $\chi^{\lambda_0} \in \mathrm{FO}_{\lambda\aleph_0}$. □

Corollary 6.14. $\mathrm{FO}_{\infty\aleph_0}(\mathrm{PFP})$ *has the Karp property.*

In some cases the closure ordinal of a least fixed point is independent of the size of the structure.

Lemma 6.15. *Let $\varphi(R, \bar{x}, \bar{y})$ be an existential first-order formula where R occurs only positively. On every structure $\mathfrak{A}$ the least fixed point $[\mathrm{lfp}\, R\bar{x} : \varphi(R, \bar{x}, \bar{y})]$ is reached after at most ω steps.*

Proof. The corresponding function $f_{\varphi} : \wp(A)^n \to \wp(A)^n$ is continuous since $f_{\varphi}(R) = \bigcup \{ f_{\varphi}(R_0) \mid R_0 \subseteq R \text{ finite} \}$. Hence, the claim follows from Lemma A3.3.12 (c). □

In Chapter E1 we will study saturated structures. One of their many properties is the fact that, for such structures, the preceding lemma holds for all first-order formulae, not only for existential ones.

Definition 6.16. A structure $\mathfrak{A}$ is $\aleph_0$-*saturated* if $\mathfrak{A}$ realises every type $\mathfrak{p} \in S^1(U)$ where $U \subseteq A$ is finite.

Lemma 6.17. *Let $\varphi(R, \bar{x})$ be a first-order formula where R occurs only positively and let $\mathfrak{A}$ be an $\aleph_0$-saturated structure. The least fixed point of φ on $\mathfrak{A}$ is reached after at most ω steps.*

Proof. Let F be the fixed-point induction of φ on $\mathfrak{A}$. If we can show that

$$\varphi(F(\omega))^{\mathfrak{A}} = \bigcup_{n<\omega} \varphi(F(n))^{\mathfrak{A}}$$

then it follows that

$$F(\omega+1) = \varphi(F(\omega))^{\mathfrak{A}} = \bigcup_{n<\omega} F(n+1) = F(\omega)\,,$$

as desired. Note that each set $F(n)$ with $n < \omega$ is definable by the n-th approximation of φ.

For the induction below we prove a slightly more general statement. We consider formulae $\varphi(R, \bar{x})$ where the relation R occurs only positively, but where we do not require the arity of R to be equal to the number of variables $\bar{x}$. With every such formula $\varphi(R, \bar{x})$ we associate the function

$$f_\varphi(R) := \{\, \bar{a} \subseteq A \mid \mathfrak{A} \vDash \varphi(R, \bar{a}) \,\}\,,$$

and we prove by induction on φ that

$$f_\varphi(\bigcup_{n<\omega} R_n) = \bigcup_{n<\omega} f_\varphi(R_n)\,,$$

for every increasing sequence $(R_n)_{n<\omega}$ of FO-definable relations.

W.l.o.g. we may assume that φ is in negation normal form. As φ is monotone in R we have

$$f_\varphi(R_n) \subseteq f_\varphi(\bigcup_n R_n)\,, \quad \text{for all } n < \omega\,.$$

This implies that

$$\bigcup_n f_\varphi(R_n) \subseteq f_\varphi(\bigcup_n R_n)\,.$$

Hence, we only need to prove the converse inclusion.

First, suppose that φ is atomic. If R does not occur in φ then there is nothing to do. Hence, assume that $\varphi = Rt_0 \dots t_{m-1}$. Then we have

$$\begin{aligned}
& \bar{a} \in f_\varphi(\bigcup_n R_n) \\
\Rightarrow\quad & \langle t_0(\bar{a}), \dots, t_{m-1}(\bar{a})\rangle \in \bigcup_n R_n \\
\Rightarrow\quad & \langle t_0(\bar{a}), \dots, t_{m-1}(\bar{a})\rangle \in R_n\,, \quad \text{for some } n < \omega \\
\Rightarrow\quad & \bar{a} \in \bigcup_n f_\varphi(R_n)\,.
\end{aligned}$$

If φ is the negation of an atom the proof is analogous.

For $\varphi = \psi \wedge \vartheta$ or $\varphi = \psi \vee \vartheta$ the claim follows immediately from inductive hypothesis.

Suppose that $\varphi = \exists y \psi(R, \bar{x}, y)$. Then we have

$$\begin{aligned}
& \bar{a} \in f_\varphi(\bigcup_n R_n) \\
\Rightarrow\quad & \bar{a}b \in f_\psi(\bigcup_n R_n)\,, \quad \text{for some } b \in A \\
\Rightarrow\quad & \bar{a}b \in \bigcup_n f_\psi(R_n)\,, \quad \text{for some } b \in A \\
\Rightarrow\quad & \bar{a} \in \bigcup_n f_\varphi(R_n)\,.
\end{aligned}$$

Finally, we consider the case that $\varphi = \forall y \psi(R, \bar{x}, y)$. For a contradiction, suppose that there is some tuple

$$\bar{a} \in f_\varphi(\bigcup_n R_n) \smallsetminus \bigcup_n f_\varphi(R_n)\,.$$

Since $\mathfrak{A} \nvDash \forall y \psi(R_n, \bar{a}, b)$ we can find elements $b_n \in A$ such that

$$\mathfrak{A} \nvDash \psi(R_n, \bar{a}, b_n)\,.$$

Let $\vartheta_n(\bar{z})$ be the formula defining R_n. We define

$$\Phi := \{\, \neg\psi(\vartheta_n, \bar{a}, y) \mid n < \omega \,\}\,,$$

where $\psi(\vartheta_n, \bar{x}, y)$ is the formula obtained from ψ by replacing every atom $R\bar{t}$ by $\vartheta_n(\bar{t})$. Since $R_k \subseteq R_n$, for $k \leq n$, we have $\vartheta_k \vDash \vartheta_n$. As ψ is monotone in R it follows that

$$\neg\psi(\vartheta_n, \bar{a}, y) \vDash \neg\psi(\vartheta_k, \bar{a}, y)\,, \quad \text{for all } k \leq n\,.$$

Therefore, every finite subset of Φ is satisfiable. Hence, Φ is a partial type over $\bar{a}$. Since $\mathfrak{A}$ is $\aleph_0$-saturated we can find some element $b_* \in A$ realising Φ. Consequently, we have

$$\bar{a}b_* \notin \bigcup_n f_\psi(R_n) = f_\psi(\bigcup_n R_n)\,.$$

Hence, $\mathfrak{A} \vDash \exists y \neg\psi(\bigcup_n R_n, \bar{a}, y)$ which implies that $\bar{a} \notin f_\varphi(\bigcup_n R_n)$. Contradiction. □

Theorem 6.18 (Barwise, Moschovakis). *Suppose that $\mathcal{K}$ is a pseudo-elementary class and $\varphi(R, \bar{x})$ a first-order formula. The following statements are equivalent:*

(1) *There exists a formula $\psi(\bar{x}) \in$ FO such that*

$$\mathfrak{A} \vDash \forall\bar{x}\big[\psi(\bar{x}) \leftrightarrow [\mathrm{lfp}\, R\bar{x} : \varphi](\bar{x})\big]\,, \quad \textit{for all } \mathfrak{A} \in \mathcal{K}\,.$$

(2) *For every $\mathfrak{A} \in \mathcal{K}$, there exists a formula $\psi(\bar{x}) \in$ FO such that*

$$\mathfrak{A} \vDash \forall\bar{x}\big[\psi(\bar{x}) \leftrightarrow [\mathrm{lfp}\, R\bar{x} : \varphi](\bar{x})\big]\,.$$

(3) *On every $\mathfrak{A} \in \mathcal{K}$ the least fixed-point of φ is reached after finitely many steps.*

(4) *There is a contant $n < \omega$ such that, on each $\mathfrak{A} \in \mathcal{K}$ the least fixed-point of φ is reached after at most n steps.*

Proof. Let $\mathcal{K}^+$ be a class such that $\mathcal{K} = \mathrm{pr}_\Sigma(\mathcal{K}^+)$ and fix a first-order theory T axiomatising $\mathcal{K}^+$. Let φ^n be the n-th approximation of φ.

(4) ⇒ (1) If, on every structure of $\mathcal{K}$, the fixed point is reached after at most n steps then we have

$$\mathfrak{A} \vDash \varphi^n(\bar{a}) \leftrightarrow [\mathrm{lfp}\, R\bar{x} : \varphi](\bar{a})\,, \quad \text{for all } \mathfrak{A} \in \mathcal{K} \text{ and all } \bar{a} \subseteq A\,.$$

Hence, we can set $\psi := \varphi^n$.

(1) ⇒ (2) is trivial.

(2) ⇒ (3) For a contradiction, suppose that on some structure $\mathfrak{A} \in \mathcal{K}^+$ the fixed point of φ is not reached in finitely many steps. Fix some $\aleph_0$-saturated elementary extension $\mathfrak{B} \succeq \mathfrak{A}$. Since

$$\mathfrak{A}|_\Sigma \vDash \exists \bar{x}[\varphi^{n+1}(\bar{x}) \wedge \neg\varphi^n(\bar{x})]\,, \quad \text{for all } n < \omega\,,$$

it follows that, on the structure $\mathfrak{B}$, the fixed point is also not reached in finitely many steps. By assumption there is a first-order formula $\psi(\bar{x})$ defining the fixed point on $\mathfrak{B}$. Hence,

$$\mathfrak{B}|_\Sigma \vDash \exists \bar{x}[\psi(\bar{x}) \wedge \neg\varphi^n(\bar{x})]\,, \quad \text{for all } n < \omega\,.$$

As $\mathfrak{B}$ is $\aleph_0$-saturated we can find some tuple $\bar{b} \subseteq B$ such that

$$\mathfrak{B}|_\Sigma \vDash \psi(\bar{b}) \wedge \bigwedge_{n<\omega} \neg\varphi^n(\bar{b})\,.$$

Note that $\mathfrak{B} \in \mathcal{K}^+$. Hence, $\psi^{\mathfrak{B}}$ is the fixed point of φ. Since the tuple $\bar{b}$ enters the fixed point at an infinite stage it follows that the fixed point is not reached in ω steps. (Note that no tuple enters the fixed point at stage ω.) This contradicts Lemma 6.17.

(3) ⇒ (4) For a contradiction, suppose that, for each $n < \omega$, there is a structure $\mathfrak{A}_n \in \mathcal{K}^+$ such that on $\mathfrak{A}_n$ the fixed-point of φ is reached after more than n steps. Setting

$$\vartheta_n := \exists \bar{x}[\varphi^{n+1}(\bar{x}) \wedge \neg\varphi^n(\bar{x})]$$

we have

$$\mathfrak{A}_n|_\Sigma \vDash \vartheta_n\,.$$

It follows that $T \nvDash \neg\vartheta_n$, for all $n < \omega$. Let $\Theta := \{\, \vartheta_n \mid n < \omega \,\}$. The theory $T \cup \Theta$ is consistent since, for every finite subset $\Theta_0 \subseteq \Theta$, we can find some n such that $\mathfrak{A}_n|_\Sigma \vDash T \cup \Theta_0$. Let $\mathfrak{B}$ be a model of $T \cup \Theta$. It follows that on $\mathfrak{B}$ the fixed-point of φ is not reached after finitely many steps. Contradiction. □

As an example of the expressive power of fixed-point logics we consider linear orders.

Lemma 6.19. *There exists a formula* $\varphi(x, y, z) \in \mathrm{FO}^3(\mathrm{LFP})[<]$ *such that, for every infinite cardinal* κ, φ *defines in the structure* $\langle \kappa, < \rangle$ *a bijection* $\kappa \times \kappa \to \kappa$.

Proof. We have shown in the proof of Theorem A4.3.8 that the formula

$$\begin{aligned}\psi(x_0 x_1, y_0 y_1) := {} & [(x_0 < y_0 \lor x_0 < y_1) \land (x_1 < y_0 \lor x_1 < y_1)] \\ & \lor [x_0 < y_0 \land x_1 = y_1 \land y_0 \leq y_1] \\ & \lor [x_0 < y_0 \land x_1 = y_0 \land y_1 \leq y_0] \\ & \lor [x_0 = y_0 \land x_1 < y_1 \land y_1 \leq y_0]\end{aligned}$$

defines a linear order on $\kappa \times \kappa$ of order type κ. The fixed-point formula

$$\begin{aligned}&\varphi(x, y, z) := \\ &\quad [\mathrm{lfp}\, R u_0 u_1 w : (\forall v_0 v_1 . \psi(\bar{v}, \bar{u}))(\exists w' . w' < w) R \bar{v} w' \\ &\qquad\qquad \land (\forall w' . w' < w)(\exists v_0 v_1 . \psi(\bar{v}, \bar{u})) R \bar{v} w'](x, y, z)\,.\end{aligned}$$

defines the corresponding bijection. □

Exercise 6.1. Let $\mathfrak{N} = \langle \mathbb{N}, < \rangle$. Construct FO(LFP)-formulae $\varphi_+(x, y, z)$ and $\varphi_\cdot(x, y, z)$ defining addition and multiplication on $\mathfrak{N}$.

To facilitate the investigation of model theoretic properties of fixed point logics we reduce them to a simpler logic, the extension of first-order logic by well-ordering quantifiers.

Lemma 6.20. $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{LFP}) =^1_{\mathrm{pc}} \mathrm{FO}_{\kappa\aleph_0}(\mathrm{wo})$.

Proof. We have seen in the example above that, for every $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{LFP})$-formula $\varphi(\bar{x}, \bar{y}, \bar{z})$, we can construct a formula $\psi(\bar{z}) \in \mathrm{FO}_{\kappa\aleph_0}(\mathrm{LFP})$ expressing that φ defines a well-order. Hence, $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{wo}) \leq \mathrm{FO}_{\kappa\aleph_0}(\mathrm{LFP})$.

For the converse, let $\varphi \in \mathrm{FO}_{\kappa\aleph_0}(\mathrm{LFP})$. For every subformula ψ of φ, we introduce a new relation symbol R_ψ and we construct a set of sentences Φ such that

$$R_\psi^{\mathfrak{A}} = \psi^{\mathfrak{A}}, \qquad \text{for every model } \mathfrak{A} \vDash \Phi .$$

The construction of Φ proceeds by induction on φ. For atomic subformulae ψ, we add the formula

$$\forall \bar{x}(R_\psi \bar{x} \leftrightarrow \psi(\bar{x}))$$

to Φ. For the inductive step we use the same formulae as in the proof of Chang's Reduction (Lemma c1.4.12), e.g., for conjuctions we use

$$\forall \bar{x}(R_{\bigwedge \Psi} \bar{x} \leftrightarrow \bigwedge_{\psi \in \Psi} R_\psi \bar{x}) .$$

The only nontrivial case is the case that $\psi = [\mathrm{lfp}\, P\bar{x} : \vartheta](\bar{x})$ is a fixed-point formula.

Let $<$ be a new binary relation symbol, s a new unary function symbol, and 0 a new constant symbol. We add to Φ the sentences

$$\begin{aligned}
&\forall u(u = 0 \vee 0 < u) \\
&\forall u(u < su \wedge \neg \exists v(u < v \wedge v < su)) \\
&\mathsf{W} uv(u < v)
\end{aligned}$$

which express that $<$ is a well-order of the universe, s is the successor function, and 0 is the minimal element. Furthermore, we add the formulae

$$\begin{aligned}
&\forall \bar{x}(\neg S_\varphi 0 \bar{x}) \\
&\forall u \forall \bar{x}\big(S_\varphi s u \bar{x} \leftrightarrow \chi[P/S_\varphi u]\big) , \\
&\forall u \forall \bar{x}\big[\forall v(sv \neq u) \rightarrow (S_\varphi u \bar{x} \leftrightarrow \exists v(v < u \wedge S_\varphi v \bar{x}))\big] , \\
&\forall \bar{x}(R_\varphi \bar{x} \leftrightarrow \exists u S_\varphi u \bar{x}) ,
\end{aligned}$$

which express that $S_\varphi = \{ \langle \alpha, \bar{a} \rangle \mid \bar{a} \in F(\alpha) \}$. Finally, we need the formula

$$\exists u \forall \bar{x} (S_\varphi s u \bar{x} \leftrightarrow S_\varphi u \bar{x}) ,$$

which expresses that the fixed point is actually reached. For the correctness of this construction note that the closure ordinal α of every $\mathrm{FO}_{\infty\aleph_0}(\mathrm{LFP})$-induction on a structure $\mathfrak{A}$ is less than $|A|^+$. Hence, we can really choose an ordering $<$ of A of order type α. □

For partial fixed points we have an analogous result where the projective reduction is replaced by a relativised reduct.

Lemma 6.21. $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{PFP}) =^1_{\mathrm{rpc}} \mathrm{FO}_{\kappa\aleph_0}(\mathrm{wo})$.

Proof. We can basically use the same construction as in the proof of Lemma 6.20. The only difference is that the closure ordinal for partial fixed points is not bounded by the size of the structure. Therefore, we cannot choose a sufficiently long well-ordering of the universe. Instead, we add a new sort w to the given structure $\mathfrak{A}$ and we choose the domain A_w large enough to contain a well-ordering $<$ of length $(2^{|A|})^+$. After performing the same construction as above in the larger structure we can take a relativised reduct to obtain the original structure $\mathfrak{A}$. □

Using this reduction we can use the Löwenheim-Skolem theorem for $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{wo})$ to derive a corresponding theorem for $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{PFP})$.

Theorem 6.22. *Let $\Delta \subseteq \mathrm{FO}^{<\omega}_{\kappa\aleph_0}(\mathrm{PFP})[\Sigma]$, for a regular cardinal κ, and set $\mu := |\Sigma| \oplus |\Delta| \oplus \kappa^-$ where $\kappa^- := \sup \{ \lambda \mid \lambda < \kappa \}$.*

For each Σ-structure $\mathfrak{A}$, every subset $X \subseteq A$, and all cardinals λ with $|X| \oplus \mu \leq \lambda \leq |A|$, there exists a Δ-substructure $\mathfrak{B} \preceq_\Delta \mathfrak{A}$ of size $|B| = \lambda$ with $X \subseteq B$.

Proof. This follows immediately by Theorem c2.3.10 and Lemma 6.21. □

We conclude this section with a proof that the logics $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{LFP})$ and $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{IFP})$ have the same expressive power.

Definition 6.23. Let $\mathfrak{A}$ be a structure, $\varphi(R, \bar{x})$ an $\mathrm{FO}_{\infty\aleph_0}(\mathrm{IFP})$-formula, and F the fixed-point induction of $[\mathrm{ifp}\, R\bar{x} : \varphi]$.

(a) The *inductive fixed-point rank* $\mathrm{rk}_\varphi(\bar{a})$ of a tuple $\bar{a} \in [\mathrm{ifp}\, R\bar{x} : \varphi]^{\mathfrak{A}}$ is the ordinal α such that $\bar{a} \in F(\alpha + 1) \setminus F(\alpha)$. For $\bar{a} \notin [\mathrm{ifp}\, R\bar{x} : \varphi]^{\mathfrak{A}}$, we set $\mathrm{rk}_\varphi(\bar{a}) := \infty$.

(b) The *stage comparison relation* $\triangleleft_\varphi$ of φ is defined by

$$\bar{a} \triangleleft_\varphi \bar{b} \quad :\text{iff} \quad \mathrm{rk}_\varphi(\bar{a}) < \mathrm{rk}_\varphi(\bar{b}) .$$

Lemma 6.24. *Let $\varphi(P, \bar{x})$ be an $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{IFP})$-formula. The stage comparison relation $\triangleleft_\varphi$ for $[\mathrm{ifp}\, P\bar{x} : \varphi]$ is $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{IFP})$-definable.*

Proof. Let $\hat{\varphi}(\bar{x}, \bar{z})$ be the formula obtained from $P\bar{x} \vee \varphi(P, \bar{x})$ by replacing every atom of the form $P\bar{t}$ by the formula $R\bar{t}\bar{z}$. We claim that $\triangleleft_\varphi$ is defined by the formula where

$$[\mathrm{ifp}\, R\bar{x}\bar{y} : \hat{\varphi}(\bar{x}, \bar{x}) \wedge \neg\hat{\varphi}(\bar{y}, \bar{x})](\bar{x}\bar{y}) .$$

Let $(R^\alpha)_\alpha$ be the fixed-point induction of this formula. A straightforward induction on α shows that

$$\langle \bar{a}, \bar{b} \rangle \in R^\alpha \quad \text{iff} \quad \bar{a} \triangleleft_\varphi \bar{b} \text{ and } \mathrm{rk}_\varphi(\bar{a}) < \alpha .$$

Hence, the result follows. □

Proposition 6.25. *Let $\varphi(P, \bar{x})$ be an $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{LFP})$-formula. The stage comparison relation $\triangleleft_\varphi$ for $[\mathrm{ifp}\, P\bar{x} : \varphi]$ is $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{LFP})$-definable.*

Proof. By $\varphi[P\bar{z}/\psi(\bar{z})/\vartheta(\bar{z})]$ we denote the formula obtained from the formula $P\bar{x} \vee \varphi(P, \bar{x})$ by replacing every atom of the form $P\bar{t}$ by

- $\psi(\bar{t})$, if this atom occurs positively in φ,
- $\vartheta(\bar{t})$, if it occurs negatively in φ.

As in the proof of the preceding lemma we would like to compute $\triangleleft_\varphi$ by the formula

$$\big[\mathrm{lfp}\, R\bar{x}\bar{y} : \varphi[P\bar{z}/R\bar{z}\bar{x}/R\bar{z}\bar{x}](\bar{x}) \wedge \neg\varphi[P\bar{z}/R\bar{z}\bar{x}/R\bar{z}\bar{x}](\bar{y})\big](\bar{x}\bar{y}) .$$

Unfortunately, this does not work since we can use R only positively in φ and only negatively in $\neg\varphi$. Instead, we construct another formula ψ computing $\triangleleft_\varphi$ that we can substitute for R at those places where we cannot use it. Again the obvious definition

$$\begin{aligned}\psi(\bar{x},\bar{y}) \coloneqq \big[\mathrm{lfp}\, S\bar{x}\bar{y} : \varphi[P\bar{z}/S\bar{z}\bar{x}/S\bar{z}\bar{x}](\bar{x})& \\ \wedge \neg\varphi[P\bar{z}/S\bar{z}\bar{x}/S\bar{z}\bar{x}](\bar{y})\big](\bar{x}\bar{y})&\end{aligned}$$

does not work. But, since ψ is used in the above formula at those places where R occurs negatively we can use R inside of ψ provided its occurrence is also negative. These considerations lead to following attempt to define $\triangleleft_\varphi$:

$$\begin{aligned}&\big[\mathrm{lfp}\, R\bar{x}\bar{y} : \varphi[P\bar{z}/R\bar{z}\bar{x}/\psi(\bar{z},\bar{x})](\bar{x}) \wedge \\ &\qquad \wedge \neg\varphi[P\bar{z}/\psi(\bar{z},\bar{x})/R\bar{z}\bar{x}](\bar{y})\big](\bar{x}\bar{y}) ,\end{aligned}$$

where $\psi(\bar{x},\bar{y})$ is the formula

$$\big[\mathrm{lfp}\, S\bar{x}\bar{y} : \varphi[P\bar{z}/S\bar{z}\bar{x}/R\bar{z}\bar{x}](\bar{x}) \wedge \neg\varphi[P\bar{z}/R\bar{z}\bar{x}/S\bar{z}\bar{x}](\bar{y})\big](\bar{x}\bar{y}) .$$

This definition is still not correct but we can repair it as follows. We claim that $\triangleleft_\varphi$ is defined by the formula $[\mathrm{lfp}\, R\bar{x}\bar{y} : \chi](\bar{x}\bar{y})$ where

$$\begin{aligned}\chi(\bar{x},\bar{y}) &\coloneqq \varphi[P\bar{z}/R\bar{z}\bar{x}/\psi(\bar{z},\bar{x})](\bar{x}) \\ &\qquad \wedge \neg\varphi[P\bar{z}/\psi(\bar{z},\bar{x})/R\bar{z}\bar{x}](\bar{y}) \\ &\qquad \wedge \forall\bar{z}(\psi(\bar{z},\bar{x}) \to R\bar{z}\bar{x}) , \\ \psi(\bar{x},\bar{y}) &\coloneqq [\mathrm{lfp}\, S\bar{x}\bar{y} : \vartheta](\bar{x}\bar{y}) , \\ \vartheta(\bar{x},\bar{y}) &\coloneqq \varphi[P\bar{z}/S\bar{z}\bar{x}/R\bar{z}\bar{x}](\bar{x}) \\ &\qquad \wedge \neg\varphi[P\bar{z}/R\bar{z}\bar{x}/S\bar{z}\bar{x} \wedge S\bar{z}\bar{y}](\bar{y}) \\ &\qquad \wedge (\forall\bar{z}.R\bar{z}\bar{x})(S\bar{z}\bar{x} \wedge S\bar{z}\bar{y}) .\end{aligned}$$

First, note that $[\mathrm{lfp}\, R\bar{x}\bar{y} : \chi] \in \mathrm{FO(LFP)}$ since S occurs only positively in ϑ and R occurs only negatively in ψ. Let $(P^\alpha)_\alpha$ be the fixed-point induction of $[\mathrm{ifp}\, P\bar{x} : \varphi]$. For $\alpha \in \mathrm{On}$, define

$$R^\alpha := \{ \langle \bar{a}, \bar{b} \rangle \mid \bar{a} \lhd_\varphi \bar{b} \text{ and } \mathrm{rk}_\varphi(\bar{a}) < \alpha \} .$$

We will show that the sequence $(R^\alpha)_\alpha$ is the fixed-point induction of χ.

Claim. *Let $(S^\beta)_\beta$ be the fixed-point induction of ψ where R is interpreted by R^α, and set $S^\infty := \bigcup_\beta S^\beta$.*

(a) *For $\beta \leq \alpha$, we have*

$$\langle \bar{a}, \bar{b} \rangle \in S^\beta \quad \text{iff} \quad \bar{a} \lhd_\varphi \bar{b} \text{ and } \mathrm{rk}_\varphi(\bar{a}) < \beta .$$

(b) *For all tuples $\bar{b}$ with rank $\mathrm{rk}_\varphi(\bar{b}) > \alpha$, there exist a tuple $\bar{a}$ with $\mathrm{rk}_\varphi(\bar{a}) = \alpha$ such that $\langle \bar{a}, \bar{b} \rangle \in S^\infty$.*

(c) *If $R^{\alpha+1} = R^\alpha$ then $S^\infty = R^\alpha$.*

(d) *If $\mathrm{rk}_\varphi(\bar{a}) < \alpha$ or $\mathrm{rk}_\varphi(\bar{b}) < \alpha$ then we have*

$$\langle \bar{a}, \bar{b} \rangle \in S^\infty \quad \text{iff} \quad \bar{a} \lhd_\varphi \bar{b} .$$

(a) We prove the claim by induction on β. The case that $\beta = 0$ is trivial and the limit step follows immediately from the inductive hypothesis. For the successor step, note that

$$\langle \bar{a}, \bar{b} \rangle \in S^{\beta+1} \quad \text{iff} \quad \langle \mathfrak{A}, R^\alpha, S^\beta \rangle \vDash \vartheta(\bar{a}, \bar{b}) .$$

First, suppose that $\gamma := \mathrm{rk}_\varphi(\bar{a}) < \beta$. By inductive hypothesis, it follows that

$$\langle \bar{c}, \bar{a} \rangle \in S^\beta \quad \text{iff} \quad \bar{c} \lhd_\varphi \bar{a} \quad \text{iff} \quad \bar{c} \in P^\gamma .$$

Since $\beta \leq \alpha$ we further have that

$$\langle \bar{c}, \bar{a} \rangle \in R^\alpha \quad \text{iff} \quad \bar{c} \lhd_\varphi \bar{a} \quad \text{iff} \quad \bar{c} \in P^\gamma .$$

Consequently, $\mathfrak{A} \vDash \varphi(P^\gamma, \bar{a})$ implies that

$$\langle \mathfrak{A}, R^\alpha, S^\beta \rangle \vDash \varphi[P\bar{z}/S\bar{z}\bar{a}/R\bar{z}\bar{a}](\bar{a}) .$$

If $\mathrm{rk}_\varphi(\bar{a}) < \mathrm{rk}_\varphi(\bar{b})$ then, by inductive hypothesis, $\langle \bar{c}, \bar{a} \rangle \in S$ implies $\langle \bar{c}, \bar{b} \rangle \in S$. Since $\beta \leq \alpha$ it follows that we have $\langle \bar{c}, \bar{a} \rangle \in S^\beta$ iff $\langle \bar{c}, \bar{a} \rangle \in R^\alpha$. Consequently, there is no tuple $\bar{c}$ such that

$$\langle \mathfrak{A}, R^\alpha, S^\beta \rangle \vDash R\bar{c}\bar{a} \wedge \neg(S\bar{c}\bar{a} \wedge S\bar{c}\bar{b}) .$$

Finally, we have

$$\langle \mathfrak{A}, R^\alpha, S^\beta \rangle \nvDash \varphi[P\bar{z}/R\bar{z}\bar{a}/S\bar{z}\bar{a} \wedge S\bar{z}\bar{b}](\bar{b})$$

since, otherwise, $\mathrm{rk}_\varphi(\bar{b}) \leq \mathrm{rk}_\varphi(\bar{a})$. It follows that $\langle \bar{a}, \bar{b} \rangle \in S^{\beta+1}$.

Next, consider the case that $\mathrm{rk}_\varphi(\bar{a}) > \mathrm{rk}_\varphi(\bar{b})$. By inductive hypothesis, $S\bar{c}\bar{a} \wedge S\bar{c}\bar{b}$ is equivalent to $S\bar{c}\bar{b}$. Consequently, choosing $\bar{c} := \bar{b}$ we can find a tuple $\bar{c}$ with

$$\langle \mathfrak{A}, R^\alpha, S^\beta \rangle \vDash R\bar{c}\bar{a} \wedge \neg(S\bar{c}\bar{a} \wedge S\bar{c}\bar{b}) .$$

Hence, $\langle \bar{a}, \bar{b} \rangle \notin S^{\beta+1}$.

Finally, suppose that $\mathrm{rk}_\varphi(\bar{a}) = \mathrm{rk}_\varphi(\bar{b})$. Then

$$\langle \mathfrak{A}, R^\alpha, S^\beta \rangle \vDash \varphi[P\bar{z}/R\bar{z}\bar{a}/S\bar{z}\bar{a} \wedge S\bar{z}\bar{b}](\bar{b})$$

and $\langle \bar{a}, \bar{b} \rangle \notin S^{\beta+1}$.

It remains to consider the case that $\mathrm{rk}_\varphi(\bar{a}) \geq \beta$. By inductive hypothesis, we have $\langle \bar{c}, \bar{a} \rangle \in S^\beta$ iff $\mathrm{rk}(\bar{c}) < \beta$. Since $\mathfrak{A} \nvDash \varphi(P^\beta, \bar{a})$ it follows that

$$\langle \mathfrak{A}, S^\beta \rangle \nvDash \varphi[P\bar{z}, S\bar{z}\bar{a}, S\bar{z}\bar{a}](\bar{a}) .$$

Note that

$$\langle \bar{c}, \bar{a} \rangle \in S^\beta \quad \text{implies} \quad \langle \bar{c}, \bar{a} \rangle \in R^\alpha .$$

Since P'' occurs only negatively in $\varphi[P\bar{z}/P'\bar{z}/P''\bar{z}]$ it therefore follows that

$$\langle \mathfrak{A}, R^\alpha, S^\beta \rangle \not\models \varphi[P\bar{z}/S\bar{z}\bar{a}/R\bar{z}\bar{a}](\bar{a}) .$$

(b) By (a), we have $S^\alpha = R^\alpha$. Let $\mathrm{rk}_\varphi(\bar{a}) \leq \alpha$. Then $\bar{a} \in P^{\alpha+1}$ implies that

$$\langle \mathfrak{A}, R^\alpha, S^\alpha \rangle \models \varphi[P\bar{z}/S\bar{z}\bar{a}/R\bar{z}\bar{a}](\bar{a}) .$$

If $\mathrm{rk}_\varphi(\bar{b}) \geq \mathrm{rk}_\varphi(\bar{a})$ then $S\bar{c}\bar{a} \wedge S\bar{c}\bar{b}$ is equivalent to $S\bar{c}\bar{a}$ and, hence, to $R\bar{c}\bar{a}$. Consequently, it follows in this case that

$$\langle \mathfrak{A}, R^\alpha, S^\alpha \rangle \models \neg\varphi[P\bar{z}, R\bar{z}\bar{b}, S\bar{z}\bar{a} \wedge S\bar{z}\bar{b}](\bar{b})$$
$$\text{iff} \quad \mathrm{rk}_\varphi(\bar{b}) > \mathrm{rk}_\varphi(\bar{a}) .$$

If, on the other hand, $\mathrm{rk}_\varphi(\bar{b}) < \mathrm{rk}_\varphi(\bar{a})$ then $\mathrm{rk}_\varphi(\bar{b}) < \alpha$ and setting $\bar{c} := \bar{b}$ we obtain a tuple such that

$$\langle \mathfrak{A}, R^\alpha, S^\alpha \rangle \models R\bar{c}\bar{a} \wedge \neg(S\bar{c}\bar{a} \wedge S\bar{c}\bar{b}) .$$

Consequently, $\langle \mathfrak{A}, R^\alpha, S^\alpha \rangle \not\models \vartheta(\bar{a}, \bar{b})$.

Finally, suppose that $\mathrm{rk}_\varphi(\bar{a}) > \alpha$. Then

$$\langle \mathfrak{A}, R^\alpha, S^\alpha \rangle \not\models \varphi[P\bar{z}/S\bar{z}\bar{a}/R\bar{z}\bar{a}](\bar{a})$$

since $\langle \bar{c}, \bar{a} \rangle \in S^\alpha$ iff $\mathrm{rk}_\varphi(\bar{c}) < \alpha$ and $\langle \bar{c}, \bar{a} \rangle \in R^\alpha$ iff $\mathrm{rk}_\varphi(\bar{c}) < \alpha$.

It follows that $S^{\alpha+1}$ contains all pairs $\langle \bar{a}, \bar{b} \rangle$ with $\mathrm{rk}_\varphi(\bar{a}) \leq \alpha$ and $\mathrm{rk}_\varphi(\bar{a}) < \mathrm{rk}_\varphi(\bar{b})$. If $\mathrm{rk}_\varphi(\bar{b}) > \alpha$ then there exists some tuple $\bar{a}$ with $\mathrm{rk}_\varphi(\bar{a}) = \alpha$ and it follows that $\langle \bar{a}, \bar{b} \rangle \in S^{\alpha+1}$.

(c) If $R^{\alpha+1} = R^\alpha$ then there are no tuples $\bar{a}$ with $\mathrm{rk}_\varphi(\bar{a}) = \alpha$. By (b) it follows that $S^{\alpha+1} = S^\alpha$. Consequently, $S^\alpha = S^\infty$.

(d) If $R^{\alpha+1} = R^\alpha$ then the claim follows from (c). Hence, we may assume that $R^\alpha \subset R^{\alpha+1}$. We show that, for every $\gamma \geq \alpha$, if $\mathrm{rk}_\varphi(\bar{a}), \mathrm{rk}_\varphi(\bar{b}) < \alpha$ then $\langle \bar{a}, \bar{b} \rangle \in S^{\gamma+1}$ implies $\langle \bar{a}, \bar{b} \rangle \in S^{\alpha+1}$. Suppose otherwise and let

γ be the minimal ordinal such that there exists a counterexample. Then we obtain a contradiction as in the proof of (a).

It remains to show that there are no tuples with $\mathrm{rk}_\varphi(\bar{a}) \geq \alpha$ and $\mathrm{rk}_\varphi(\bar{b}) < \alpha$ such that $\langle \bar{a}, \bar{b} \rangle \in S^{\beta+1}$, for some $\beta \geq \alpha$. Suppose otherwise and let β be the minimal ordinal such that there exists a counterexample $\langle \bar{a}, \bar{b} \rangle \in S^{\beta+1}$. Then

$$\langle \mathfrak{A}, R^\alpha, S^\beta \rangle \vDash \vartheta(\bar{a}, \bar{b}) .$$

Since $\mathrm{rk}_\varphi(\bar{a}) \geq \alpha$, we have $\langle \bar{c}, \bar{a} \rangle \in R^\alpha$ iff $\mathrm{rk}_\varphi(\bar{c}) < \alpha$. By minimality of β, it follows that we have

$$\langle \mathfrak{A}, R^\alpha, S^\beta \rangle \vDash S\bar{c}\bar{a} \wedge S\bar{c}\bar{b} \quad \text{iff} \quad \mathrm{rk}_\varphi(\bar{c}) < \mathrm{rk}_\varphi(\bar{b}) .$$

If $\mathrm{rk}_\varphi(\bar{b}) < \alpha$ then setting $\bar{c} := \bar{b}$ we obtain a tuple $\bar{c}$ such that

$$\langle \mathfrak{A}, R^\alpha, S^\beta \rangle \vDash R\bar{c}\bar{a} \wedge \neg(S\bar{c}\bar{a} \wedge S\bar{c}\bar{b}) .$$

Consequently, $\langle \mathfrak{A}, R^\alpha, S^\beta \rangle \nvDash \vartheta(\bar{a}, \bar{b})$. Contradiction.

Similarly, if $\mathrm{rk}_\varphi(\bar{b}) = \alpha$ then

$$\langle \mathfrak{A}, R^\alpha, S^\beta \rangle \vDash \varphi[P\bar{z}/R\bar{z}\bar{a}/S\bar{z}\bar{a} \wedge S\bar{z}\bar{b}](\bar{b}) ,$$

and again $\langle \mathfrak{A}, R^\alpha, S^\beta \rangle \nvDash \vartheta(\bar{a}, \bar{b})$. This contradiction concludes the proof of the claim.

To finish the proof of the lemma we still have to show that $(R^\alpha)_\alpha$ is the fixed-point induction of χ. We prove this statement by induction on α. For $\alpha = 0$ and for limit ordinals the proof is trivial. For the successor step we show that

$$\langle \mathfrak{A}, R^\alpha \rangle \vDash \chi(\bar{a}, \bar{b}) \quad \text{iff} \quad \bar{a} \triangleleft_\varphi \bar{b} \text{ and } \mathrm{rk}_\varphi(\bar{a}) \leq \alpha .$$

First, we consider the case that $\mathrm{rk}_\varphi(\bar{a}) \leq \alpha$. Then we have $\langle \bar{c}, \bar{a} \rangle \in R^\alpha$ iff $\bar{c} \triangleleft_\varphi \bar{a}$. By statement (d) above, it follows that $\langle \bar{c}, \bar{a} \rangle \in R^\alpha$ iff $\psi(\bar{c}, \bar{a})$. Consequently, we have

$$\langle \mathfrak{A}, R^\alpha \rangle \vDash \varphi[P\bar{z}/R\bar{z}\bar{a}/\psi(\bar{z}, \bar{a})](\bar{a})$$

and every tuple $\bar{c}$ satisfies $\psi(\bar{c}, \bar{a}) \rightarrow R^\alpha \bar{c}\bar{a}$. Since

$$\langle \mathfrak{A}, R^\alpha \rangle \vDash \neg\varphi[P\bar{z}/\psi(\bar{z}, \bar{a})/R\bar{z}\bar{a}](\bar{b})$$

it follows that

$$\langle \mathfrak{A}, R^\alpha \rangle \vDash \chi(\bar{a}, \bar{b}) \quad \text{iff} \quad \bar{a} \vartriangleleft_\varphi \bar{b}\,.$$

It remains to consider the case that $\mathrm{rk}_\varphi(\bar{a}) > \alpha$. Then we have $\langle \bar{c}, \bar{a} \rangle \in R^\alpha$ iff $\mathrm{rk}_\varphi(\bar{c}) < \alpha$. If $P^\alpha = P^{\alpha+1}$ then, by statement (c) above, it follows that $\psi(\bar{x}, \bar{y})$ defines R^α and

$$\langle \mathfrak{A}, R^\alpha \rangle \nvDash \varphi[P\bar{z}/R\bar{z}\bar{a}/\psi(\bar{z}, \bar{a})](\bar{a})\,.$$

Hence, $\langle \mathfrak{A}, R^\alpha \rangle \nvDash \chi(\bar{a}, \bar{b})$.

If, on the other hand, $P^\alpha \subset P^{\alpha+1}$ then, by (b), there is a tuple $\bar{c} \vartriangleleft_\varphi \bar{a}$ with $\mathrm{rk}_\varphi(\bar{c}) = \alpha$. Consequently,

$$\langle \mathfrak{A}, R^\alpha \rangle \nvDash \forall \bar{z}(\psi(\bar{z}, \bar{a}) \rightarrow R\bar{z}\bar{a})\,. \qquad \square$$

Theorem 6.26 (Gurevich, Kreutzer, Shelah). $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{LFP}) = \mathrm{FO}_{\kappa\aleph_0}(\mathrm{IFP})$.

Proof. Let $[\mathrm{ifp}\, R\bar{x} : \varphi]$ be an $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{IFP})$-formula. By induction we may assume that $\varphi \in \mathrm{FO}_{\kappa\aleph_0}(\mathrm{LFP})$. By Proposition 6.25, there is an $\mathrm{FO}_{\kappa\aleph_0}(\mathrm{LFP})$-formula defining the stage comparison relation $\vartriangleleft_\varphi$. Note that we have

$$\mathrm{ifp}\, f = f(\mathrm{dom} \vartriangleleft_f) \cup \mathrm{dom} \vartriangleleft_f\,, \quad \text{for every function } f\,.$$

Hence, it follows that

$$[\mathrm{ifp}\, R\bar{x} : \varphi](\bar{x}) \equiv \varphi[R\bar{z}/\exists \bar{y}(\bar{z} \vartriangleleft_\varphi \bar{y})](\bar{x})\,,$$

where $\varphi[R\bar{z}/\vartheta(\bar{z})]$ denotes the formula obtained from $R\bar{x} \vee \varphi(R, \bar{x})$ by replacing every atom of the form $R\bar{t}$ by the formula $\vartheta(\bar{t})$. $\square$

Part D.

Axiomatisation and Definability

D1. *Quantifier elimination*

1. *Preservation theorems*

In Section C2.1 we have seen that several fragments of first-order logic are preserved under various operations. In this section we will show the converse. A preservation theorem is a result that characterises a semantic property of a formula by a syntactic condition. The general form of such a theorem is the statement:

> Let $\varphi \in L_+$. The class $\mathrm{Mod}_{L_+}(\varphi)$ has the property P if and only if there exists a formula $\psi \in L_-$ such that $\varphi \equiv \psi$.

Here P is an arbitrary property and L_+ and L_- are logics where usually we have $L_- \subset L_+$.

We will mostly be interested in closure properties. We consider a relation $\sqsubseteq$ with the property that $\mathrm{Mod}_{L_-}(\psi)$ is closed under $\sqsubseteq$, for every L_--formula ψ, i.e.,

$$\mathfrak{A} \vDash \psi \text{ and } \mathfrak{A} \sqsubseteq \mathfrak{B} \quad \text{implies} \quad \mathfrak{B} \vDash \psi\,.$$

Further, we assume that φ is an L_+-formula such that $\mathrm{Mod}_{L_+}(\varphi)$ is closed under $\sqsubseteq$. We want to find a formula $\psi \in L_-$ with $\psi \equiv \varphi$.

One way to prove that such a formula exists is the following. For a contradiction, we suppose that we can find structures $\mathfrak{A} \equiv_{L_-} \mathfrak{B}$ such that $\mathfrak{A} \vDash \varphi$ and $\mathfrak{B} \nvDash \varphi$. Given $\mathfrak{A}$ and $\mathfrak{B}$ we construct a structure $\mathfrak{C}$ such that

$$\mathfrak{A} \sqsubseteq \mathfrak{C} \quad \text{and} \quad \mathfrak{B} \equiv_{L_+} \mathfrak{C}\,.$$

This leads to a contradiction since, on the one hand, $\mathfrak{A} \vDash \varphi$ and $\mathfrak{A} \sqsubseteq \mathfrak{C}$ implies that $\mathfrak{C} \vDash \varphi$. But, on the other hand, $\mathfrak{B} \nvDash \varphi$ and $\mathfrak{B} \equiv_{L_+} \mathfrak{C}$ implies that $\mathfrak{C} \nvDash \varphi$.

Lemma 1.1. *Let T be a first-order theory and $\mathfrak{A}$ a structure.*

$$\mathfrak{A} \vDash T_{\forall}^{\vDash} \quad \text{iff} \quad \textit{there exists an embedding } \mathfrak{A} \to \mathfrak{B} \textit{ into some model } \mathfrak{B} \vDash T\,.$$

Proof. ($\Leftarrow$) Let $\mathfrak{A} \to \mathfrak{B}$ be an embedding and $\mathfrak{B} \vDash T$. Replacing $\mathfrak{A}$ by an isomorphic copy we may assume that $\mathfrak{A} \subseteq \mathfrak{B}$. Let $\varphi \in T_{\forall}^{\vDash}$. Since $T \vDash T_{\forall}^{\vDash}$ we have $\mathfrak{B} \vDash \varphi$. By Lemma C2.1.6, it follows that $\mathfrak{A} \vDash \varphi$.

($\Rightarrow$) Note that every function preserving $\exists$-formulae is an embedding. Therefore, this direction follows from Corollary C2.5.6 if we set $\Delta :=$ $\forall$. ☐

Theorem 1.2 (Łoś, Tarski). *For a first-order theory T and a set Φ of sentences, the following statements are equivalent:*

(1) $\mathfrak{B} \vDash \Phi$ *implies* $\mathfrak{A} \vDash \Phi$*, for all models* $\mathfrak{A} \subseteq \mathfrak{B}$ *of* T*.*

(2) Φ *is equivalent modulo T to a set of first-order $\forall$-formulae.*

Proof. The implication (2) $\Rightarrow$ (1) was proved in Lemma C2.1.6. For the other direction, we claim that $\Psi := (T \cup \Phi)_{\forall}^{\vDash}$ is equivalent to Φ. Clearly, if $\mathfrak{A} \vDash \Phi$ and $\mathfrak{A} \vDash T$ then $\mathfrak{A} \vDash \Psi$. On the other hand, by Lemma 1.1, we have

$$\mathfrak{A} \vDash \Psi \quad \text{iff} \quad \mathfrak{A} \subseteq \mathfrak{B} \text{ for some } \mathfrak{B} \vDash T \cup \Phi\,.$$

By (1), it follows that $\mathfrak{A} \vDash \Psi$ implies $\mathfrak{A} \vDash \Phi$. Therefore, $\Phi \equiv \Psi$ modulo T. ☐

Dualising the statement of the Theorem of Łoś and Tarski we obtain a characterisation of formulae preserved by embeddings.

Corollary 1.3. *Let T be a first-order theory. A formula $\varphi \in$ FO is preserved by embeddings between models of T if and only if φ is equivalent modulo T to an $\exists$-formula.*

Proof. Since $\neg\varphi$ is preserved in substructures it follows by Theorem 1.2 that we can find a set Φ of $\forall$-formulae with $\Phi \equiv \neg\varphi$. By the Compactness Theorem, there exists a finite subset $\Phi_0 \subseteq \Phi$ such that $\Phi_0 \models \neg\varphi$. Hence, $\neg\varphi \equiv \bigwedge \Phi_0$ and $\varphi \equiv \neg \bigwedge \Phi_0$. The latter is equivalent to an $\exists$-formula. □

We can extend the Theorem of Łoś and Tarski to pseudo-elementary classes.

Theorem 1.4. *If a class $\mathcal{K} \in \mathrm{RPC}_\infty(\mathrm{FO}, \Sigma)$ is closed under substructures then $\mathcal{K}$ is $\forall[\Sigma]$-axiomatisable.*

Proof. By Theorem C5.4.14, there exists a set $\Phi \subseteq \mathrm{FO}[\Gamma]$ such that

$$\mathcal{K} = \mathrm{pr}_\Sigma(\mathrm{Mod}(\Phi)) .$$

Let $T := \Phi^{\models}_{\forall} \cap \mathrm{FO}[\Sigma]$. Clearly, $\mathcal{K} \subseteq \mathrm{Mod}(T)$. It remains to prove the converse. Suppose that $\mathfrak{A} \models T$. Let $\Delta := \mathrm{QF}^{<\omega}[\Sigma]$ and set

$$\Psi := \mathrm{Th}_\Delta(\mathfrak{A}_A) \cup \Phi .$$

We show that Ψ is satisfiable. Suppose otherwise. Then there is some quantifier-free formula $\psi(\bar a)$ with parameters $\bar a \subseteq A$ such that

$$\mathfrak{A} \models \psi(\bar a) \quad \text{and} \quad \Phi \models \neg\psi(\bar a) .$$

Consequently, $\Phi \models \forall \bar x \neg\psi(\bar x)$. Since this sentence is in T it follows that $\mathfrak{A} \models \forall \bar x \neg\psi(\bar x)$. Contradiction.

Let $\mathfrak{B}$ be a model of Ψ. Since $\mathfrak{B} \models \mathrm{Th}_\Delta(\mathfrak{A}_A)$ there exists an embedding $\mathfrak{A} \to \mathfrak{B}$. Furthermore, we have $\mathfrak{B} \in \mathcal{K}$. As $\mathcal{K}$ is closed under substructures and isomorphisms, it follows that $\mathfrak{A} \in \mathcal{K}$. □

Example. As an application we consider representable groups. Let $0 < n < \omega$. We say that a group $\mathfrak{G}$ has a *faithful n-linear representation* if it can be embedded into $\mathrm{GL}_n(\mathfrak{K})$, the group of all invertible $n \times n$ matrices over some field $\mathfrak{K}$.

Claim. *A group $\mathfrak{G}$ has a faithful n-linear representation if and only if every finitely generated subgroup of $\mathfrak{G}$ has such a representation.*

($\Rightarrow$) Clearly, if $\mathfrak{G}$ can be embedded into $\mathrm{GL}_n(\mathfrak{K})$ then the same is true for all subgroups of $\mathfrak{G}$.

($\Leftarrow$) Let $\mathcal{K}_n$ be the class of all groups with a faithful n-linear representation. Then $\mathcal{K}_n$ is closed under substructures. Furthermore, we have $\mathcal{K}_n \in \mathrm{PC}_1(\mathrm{FO}, \{\cdot, ^{-1}, e\})$. By the preceding lemma, it follows that $\mathcal{K}_n = \mathrm{Mod}(T)$, for some $T \subseteq \forall$.

Suppose that $\mathfrak{G} \notin \mathcal{K}_n$. Then there is some formula $\forall \bar{x} \varphi(\bar{x}) \in T$ such that $\mathfrak{G} \vDash \neg \forall \bar{x} \varphi$. Fix some $\bar{a} \subseteq G$ with $\mathfrak{G} \vDash \neg \varphi(\bar{a})$. Setting $\mathfrak{G}_0 := \langle\!\langle \bar{a} \rangle\!\rangle_{\mathfrak{G}}$ it follows that $\mathfrak{G}_0 \vDash \neg \varphi(\bar{a})$. Hence, we have found a finitely generated subgroup with $\mathfrak{G}_0 \notin \mathcal{K}_n$.

We conclude this section with a characterisation of classes closed under unions of chains.

Theorem 1.5 (Chang, Łoś, Suszko). *For a first-order theory T and a set Φ of sentences, the following statements are equivalent:*

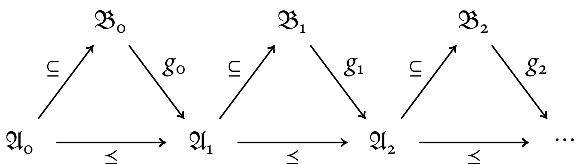
(1) *If $(\mathfrak{A}_i)_{i<\alpha}$ is a chain such that $\bigcup_i \mathfrak{A}_i \vDash T$ and $\mathfrak{A}_i \vDash T \cup \Phi$, for all $i < \alpha$, then $\bigcup_i \mathfrak{A}_i \vDash \Phi$.*

(2) *Φ is equivalent modulo T to a set of first-order $\forall\exists$-formulae.*

Proof. (2) $\Rightarrow$ (1) was already proved in Lemma C2.1.8. For the other direction, set $\Psi := (T \cup \Phi)^{\vDash}_{\forall\exists}$. It is sufficient to show that $T \cup \Psi \vDash \Phi$.

We prove that every model $\mathfrak{D} \vDash T \cup \Psi$ is elementary equivalent to the union $\mathfrak{C} := \bigcup_{i<\omega} \mathfrak{A}_i$ of a chain $(\mathfrak{A}_i)_{i<\omega}$ where $\mathfrak{C} \vDash T$ and $\mathfrak{A}_i \vDash T \cup \Phi$, for all $i < \omega$. Since Φ is closed under unions of chains it follows that $\mathfrak{C} \vDash \Phi$, which implies that $\mathfrak{D} \vDash \Phi$.

Fix an arbitrary model $\mathfrak{D} \vDash T \cup \Psi$. By induction on i, we construct an elementary chain $(\mathfrak{A}_i)_{i<\omega}$, extensions $\mathfrak{B}_i \supseteq \mathfrak{A}_i$, and embeddings $g_i : \mathfrak{B}_i \to \mathfrak{A}_{i+1}$ such that the following diagram commutes:



Furthermore, we ensure that

$$\mathfrak{B}_i \vDash T \cup \Phi \quad \text{and} \quad \langle \mathfrak{B}_i, \bar{a}^i \rangle \leq_{\forall\exists} \langle \mathfrak{A}_i, \bar{a}^i \rangle,$$

where $\bar{a}^i$ is some enumeration of A_i.

We start with $\mathfrak{A}_0 := \mathfrak{D}$. Suppose that $\mathfrak{A}_i$ has already been defined. $\mathfrak{A}_0 \leq \mathfrak{A}_i$ implies that $\mathfrak{A}_i \vDash \Psi$. If we set $\Delta := \forall\exists$ in Corollary c2.5.6 then we obtain an extension $\mathfrak{B}_i \supseteq \mathfrak{A}_i$ such that

$$\mathfrak{B}_i \vDash T \cup \Phi \quad \text{and} \quad \langle \mathfrak{A}_i, \bar{a}^i \rangle \leq_{\exists\forall} \langle \mathfrak{B}_i, \bar{a}^i \rangle,$$

that is, $\langle \mathfrak{B}_i, \bar{a}^i \rangle \leq_{\forall\exists} \langle \mathfrak{A}_i, \bar{a}^i \rangle$. Since $\exists \subseteq \forall\exists$, we can use Corollary c2.5.4 to find an elementary extension $\mathfrak{A}_{i+1} \succeq \mathfrak{A}_i$ and an embedding $g_i : \mathfrak{B}_i \to \mathfrak{A}_{i+1}$ with $g_i \restriction A_i = \mathrm{id}_{A_i}$.

Let $\mathfrak{C} := \bigcup_{i<\omega} \mathfrak{A}_i = \bigcup_{i<\omega} g_i(\mathfrak{B}_i)$. Since $(\mathfrak{A}_i)_i$ is an elementary chain it follows that $\mathfrak{A}_0 \leq \mathfrak{C}$. Hence, we have found a model $\mathfrak{C} \vDash T$ that is the union of a chain of models of $T \cup \Phi$. □

2. *Quantifier elimination*

Some theories, like the theory of dense linear orders or the theory of algebraically closed fields, have the pleasant property that every formula is equivalent to a quantifier-free one. We can use this fact to deduce some useful information about the theory.

First of all, we gain a better understanding of which relations are definable since we only need to consider relations definable by quantifier-free formulae. For instance, every definable relation of an algebraically closed field is given by finitely many equations and inequations between polynomials.

Secondly, we can sometimes use this fact to prove that a theory is complete. Since every sentence is equivalent to a quantifier-free one we only have to check that, for every quantifier-free sentence φ, the theory determines whether φ does hold or not. In particular, if the signature contains neither constant symbols nor 0-ary relation symbols then the only quantifier-free sentences are true and false and this question becomes trivial.

Definition 2.1. (a) Let L be a logic, $\Delta, \Gamma \subseteq L$, and $\mathcal{K}$ a class of L-interpretations. We say that Γ is a *Δ-elimination set over $\mathcal{K}$* if, for all sets $\Phi \subseteq \Delta$ there exists a set $\Psi \subseteq \Gamma$ such that

$$\mathfrak{J} \vDash \Phi \quad \text{iff} \quad \mathfrak{J} \vDash \Psi\,, \qquad \text{for all } \mathfrak{J} \in \mathcal{K}\,.$$

(b) We say that a class of Σ-structures $\mathcal{K}$ admits *quantifier elimination for* $\mathrm{FO}_{\kappa\aleph_0}$ if $\mathrm{QF}^{<\omega}_{\kappa\aleph_0}[\Sigma]$ is an $\mathrm{FO}^{<\omega}_{\kappa\aleph_0}[\Sigma]$-elimination set over $\mathcal{K}$. In particular, we say that a first-order theory T admits *quantifier elimination* if $\mathrm{Mod}(T)$ admits quantifier elimination for FO.

In terms of type spaces we obtain the following characterisation.

Lemma 2.2. *Let L be a logic, $T \subseteq L$ a theory, and $\Gamma \subseteq \Delta \subseteq L$ fragments of L/T that are both closed under disjunctions. The following statements are equivalent.*

(1) *Γ is an Δ-elimination set over T.*

(2) *The function $\mathfrak{S}(i) : \mathfrak{S}((L/T)|_\Delta) \to \mathfrak{S}((L/T)|_\Gamma)$ corresponding to the inclusion map $i : L|_\Gamma \to L|_\Delta$ is a homeomorphism.*

Proof. Replacing L by L/T we may w.l.o.g. assume that $T = \varnothing$. Further, note that $S(i)(\mathfrak{p}) = \mathfrak{p} \cap \Gamma$ and that, according to Lemma C3.2.2, the closed sets of $\mathfrak{S}(L|_\Delta)$ and $\mathfrak{S}(L|_\Gamma)$ are of the form $\langle \Phi \rangle_{L|_\Delta}$ and $\langle \Psi \rangle_{L|_\Gamma}$, for $\Phi \subseteq \Delta$ and $\Psi \subseteq \Gamma$.

(1) $\Rightarrow$ (2) Suppose that Γ is a Δ-elimination set. We have to prove that $S(i)$ is continuous and that it has a continuous inverse. It follows from Proposition C3.2.11 that $S(i)$ is a continuous surjection. To prove that it is also injective suppose that $\mathfrak{p}, \mathfrak{q} \in S(\Delta)$ are two types with $S(i)(\mathfrak{p}) = S(i)(\mathfrak{q})$. By assumption there exist sets $\Phi, \Psi \subseteq \Gamma$ such that $\mathfrak{p} \equiv \Phi$ and $\mathfrak{q} \equiv \Psi$. Consequently, we have

$$\Phi \subseteq \mathfrak{p}^{\vDash}_{\Gamma} = \mathfrak{p} \cap \Gamma = S(i)(\mathfrak{p}) = S(i)(\mathfrak{q}) = \mathfrak{q} \cap \Gamma \subseteq \mathfrak{q}\,.$$

Hence, $\mathfrak{p} = \Phi^{\vDash}_{\Delta} \subseteq \mathfrak{q}^{\vDash}_{\Delta} = \mathfrak{q}$. By symmetry, we also have $\mathfrak{q} \subseteq \mathfrak{p}$. It follows that $\mathfrak{p} = \mathfrak{q}$, as desired.

We have shown that $S(i)$ has an inverse. It remains to prove that $S(i)^{-1}$ is continuous. Let $\langle \Phi \rangle$ be a closed subset of $S(\Delta)$. We have to show that $(S(i)^{-1})^{-1}[\langle \Phi \rangle] = S(i)[\langle \Phi \rangle]$ is closed in $S(\Gamma)$. By assumption there is a set $\Psi \subseteq \Gamma$ with $\Phi \equiv \Psi$. We claim that $S(i)[\langle \Phi \rangle] = \langle \Psi \rangle$.

First, suppose that $\mathfrak{p} \in \langle \Phi \rangle$. Then $\Psi \subseteq \mathfrak{p}$ and

$$S(i)(\mathfrak{p}) = p \cap \Gamma \supseteq \Psi .$$

Hence, $S(i)(\mathfrak{p}) \in \langle \Psi \rangle$. Conversely, suppose that $\mathfrak{p} \in \langle \Psi \rangle$. Then $\Psi \subseteq \mathfrak{p} \subseteq S(i)^{-1}(\mathfrak{p})$ implies that $\Phi \subseteq S(i)^{-1}(\mathfrak{p})$. Hence, $S(i)^{-1}(\mathfrak{p}) \in \langle \Phi \rangle$, i.e., $\mathfrak{p} \in S(i)[\langle \Phi \rangle]$

(2) $\Rightarrow$ (1) Suppose that $S(i)$ is a homeomorphism. To show that Γ is a Δ-elimination set let $\Phi \subseteq \Delta$. Since $\langle \Phi \rangle$ is a closed subset of $S(\Delta)$ it follows that $C := S(i)[\langle \Phi \rangle]$ is a closed subset of $S(\Gamma)$. Hence, there exists a set $\Psi \subseteq \Gamma$ such that $C = \langle \Psi \rangle$. We claim that $\Phi \equiv \Psi$.

First, suppose that $\mathfrak{I} \vDash \Phi$ and let $\mathfrak{p} := \mathrm{Th}_{\Delta}(\mathfrak{I})$. Then $\mathfrak{p} \in \langle \Phi \rangle$ implies that

$$\mathrm{Th}_{\Gamma}(\mathfrak{I}) = \mathfrak{p} \cap \Gamma = S(i)(\mathfrak{p}) \in \langle \Psi \rangle .$$

Hence, $\mathfrak{I} \vDash \Psi$. Conversely, suppose that $\mathfrak{I} \vDash \Psi$ and let $\mathfrak{p} := \mathrm{Th}_{\Delta}(\mathfrak{I})$. Then $S(i)(\mathfrak{p}) = \mathfrak{p} \cap \Gamma \in \langle \Psi \rangle$. Hence, we have $\mathfrak{p} = S(i)^{-1}(\mathfrak{p} \cap \Gamma) \in \langle \Phi \rangle$ and, therefore, $\mathfrak{I} \vDash \Phi$. ☐

For first-order logic we can get a slightly stronger result.

Lemma 2.3. *Let $T \subseteq \mathrm{FO}^{\bar{s}}[\Sigma]$ be a first-order theory and $\Delta \subseteq \Phi \subseteq \mathrm{FO}^{\bar{s}}[\Sigma]$ sets of formulae. If*

$$\mathfrak{p}|_{\Delta} = \mathfrak{q}|_{\Delta} \quad \textit{implies} \quad \mathfrak{p}|_{\Phi} = \mathfrak{q}|_{\Phi} , \quad \textit{for all } \mathfrak{p}, \mathfrak{q} \in S^{\bar{s}}(T) ,$$

then every formula of Φ is equivalent modulo T to a finite boolean combination of formulae of Δ.

Proof. Let Δ_+ and Φ_+ be the boolean closures of, respectively, Δ and Φ. The inclusion $i : \Delta_+ \to \Phi_+$ induces an injective homomorphism

$$f : \mathfrak{Lb}(\Delta_+/T) \to \mathfrak{Lb}(\Phi_+/T)\,.$$

By Corollary B5.6.11, we obtain a surjective continuous map

$$\mathrm{spec}(f) : \mathfrak{S}_{\Phi_+}(T) \to \mathfrak{S}_{\Delta_+}(T) : \mathfrak{p} \mapsto \mathfrak{p}|_{\Delta_+}\,.$$

By assumption, this map is injective. Hence, $\mathrm{spec}(f)$ is in fact an isomorphism. By Corollary B5.6.11 it follows that so is f. Consequently, for every formula $\varphi \in \Phi_+$, there is some formula $\delta \in \Delta_+$ with

$$[\varphi]_{\equiv_T} = f([\delta]_{\equiv_T}) = [i(\delta)]_{\equiv_T} = [\delta]_{\equiv_T}\,.$$

It follows that $\varphi \equiv \delta$ modulo T, as desired. □

If Γ is a Δ-elimination set and the logic in question is compact then it follows that every Δ-formula is equivalent to a single Γ-formula. In particular, if a theory T admits quantifier elimination then every first-order formula is equivalent modulo T to a quantifier-free one.

Lemma 2.4. *Let $\Delta, \Gamma, T \subseteq$ FO sets of first-order formulae where Γ is closed under conjunctions. Γ is a Δ-elimination set over T if and only if, for every formula $\varphi \in \Delta$, there exists a formula $\psi \in \Gamma$ such that $\varphi \equiv \psi$ modulo T.*

Proof. ($\Leftarrow$) is trivial. For ($\Rightarrow$), let $\varphi \in \Delta$. By assumption, there exists a set $\Psi \subseteq \Gamma$ such that $\varphi \equiv \Psi$ modulo T. By compactness, we can find a finite subset $\Psi_0 \subseteq \Psi$ such that $T \cup \Psi_0 \vDash \varphi$. If we set $\psi := \bigwedge \Psi_0 \in \Gamma$ then we have $T \vDash \varphi \leftrightarrow \psi$. □

Lemma 2.5. *Let T be a first-order theory and $\varphi(\bar{x})$ a formula. The following statements are equivalent:*

(1) *There exists a quantifier-free formula $\psi(\bar{x})$ that is equivalent to φ modulo T.*

(2) *For all models* $\mathfrak{A}$ *and* $\mathfrak{B}$ *of* T *and all* $\bar{a} \in A^{<\omega}$ *and* $\bar{b} \in B^{<\omega}$ *with* $\langle \mathfrak{A}, \bar{a} \rangle \equiv_0 \langle \mathfrak{B}, \bar{b} \rangle$, *we have*

$$\mathfrak{A} \vDash \varphi(\bar{a}) \quad \text{iff} \quad \mathfrak{B} \vDash \varphi(\bar{b}) .$$

Proof. (1) ⇒ (2) Suppose that $\langle \mathfrak{A}, \bar{a} \rangle \equiv_0 \langle \mathfrak{B}, \bar{b} \rangle$. By (1), there exists a quantifier-free formula $\psi(\bar{x}) \equiv \varphi$ modulo T. It follows that

$$\begin{aligned} \mathfrak{A} \vDash \varphi(\bar{a}) \quad &\text{iff} \quad \mathfrak{A} \vDash \psi(\bar{a}) \\ &\text{iff} \quad \mathfrak{B} \vDash \psi(\bar{b}) \quad \text{iff} \quad \mathfrak{B} \vDash \varphi(\bar{b}) . \end{aligned}$$

(2) ⇒ (1) Let Φ the closure of QF $\cup \{\varphi\}$ under boolean operations. Condition (2) can be written as

$$\mathfrak{p}|_{\mathrm{QF}} = \mathfrak{q}|_{\mathrm{QF}} \quad \text{implies} \quad \mathfrak{p}|_{\Phi} = \mathfrak{q}|_{\Phi} , \qquad \text{for all } \mathfrak{p}, \mathfrak{q} \in S^n(T) .$$

Consequently the claim follows by Lemma 2.3. □

Theorem 2.6. *Let* T *be a first-order theory. The following statements are equivalent:*

(1) T *admits quantifier elimination.*

(2) *For all models* $\mathfrak{A}$ *and* $\mathfrak{B}$ *of* T *and all* $\bar{a} \in A^{<\omega}$ *and* $\bar{b} \in B^{<\omega}$, *we have*

$$\langle \mathfrak{A}, \bar{a} \rangle \equiv_0 \langle \mathfrak{B}, \bar{b} \rangle \quad \text{iff} \quad \langle \mathfrak{A}, \bar{a} \rangle \equiv_{\mathrm{FO}} \langle \mathfrak{B}, \bar{b} \rangle .$$

(3) *For all models* $\mathfrak{A}$ *and* $\mathfrak{B}$ *of* T, *each quantifier-free formula* $\varphi(\bar{x}, y)$, *and all elements* $\bar{a} \in A^{<\omega}$ *and* $\bar{b} \in B^{<\omega}$ *with* $\langle \mathfrak{A}, \bar{a} \rangle \equiv_0 \langle \mathfrak{B}, \bar{b} \rangle$ *we have*

$$\mathfrak{A} \vDash \exists y \varphi(\bar{a}, y) \quad \textit{implies} \quad \mathfrak{B} \vDash \exists y \varphi(\bar{b}, y) .$$

Proof. (1) ⇔ (2) follows from Lemma 2.5 and (2) ⇒ (3) is trivial.

(3) ⇒ (1) W.l.o.g. we may assume that φ is written without universal quantifiers. By induction on φ, we construct a quantifier-free formula φ°

with $\varphi^\circ \equiv \varphi$ modulo T. If φ is quantifier-free we are done. For boolean combinations we can set

$$(\neg\varphi)^\circ := \neg\varphi^\circ, \quad (\varphi \vee \psi)^\circ := \varphi^\circ \vee \psi^\circ, \quad (\varphi \wedge \psi)^\circ := \varphi^\circ \wedge \psi^\circ.$$

Finally, suppose that $\varphi = \exists y \psi(\bar{x}, y)$. By (3) and Lemma 2.5, we can find a quantifier-free formula φ° such that $\varphi^\circ \equiv \exists y \psi^\circ(\bar{x}, y)$ modulo T. □

A useful simple criterion for quantifier elimination is the following one.

Definition 2.7. (a) Let T be a theory and $\mathfrak{A}$ a model of $T_\forall^\vDash$. An *algebraic prime model* of T over $\mathfrak{A}$ is an embedding $f : \mathfrak{A} \to \mathfrak{B}$ into a model of T such that any other embedding $g : \mathfrak{A} \to \mathfrak{C}$ into a model of T factorises as $g = h \circ f$, for some embedding $h : \mathfrak{B} \to \mathfrak{C}$. We say that T has *algebraic prime models* if, for every $\mathfrak{A} \vDash T_\forall^\vDash$, there is an algebraic prime model of T over $\mathfrak{A}$.

(b) Let $\mathfrak{A} \subseteq \mathfrak{B}$. We say that $\mathfrak{A}$ is *simply closed* in $\mathfrak{B}$ if, for every quantifier-free formula $\varphi(\bar{x}, y)$ and all elements $\bar{a} \subseteq A$

$$\mathfrak{B} \vDash \exists y \varphi(\bar{a}, y) \quad \text{implies} \quad \mathfrak{A} \vDash \exists y \varphi(\bar{a}, y)\,.$$

Proposition 2.8. *Let T be a first-order theory with algebraic prime models such that, whenever $\mathfrak{A} \subseteq \mathfrak{B}$ are both models of T then $\mathfrak{A}$ is simply closed in $\mathfrak{B}$. Then T admits quantifier elimination.*

Proof. Let $\mathfrak{A}$ and $\mathfrak{B}$ be models of T and suppose that

$$\langle \mathfrak{A}, \bar{a} \rangle \equiv_0 \langle \mathfrak{B}, \bar{b} \rangle\,.$$

By Theorem 2.6, it is sufficient to show that

$$\mathfrak{A} \vDash \exists y \varphi(\bar{a}, y) \quad \text{implies} \quad \mathfrak{B} \vDash \exists y \varphi(\bar{b}, y)\,,$$

for every quantifier-free formula φ. Let $f : \langle\!\langle \bar{a} \rangle\!\rangle_{\mathfrak{A}} \to \mathfrak{C}$ be the algebraic prime model of T over $\langle\!\langle \bar{a} \rangle\!\rangle_{\mathfrak{A}}$. Since $\langle\!\langle \bar{a} \rangle\!\rangle_{\mathfrak{A}} \cong \langle\!\langle \bar{b} \rangle\!\rangle_{\mathfrak{B}}$ we obtain an embedding $g : \langle\!\langle \bar{b} \rangle\!\rangle \to \mathfrak{C}$ with $g(\bar{b}) = f(\bar{a})$. By definition of an algebraic prime model there exist embeddings $h : \mathfrak{C} \to \mathfrak{A}$ and $k : \mathfrak{C} \to \mathfrak{B}$ such that $h(f(\bar{a})) = \bar{a}$ and $k(g(\bar{b})) = \bar{b}$.

Suppose that $\mathfrak{A} \vDash \varphi(\bar{a}, b)$. By assumption $\mathfrak{C}$ is simply closed in $\mathfrak{A}$. Hence,

$$\mathfrak{C} \vDash \varphi(f(\bar{a}), c), \quad \text{for some } c \in C.$$

It follows that $\mathfrak{B} \vDash \varphi(k(f(\bar{a})), k(c))$. Since $k(f(\bar{a})) = k(g(\bar{b})) = \bar{b}$ this implies that

$$\mathfrak{B} \vDash \varphi(\bar{b}, k(c)),$$

as desired. □

Similar to the characterisation of Theorem 2.6 above we can describe infinitary first-order theories admitting quantifier elimination.

Theorem 2.9. *Let $\mathcal{K}$ be a class of structures. The following statements are equivalent:*

(1) *$\mathcal{K}$ admits quantifier elimination for* $\mathrm{FO}_{\infty\aleph_0}$.

(2) $\mathfrak{A} \sqsubseteq_0^{\aleph_0} \mathfrak{B}$ *for all structures* $\mathfrak{A}, \mathfrak{B} \in \mathcal{K}$.

(3) $\mathfrak{A} \cong_0^{\aleph_0} \mathfrak{B}$ *for all structures* $\mathfrak{A}, \mathfrak{B} \in \mathcal{K}$.

(4) *For all structures $\mathfrak{A}, \mathfrak{B} \in \mathcal{K}$ and all $\bar{a} \in A^{<\omega}$ and $\bar{b} \in B^{<\omega}$, we have*

$$\langle \mathfrak{A}, \bar{a} \rangle \equiv_0 \langle \mathfrak{B}, \bar{b} \rangle \quad \text{iff} \quad \langle \mathfrak{A}, \bar{a} \rangle \equiv_\infty \langle \mathfrak{B}, \bar{b} \rangle .$$

Proof. (1) ⇒ (4) Suppose that $\langle \mathfrak{A}, \bar{a} \rangle \equiv_0 \langle \mathfrak{B}, \bar{b} \rangle$. By (1), there exists a set $\Phi(\bar{x}) \subseteq \mathrm{QF}^{<\omega}_{\infty\aleph_0}$ such that $\Phi(\bar{a})$ is equivalent to $\mathrm{tp}_{\mathrm{FO}_{\infty\aleph_0}}(\bar{a}/\mathfrak{A})$ on structures of $\mathcal{K}$. Hence, $\mathfrak{A} \vDash \Phi(\bar{a})$ implies that $\mathfrak{B} \vDash \Phi(\bar{b})$, and it follows that $\mathrm{tp}_{\mathrm{FO}_{\infty\aleph_0}}(\bar{b}/\mathfrak{B}) = \mathrm{tp}_{\mathrm{FO}_{\infty\aleph_0}}(\bar{a}/\mathfrak{A})$.

(4) ⇒ (1) Let $\varphi(\bar{x}) \in \mathrm{FO}_{\infty\aleph_0}$. For each pair of types $\mathfrak{p} \in \langle\varphi\rangle$ and $\mathfrak{q} \in \langle\neg\varphi\rangle$ there exists a quantifier-free formula $\psi_{\mathfrak{pq}}$ such that $\psi_{\mathfrak{pq}} \in \mathfrak{p}$ and $\neg\psi_{\mathfrak{pq}} \in \mathfrak{q}$. It follows that the formula

$$\bigvee_{\mathfrak{p}\in\langle\varphi\rangle} \bigwedge_{\mathfrak{q}\in\langle\neg\varphi\rangle} \psi_{\mathfrak{pq}}$$

is equivalent to φ on structures of $\mathcal{K}$. (Note that the above disjunction and the conjunctions are over sets of formulae since, up to logical equivalence, the number of quantifier-free formulae with a given number of free variables can be bounded in terms of the size of the signature.)

(2) ⇒ (3) $\mathfrak{A} \sqsubseteq_0^{\aleph_0} \mathfrak{B}$ and $\mathfrak{B} \sqsubseteq_0^{\aleph_0} \mathfrak{A}$ implies that $\mathfrak{A} \cong_0^{\aleph_0} \mathfrak{B}$.

(3) ⇒ (4) Suppose that $\mathfrak{A} \cong_0^{\aleph_0} \mathfrak{B}$. Then $\mathrm{pIso}_{\aleph_0}(\mathfrak{A}, \mathfrak{B}) = I_\infty^{\aleph_0}(\mathfrak{A}, \mathfrak{B})$. Hence,

$$\langle \mathfrak{A}, \bar{a} \rangle \equiv_0 \langle \mathfrak{B}, \bar{b} \rangle \quad \text{implies} \quad \langle \mathfrak{A}, \bar{a} \rangle \equiv_\infty \langle \mathfrak{B}, \bar{b} \rangle .$$

(4) ⇒ (2) We have to show that $\mathrm{pIso}_{\aleph_0}(\mathfrak{A}, \mathfrak{B})$ has the forth property with respect to itself. Since $\mathrm{pIso}_{\aleph_0}(\mathfrak{A}, \mathfrak{B}) = I_\infty^{\aleph_0}(\mathfrak{A}, \mathfrak{B})$ and the latter set has the back-and-forth property with respect to itself the claim follows. □

Corollary 2.10. *Let T be a first-order theory. If T admits quantifier elimination for* $\mathrm{FO}_{\infty\aleph_0}$ *then it also admits quantifier elimination for* FO.

Proof. Suppose that $\mathfrak{A}$ and $\mathfrak{B}$ are models of T with

$$\langle \mathfrak{A}, \bar{a} \rangle \equiv_0 \langle \mathfrak{B}, \bar{b} \rangle .$$

If T admits quantifier elimination for $\mathrm{FO}_{\infty\aleph_0}$ then it follows that

$$\langle \mathfrak{A}, \bar{a} \rangle \equiv_\infty \langle \mathfrak{B}, \bar{b} \rangle .$$

In particular, we have

$$\langle \mathfrak{A}, \bar{a} \rangle \equiv_{\mathrm{FO}} \langle \mathfrak{B}, \bar{b} \rangle .$$

By Theorem 2.6 it follows that T admits quantifier elimination. □

Example. (a) In Corollary C4.4.7 we have shown that we have $\mathfrak{A} \cong_0^{\aleph_0} \mathfrak{B}$ for all open dense linear orders $\mathfrak{A}$ and $\mathfrak{B}$. By the preceding theorem, it follows that the class of open dense linear orders admits quantifier elimination for $\mathrm{FO}_{\infty\aleph_0}$.

(b) Further examples like the theory of algebraically closed fields will be treated in the sections below.

Exercise 2.1. Let $\mathfrak{Z} := \langle \mathbb{Z}, s \rangle$ where $s : x \mapsto x + 1$ is the successor function. Prove that $\mathrm{Th}(\mathfrak{Z})$ admits quantifier-elimination.

To check whether a theory T admits quantifier elimination for $\mathrm{FO}_{\infty\aleph_0}$ the most useful characterisation is statement (2) of Theorem 2.9. In fact, we do not need to consider all models of T, only sufficiently large ones.

Lemma 2.11. *Let L be a logic and $\Gamma, \Delta \subseteq L$ sets such that Γ is a Δ-elimination set over $\mathcal{K}_0$. If $\mathcal{K}$ is a class of L-interpretations such that, for every $\mathfrak{J} \in \mathcal{K}$, there exists some $\mathfrak{J}_0 \in \mathcal{K}_0$ with $\mathfrak{J}_0 \equiv_L \mathfrak{J}$ then Γ is a Δ-elimination set over $\mathcal{K}$.*

Proof. Given $\Phi \subseteq \Delta$ there exists a set $\Psi \subseteq \Gamma$ such that

$$\mathfrak{J} \models \Phi \quad \text{iff} \quad \mathfrak{J} \models \Psi \,, \qquad \text{for all } \mathfrak{J} \in \mathcal{K}_0 \,.$$

We claim that these sets are also equivalent for all interpretations in $\mathcal{K}$. Let $\mathfrak{J} \in \mathcal{K}$. By assumption, there exists an interpretation $\mathfrak{J}_0 \in \mathcal{K}_0$ with $\mathfrak{J}_0 \equiv_L \mathfrak{J}$. Consequently, we have

$$\mathfrak{J} \models \Phi \quad \text{iff} \quad \mathfrak{J}_0 \models \Phi \quad \text{iff} \quad \mathfrak{J}_0 \models \Psi \quad \text{iff} \quad \mathfrak{J} \models \Psi \,. \qquad \square$$

Corollary 2.12. *Let T be a first-order theory and $\mathcal{K} \subseteq \mathrm{Mod}(T)$ a class such that, for every model $\mathfrak{A} \models T$, there is some structure $\mathfrak{B} \in \mathcal{K}$ with $\mathfrak{A} \preceq \mathfrak{B}$. If $\mathfrak{A} \sqsubseteq_0^{\aleph_0} \mathfrak{B}$, for all $\mathfrak{A}, \mathfrak{B} \in \mathcal{K}$, then T admits quantifier elimination.*

If we replace in the proof of Theorem 2.9 all quantifier-free formulae by arbitrary first-order formulae we obtain the following result.

Theorem 2.13. *Let $\mathcal{K}$ be a class of structures. The following statements are equivalent:*

(1) *Over the class $\mathcal{K}$ every $\mathrm{FO}^{<\omega}_{\infty\aleph_0}$-formula is equivalent to an infinite boolean combination of first-order formulae.*

(2) $\mathfrak{A} \sqsubseteq_{\mathrm{FO}}^{\aleph_0} \mathfrak{B}$ *for all structures $\mathfrak{A}, \mathfrak{B} \in \mathcal{K}$.*

(3) $\mathfrak{A} \cong_{\mathrm{FO}}^{\aleph_0} \mathfrak{B}$ *for all structures $\mathfrak{A}, \mathfrak{B} \in \mathcal{K}$.*

(4) *For all structures* $\mathfrak{A}, \mathfrak{B} \in \mathcal{K}$ *and all* $\bar{a} \in A^{<\omega}$ *and* $\bar{b} \in B^{<\omega}$, *we have*

$$\langle \mathfrak{A}, \bar{a} \rangle \equiv_{\mathrm{FO}} \langle \mathfrak{B}, \bar{b} \rangle \quad \text{iff} \quad \langle \mathfrak{A}, \bar{a} \rangle \equiv_{\infty} \langle \mathfrak{B}, \bar{b} \rangle .$$

We conclude this section with a closer look at quantifier elimination for the quantifier $\exists^{\aleph_0}$.

Definition 2.14. A first-order theory $T \subseteq \mathrm{FO}^0[\Sigma]$ is *graduated* if, for every formula $\varphi(\bar{x}, y) \in \mathrm{FO}^{<\omega}[\Sigma]$, there exists a number $k < \omega$ such that, for every model $\mathfrak{A}$ of T and all parameters $\bar{a} \subseteq A$,

$$\left|\varphi(\bar{a}, y)^{\mathfrak{A}}\right| < \aleph_0 \quad \text{implies} \quad \left|\varphi(\bar{a}, y)^{\mathfrak{A}}\right| \leq k .$$

Theorem 2.15. *A theory* $T \subseteq \mathrm{FO}^0[\Sigma]$ *is graduated if and only if* FO *is an* $\mathrm{FO}(\exists^{\aleph_0})$*-elimination set over* T.

Proof. ($\Rightarrow$) For every formula $\varphi \in \mathrm{FO}(\exists^{\aleph_0})$, we construct an equivalent first-order formula by induction on φ. Suppose that $\varphi = \exists^{\aleph_0} y \psi(\bar{x}, y)$. By inductive hypothesis, we may assume that ψ is a first-order formula. Since T is graduated there exists a number $k < \omega$ such that

$$\varphi(\bar{x}) \equiv \exists y_0 \cdots \exists y_k \Big[\bigwedge_{0 \leq i < l \leq k} y_i \neq y_l \wedge \bigwedge_{i \leq k} \psi(\bar{x}, y_i) \Big] .$$

($\Leftarrow$) For a contradiction, suppose that T is not graduated but FO is an $\mathrm{FO}(\exists^{\aleph_0})$-elimination set over T. Then there exists a formula $\varphi(\bar{x}, y)$ such that, for every $n < \omega$, there is a model $\mathfrak{A}_n$ of T and parameters $\bar{a}_n \subseteq A_n$ such that

$$n < \left|\varphi(\bar{a}_n, y)^{\mathfrak{A}_n}\right| < \aleph_0 .$$

By assumption there exists a set $\Phi \subseteq \mathrm{FO}$ such that $\neg\exists^{\aleph_0} y \varphi \equiv \Phi$. Then the set

$$\Psi := \Phi \cup \big\{ \exists y_0 \cdots \exists y_n [\textstyle\bigwedge_{i<l} y_i \neq y_l \wedge \bigwedge_i \varphi(\bar{x}, y_i)] \bigm| n < \omega \big\}$$

is inconsistent. On the other hand, for every finite subset $\Psi_0 \subseteq \Psi$, there is some number $m < \omega$ such that

$$\Psi_0 \subseteq \Phi \cup \{ \exists y_0 \cdots \exists y_n [\bigwedge_{i<l} y_i \neq y_l \wedge \bigwedge_i \varphi(\bar{x}, y_i)] \mid n < m \} .$$

Consequently, $\mathfrak{A}_m \vDash \Psi_0(\bar{a}_m)$. By the Compactness Theorem, it follows that Ψ is satisfiable. Contradiction. □

3. *Existentially closed structures*

In this section we study classes where each structure passes the Tarski-Vaught Test.

Definition 3.1. (a) A first-order formula is *primitive* if it is of the form

$$\varphi(\bar{x}) = \exists \bar{y} \bigwedge_{i<n} \psi_i(\bar{x}, \bar{y}) ,$$

where each ψ_i is a literal.

(b) Let $\mathcal{K}$ be a class of structures. A structure $\mathfrak{A} \in \mathcal{K}$ is *existentially closed* (in $\mathcal{K}$) if, for every extension $\mathfrak{B} \supseteq \mathfrak{A}$ with $\mathfrak{B} \in \mathcal{K}$, we have

$$\mathfrak{B} \vDash \varphi(\bar{a}) \quad \text{implies} \quad \mathfrak{A} \vDash \varphi(\bar{a}) ,$$

for each primitive formula $\varphi(\bar{x})$ and all parameters $\bar{a} \subseteq A$.

(c) We call a theory T *existentially closed,* or *model-complete,* if every model of T is existentially closed in $\mathrm{Mod}(T)$. A theory T_{ec} is the *existential closure,* or *model companion,* of the theory T if

$$\mathrm{Mod}(T_{ec}) = \{ \mathfrak{A} \in \mathrm{Mod}(T) \mid \mathfrak{A} \text{ is existentially closed in } \mathrm{Mod}(T) \} .$$

Remark. The existential closure of a theory does not necessarily exist since the class

$$\mathcal{K} := \{ \mathfrak{A} \in \mathrm{Mod}(T) \mid \mathfrak{A} \text{ is existentially closed} \}$$

does not need to be axiomatisable. But if it exists then it is unique since $\mathrm{Mod}(T_0) = \mathcal{K} = \mathrm{Mod}(T_1)$ implies that $T_0 \equiv T_1$.

Theorem 3.2. *Let T be a first-order theory. The following statements are equivalent:*

(1) *T is existentially closed.*

(2) *$\mathfrak{B} \vDash \varphi(\bar{a})$ implies $\mathfrak{A} \vDash \varphi(\bar{a})$, for all models $\mathfrak{A} \subseteq \mathfrak{B}$ of T, all parameters $\bar{a} \subseteq A$, and every first-order formula φ.*

(3) *Every embedding between models of T is elementary.*

(4) *For every formula φ, there exists a universal formula ψ such that $T \vDash \varphi \leftrightarrow \psi$.*

(5) *For every primitive formula φ, there exists a universal formula ψ such that $T \vDash \varphi \leftrightarrow \psi$.*

Proof. (4) ⇒ (3) follows from the fact that universal formulae are preserved in substructures.

(3) ⇒ (2) If $\mathfrak{A} \subseteq \mathfrak{B}$ and $\mathfrak{B} \vDash \varphi(\bar{a})$, for $\bar{a} \subseteq A$, then $\mathfrak{A} \preceq \mathfrak{B}$ implies that $\mathfrak{A} \vDash \varphi(\bar{a})$.

(2) ⇒ (1) is trivial.

(1) ⇒ (5) Let φ be a primitive formula. By (1), the negation $\neg\varphi$ is preserved by embeddings between models of T. Hence, we can use Corollary 1.3 to find an existential formula ψ equivalent to $\neg\varphi$ modulo T. The negation $\neg\psi$ is the desired universal formula.

(5) ⇒ (4) W.l.o.g. we may assume that φ is in prenex normal form, say, $\varphi = Q_0 x_0 \cdots Q_{n-1} x_{n-1} \psi$ with ψ quantifier-free. We prove the claim by induction on n. By inductive hypothesis, there exists a universal formula $\forall \bar{y} \vartheta$ equivalent to $Q_1 x_1 \cdots Q_{n-1} x_{n-1} \psi$. If $Q_0 = \forall$ then $\forall x_0 \forall \bar{y} \vartheta$ is the desired formula. Suppose that $Q_0 = \exists$. Let $\bigvee_i \chi_i$ be the disjunctive normal form of $\neg\vartheta$. By (5), there exists a universal formula $\forall \bar{z}^i \eta_i$ that is equivalent to $\exists \bar{y} \chi_i$. Consequently, we have

$$\exists \bar{y} \neg\vartheta \equiv \bigvee_i \exists \bar{y} \chi_i \equiv \bigvee_i \forall \bar{z}^i \eta_i \equiv \forall \bar{z}^0 \cdots \forall \bar{z}^m \bigvee_i \eta_i \,.$$

Let $\bar{z} = \bar{z}^0 \dots \bar{z}^m$ and let $\bigvee_i \beta_i$ be the disjunctive normal form of $\bigwedge_i \neg\eta_i$. It follows that

$$\varphi = \exists x_0 \forall \bar{y} \vartheta \equiv \exists x_0 \exists \bar{z} \bigwedge_i \neg\eta_i \equiv \bigvee_i \exists x_0 \exists \bar{z} \beta_i \,.$$

Applying (5) again, we obtain universal formula $\forall \bar{y}^i \gamma_i$ equivalent to $\exists x_0 \exists \bar{z} \beta_i$. Hence,

$$\varphi \equiv \bigvee_i \forall \bar{y}^i \gamma_i \equiv \forall \bar{y}^0 \cdots \forall \bar{y}^k \bigvee_i \gamma_i \,,$$

as desired. □

Corollary 3.3. *Let T be a first-order theory.*

(a) *If T admits quantifier elimination then it is existentially closed.*

(b) *If T has algebraic prime models then it is existentially closed if and only if it admits quantifier elimination.*

(c) *If T is a Skolem theory then it is existentially closed.*

Proof. (a) and (c) follow from Theorem 3.2 (4). (b) follows from (a) and Proposition 2.8. □

Example. The theory of open dense linear orders is existentially closed. Other examples such as the theory of divisible abelian groups and the theory of algebraically closed fields will be treated below.

Let us give some basic properties of existentially closed theories. We start with a partial converse of Corollary 3.3 (a).

Lemma 3.4. *Let T be a theory such that* $\mathrm{Mod}(T)$ *is closed under substructures. Then T is existentially closed if and only if T admits quantifier elimination.*

Proof. We have already seen that every theory admitting quantifier elimination is existentially closed. For the converse, suppose that T is existentially closed. We apply Theorem 2.6 (3). Suppose that $\mathfrak{A}$ and $\mathfrak{B}$ are models of T with elements $\bar{a} \subseteq A$ and $\bar{b} \subseteq B$ such that

$$\langle \mathfrak{A}, \bar{a} \rangle \equiv_0 \langle \mathfrak{B}, \bar{b} \rangle \,.$$

Let $\varphi(\bar{x}, y)$ be a quantifier-free formula such that

$$\mathfrak{A} \vDash \exists y \varphi(\bar{a}, y) \,.$$

Since $\mathrm{Mod}(T)$ is closed under substructures, we have $\langle\langle \bar a \rangle\rangle_{\mathfrak A} \vDash T$. By Theorem 3.2, it follows that $\langle\langle \bar a \rangle\rangle_{\mathfrak A} \preceq \mathfrak A$. Hence,

$$\langle\langle \bar a \rangle\rangle_{\mathfrak A} \vDash \exists y \varphi(\bar a, y) .$$

Fix some element $c \in \langle\langle \bar a \rangle\rangle_{\mathfrak A}$ such that $\langle\langle \bar a \rangle\rangle_{\mathfrak A} \vDash \varphi(\bar a, c)$. There exists some term t such that $c = t^{\langle\langle \bar a \rangle\rangle_{\mathfrak A}}(\bar a)$. Therefore, we have

$$\langle\langle \bar a \rangle\rangle_{\mathfrak A} \vDash \varphi(\bar a, t(\bar a)) .$$

It follows that

$$\langle\langle \bar b \rangle\rangle_{\mathfrak B} \vDash \varphi(\bar b, t(\bar b)) .$$

Consequently, $\mathfrak B \vDash \exists y \varphi(\bar b, y)$. □

Lemma 3.5. *Let T be an existentially closed theory. Then T is the existential closure of $T_\forall^{\vDash}$.*

Proof. Consider structures $\mathfrak A \subseteq \mathfrak B$ where $\mathfrak A$ is a model of T and $\mathfrak B$ a model of $T_\forall^{\vDash}$. Suppose that $\mathfrak B \vDash \varphi(\bar a)$ where $\varphi(\bar x)$ is a primitive formula and $\bar a \subseteq A$. We have to show that $\mathfrak A \vDash \varphi(\bar a)$. By Lemma 1.1, we can find a model $\mathfrak C$ of T with $\mathfrak B \subseteq \mathfrak C$. Since existential formulae are preserved in extensions it follows that $\mathfrak C \vDash \varphi(\bar a)$. As T is existentially closed and we have $\mathfrak A \subseteq \mathfrak C$, it follows that $\mathfrak A \preceq \mathfrak C$. Hence, $\mathfrak A \vDash \varphi(\bar a)$, as desired. □

Lemma 3.6. *If T is existentially closed then $T \equiv T_{\forall\exists}^{\vDash}$.*

Proof. If T is existentially closed then every chain is elementary. Hence, $\mathrm{Mod}(T)$ is closed under unions of chains and the claim follows by Theorem 1.5. □

For $\forall\exists$-theories, one can embed every model into an existentially closed one.

Proposition 3.7. *Let $T \subseteq \forall\exists$ be a first-order theory and $\mathfrak A$ an infinite Σ-structure with $\mathfrak A \vDash T_\forall^{\vDash}$. Then there exists an existentially closed model $\mathfrak B$ of T of size $|B| = |A| \oplus |\Sigma|$ such that $\mathfrak A \subseteq \mathfrak B$.*

Proof. By Lemma 1.1, there exists a model $\mathfrak{C}$ of T with $\mathfrak{A} \subseteq \mathfrak{C}$. By the Theorem of Löwenheim and Skolem we may choose $\mathfrak{C}$ of size $|C| = |A| \oplus |\Sigma|$. To conclude the proof we construct an existentially closed elementary extension $\mathfrak{B} \succeq \mathfrak{C}$ of size $|B| = |C|$. The construction is similar to the one used in Theorem C2.3.6 to find a Skolem theory.

Claim. *For every infinite model $\mathfrak{A} \vDash T$, there exists an extension $\mathfrak{A}^+ \supseteq \mathfrak{A}$ of size $|A^+| = |A| \oplus |\Sigma|$ such that $\mathfrak{A}^+ \vDash T$ and, for every $\exists$-formula $\varphi(\bar{x})$ and all $\bar{a} \subseteq A$,*

$$\mathfrak{A}^+ \nvDash \varphi(\bar{a}) \quad \textit{implies} \quad \mathfrak{B} \nvDash \varphi(\bar{a}), \quad \textit{for all } \mathfrak{B} \supseteq \mathfrak{A}^+.$$

When we have proved the claim then we can find the desired existentially closed structure $\mathfrak{B} \succeq \mathfrak{C}$ as follows. We define an increasing chain $(\mathfrak{B}_n)_{n<\omega}$ by

$$\mathfrak{B}_0 := \mathfrak{C} \quad \text{and} \quad \mathfrak{B}_{n+1} := (\mathfrak{B}_n)^+.$$

Since $T \subseteq \forall\exists$ it follows that $\mathfrak{B} := \bigcup_n \mathfrak{B}_n$ is a model of T. By definition, we have $\mathfrak{C} \subseteq \mathfrak{B}$ and

$$|B| = \sup_n |B_n| \leq \aleph_0 \oplus |C| = |C|.$$

It remains to show that $\mathfrak{B}$ is existentially closed. If $\varphi(\bar{x})$ is an $\exists$-formula and $\bar{a} \subseteq B$ then there is some index $n < \omega$ such that $\bar{a} \subseteq B_n$. Consequently, if there exists a model $\mathfrak{D} \supseteq \mathfrak{B}$ of T with $\mathfrak{D} \vDash \varphi(\bar{a})$ then, by construction of $\mathfrak{B}_{n+1} = \mathfrak{B}_n^+$, we have $\mathfrak{B}_{n+1} \vDash \varphi(\bar{a})$. Since φ is existential and $\mathfrak{B}_{n+1} \subseteq \mathfrak{B}$ it follows that $\mathfrak{B} \vDash \varphi(\bar{a})$, as desired.

It remains to prove the above claim. Let $\kappa := |A| \oplus |\Sigma|$ and fix an enumeration $\langle \varphi_\alpha, \bar{a}_\alpha \rangle_{\alpha<\kappa}$ of all pairs $\langle \varphi, \bar{a} \rangle$ where $\varphi \in \exists$ and $\bar{a} \in A^{<\omega}$. We define an increasing sequence $(\mathfrak{A}_\alpha)_{\alpha<\kappa}$ of models of T as follows. We start with $\mathfrak{A}_0 := \mathfrak{A}$ and, for limit ordinals δ, we set $\mathfrak{A}_\delta := \bigcup_{\alpha<\delta} \mathfrak{A}_\alpha$. For the successor step, we distinguish two cases. If there is some extension $\mathfrak{B} \supseteq \mathfrak{A}_\alpha$ with $\mathfrak{B} \vDash \varphi_\alpha(\bar{a}_\alpha)$ then, by the Theorem of Löwenheim and Skolem, we can choose such an extension of size $|B| \leq |A_\alpha| \oplus |\Sigma|$ and we set $\mathfrak{A}_{\alpha+1} := \mathfrak{B}$. Otherwise, we set $\mathfrak{A}_{\alpha+1} := \mathfrak{A}_\alpha$.

We claim that $\mathfrak{A}^+ := \bigcup_\alpha \mathfrak{A}_\alpha$ is the desired structure. By induction on α, it follows that $|A_\alpha| \leq \kappa$. Hence, $|A^+| \leq \kappa$. Furthermore, if there exists an extension $\mathfrak{B} \supseteq \mathfrak{A}$ such that $\mathfrak{B} \vDash \varphi(\bar{a})$, for some $\varphi \in \exists$ and $\bar{a} \subseteq A$, then there exits an index α with $\varphi = \varphi_\alpha$ and $\bar{a} = \bar{a}_\alpha$. Hence, $\mathfrak{A}_{\alpha+1}$ is some extension of $\mathfrak{A}_\alpha$ with $\mathfrak{A}_{\alpha+1} \vDash \varphi(\bar{a})$. Since φ is existential and $\mathfrak{A}_{\alpha+1} \subseteq \mathfrak{A}^+$ it follows that $\mathfrak{A}^+ \vDash \varphi(\bar{a})$, as desired. □

Example. A field is existentially closed if and only if it is algebraically closed. Since the theory of fields is $\forall\exists$-axiomatisable it follows that every field has an algebraically closed extension.

Lemma 3.8. *Let $T \subseteq \forall\exists$ be a theory with existential closure T_{ec}.*

(a) *Every model of T_{ec} is a model of T.*

(b) *Every model of T has an extension that is a model of T_{ec}.*

Proof. (a) holds by definition of an existential closure and (b) follows from Proposition 3.7. □

Corollary 3.9. *If T_{ec} is the existential closure of a theory $T \subseteq \forall\exists$ then*

$$T_\forall^\vDash = (T_{\text{ec}})_\forall^\vDash \quad \textit{and} \quad (T_{\text{ec}})_\forall^\vDash \subseteq T \subseteq T_{\text{ec}} \,.$$

Proof. The equation $T_\forall^\vDash = (T_{\text{ec}})_\forall^\vDash$ follows by the preceding lemma and Lemma 1.1. Hence, we have $(T_{\text{ec}})_\forall^\vDash = T_\forall^\vDash \subseteq T$. Finally, $\text{Mod}(T_{\text{ec}}) \subseteq \text{Mod}(T)$ implies $T \subseteq T_{\text{ec}}$. □

4. *Abelian groups*

As a simple example of existentially closed theories we consider theories of abelian groups.

Definition 4.1. Let $\mathfrak{G} = \langle G, \cdot, ^{-1}, e \rangle$ be a group. A *torsion element* of $\mathfrak{G}$ is an element $a \neq e$ such that $a^n = e$, for some $0 < n < \omega$. The set of all torsion elements of $\mathfrak{G}$ (including e) is denoted by

$$\text{tor}(\mathfrak{G}) := \{\, a \in G \mid a^n = e \text{ for some } n > 0 \,\} \,.$$

We say that $\mathfrak{G}$ is *torsion-free* if $\operatorname{tor}(\mathfrak{G}) = \{e\}$.

Example. $\operatorname{tor}\langle \mathbb{R}/\mathbb{Z}, +, -, 0\rangle = \langle \mathbb{Q}/\mathbb{Z}, +, -, 0\rangle$.

Lemma 4.2. *If $\mathfrak{G}$ is an abelian group then $\operatorname{tor}(\mathfrak{G})$ is a normal subgroup of $\mathfrak{G}$.*

Proof. In an abelian group every subgroup is normal. Hence, we only need to show that $\operatorname{tor}(\mathfrak{G})$ is closed under the group operations. Let $a, b \in \operatorname{tor}(\mathfrak{G})$. Then there are numbers $m, n > 0$ such that $a^m = e$ and $b^n = e$. Consequently, we have

$$(ab^{-1})^{mn} = a^{mn}(b^{mn})^{-1} = e^n(e^m)^{-1} = e\,,$$

which implies that $ab^{-1} \in \operatorname{tor}(\mathfrak{G})$. □

Corollary 4.3. *Every abelian group $\mathfrak{G}$ can we written as direct sum*

$$\mathfrak{G} \cong \mathfrak{H} \oplus \operatorname{tor}(\mathfrak{G}) \quad \textit{where } \mathfrak{H} \textit{ is torsion-free.}$$

Definition 4.4. An *ordered group* is a structure $\mathfrak{G} = \langle G, \circ, ^{-1}, e, <\rangle$ such that $\langle G, \circ, ^{-1}, e\rangle$ forms a group, $<$ is a linear order on G, and we have

$$a < b \quad \text{implies} \quad ac < bc \quad \text{and} \quad ca < cb\,, \quad \text{for all } a, b, c \in G\,.$$

Exercise 4.1. Prove that there are exactly two orderings $\sqsubseteq$ on $\mathbb{Q}$ such that $\langle \mathbb{Q}, +, \sqsubseteq\rangle$ is an ordered group.

Lemma 4.5. *Every ordered group is torsion-free.*

Proof. For a contradiction, suppose that there is some element $a \neq e$ such that $a^n = e$, for some $n > 0$. If $a > e$ then we have $a^{k+1} > a^k$, for all k. It follows that $e < a < \dots < a^n = e$. Contradiction. Similarly, $a < e$ implies that $e > a > \dots > a^n = e$. □

Definition 4.6. (a) An abelian group $\mathfrak{G} = \langle G, +, -, 0\rangle$ is *divisible* if, for every element $a \in G$ and all numbers $0 < n < \omega$, there exists an element $b \in G$ with $nb = a$. We denote this element by a/n.

(b) Let DAG be the theory of all divisible torsion-free abelian groups with more than one element. Let ODAG be the theory of all ordered divisible abelian groups with more than one element.

If $\mathfrak{G}$ is divisible and torsion-free we can define an action $\mathbb{Q} \times G \to G$ by setting $\frac{m}{n} \cdot a := m(a/n)$.

Lemma 4.7. *Every divisible torsion-free abelian group $\mathfrak{G}$ is a $\mathbb{Q}$-module.*

Exercise 4.2. Let $\mathfrak{G}$ be a divisible abelian group that is not torsion-free. Show that $\mathfrak{G}$ is no $\mathbb{Q}$-module under the above action.

Theorem 4.8. *For every divisible torsion-free abelian group $\mathfrak{G}$ there is a cardinal κ such that $\mathfrak{G} \cong \langle \mathbb{Q}, + \rangle^{(\kappa)}$.*

Proof. $\mathfrak{G}$ is a $\mathbb{Q}$-module, that is, a $\mathbb{Q}$-vector space. By Theorem B6.4.12, we have $\mathfrak{G} \cong \mathbb{Q}^{(\kappa)}$ where κ is the dimension of $\mathfrak{G}$. □

Corollary 4.9. *For every divisible torsion-free abelian group $\mathfrak{G}$ there exists a linear order $<$ such that $\langle \mathfrak{G}, < \rangle$ is an ordered group.*

Proof. We can take the lexicographic order on $\mathbb{Q}^{(\kappa)}$. □

Every abelian group can be embedded into a divisible one.

Definition 4.10. Let $\mathfrak{G}$ be an abelian group. The *divisible closure* of $\mathfrak{G}$ is the group $\operatorname{div}(\mathfrak{G})$ with universe

$$\operatorname{div}(G) := \{ \langle a, n \rangle \mid a \in G,\ 0 < n < \omega \} / {\sim}$$

where

$$\langle a, m \rangle \sim \langle b, n \rangle \quad :\text{iff} \quad na = mb\,.$$

We denote the $\sim$-class of $\langle a, n \rangle$ by a/n. The group operations of $\operatorname{div}(\mathfrak{G})$ are given by

$$a/m + b/n := (na + mb)/mn \quad \text{and} \quad -(a/m) := (-a)/m\,.$$

Theorem 4.11. *Let $\mathfrak{G}$ be an abelian group.*

(a) *The divisible closure* $\operatorname{div}(\mathfrak{G})$ *of* $\mathfrak{G}$ *is a divisible abelian group.*

(b) *If* $\mathfrak{G}$ *is torsion-free then so is* $\operatorname{div}(\mathfrak{G})$.

(c) *If* $\mathfrak{G}$ *is ordered then so is* $\operatorname{div}(\mathfrak{G})$.

(d) *The embedding* $\mathfrak{G} \to \operatorname{div}(\mathfrak{G}) : a \mapsto a/1$ *is an algebraic prime model for the theory* DAG *and* ODAG, *respectively.*

Proof. (a) If $a/m = a'/m'$ then we have $a/m + b/n = a'/m' + b/n$ since $m'a = ma'$ implies that

$$\begin{aligned} m'n(na + mb) &= m'n^2a + mm'nb \\ &= mn^2a' + mm'nb = mn(na' + m'b)\,. \end{aligned}$$

Hence, + is well-defined. In a similar way one shows that − is also well-defined and that $\operatorname{div}(\mathfrak{G})$ forms an abelian group with unit $0/1$.

Note that $\operatorname{div}(\mathfrak{G})$ is divisible since $n(a/mn) = (na/mn) = a/m$.

(b) Suppose that $n(a/m) = 0/1$. Then we have $na = m0 = 0$, which implies that $a = 0$ since $\mathfrak{G}$ is torsion-free.

(c) We define the order on $\operatorname{div}(\mathfrak{G})$ by setting

$$a/m < b/n \quad :\text{iff} \quad na < mb\,.$$

To see that this definition turns $\operatorname{div}(\mathfrak{G})$ into an ordered group note that $na < mb$ implies

$$nk(ka + mc) < mk(kb + nc)\,.$$

Consequently,

$$a/m < b/n \quad \text{implies} \quad a/m + c/k < b/n + c/k\,.$$

(d) Let $g : \mathfrak{G} \to \mathfrak{H}$ be some embedding of $\mathfrak{G}$ into a model of DAG or ODAG. Then we obtain an embedding $\operatorname{div}(\mathfrak{G}) \to \mathfrak{H}$ by mapping $a/n \in \operatorname{div}(G)$ to the unique element $b \in H$ with $nb = g(a)$. □

Corollary 4.12. *Every abelian group can be embedded into a divisible abelian group.*

Corollary 4.13. *For every torsion-free abelian group* $\mathfrak{G}$, *there exists a cardinal* κ *such that* $\mathfrak{G}$ *can be embedded into* $\mathbb{Q}^{(\kappa)}$.

Corollary 4.14. DAG *and* ODAG *have algebraic prime models.*

In order to prove that DAG and ODAG admit quantifier elimination it remains to check that subgroups are simply closed.

Lemma 4.15. *If* $\mathfrak{G} \subseteq \mathfrak{H}$ *are torsion-free divisible abelian groups then* $\mathfrak{G}$ *is simply closed in* $\mathfrak{H}$. *The same holds if* $\mathfrak{G}$ *and* $\mathfrak{H}$ *are ordered.*

Proof. We have to show that

$$\mathfrak{H} \vDash \exists y \varphi(\bar{a}, y) \quad \text{implies} \quad \mathfrak{G} \vDash \exists y \varphi(\bar{a}, y),$$

for every quantifier-free formula φ and all $\bar{a} \subseteq G$. Suppose that $\varphi = \bigvee_i \bigwedge_k \psi_{ik}$ is in disjunctive normal form. If $\mathfrak{H} \vDash \varphi(\bar{a}, b)$ then there is some i such that $\mathfrak{H} \vDash \bigwedge_k \psi_{ik}(\bar{a}, b)$. Since each atomic formula can be written as

$$\sum_i m_i x_i + ny = 0 \quad \text{or} \quad \sum_i m_i x_i + ny < 0, \quad \text{for } m_i, n \in \mathbb{Z},$$

we may therefore assume that

$$\begin{aligned} \varphi = \bigwedge_k \sum_i m_{ki} x_i + n_k y = 0 \wedge \bigwedge_k \sum_i m'_{ki} x_i + n'_k y < 0 \\ \wedge \bigwedge_k \sum_i m''_{ki} x_i + n''_k y \neq 0 . \end{aligned}$$

Set $c_k := \sum_i m_{ki} a_i$, $c'_k := \sum_i m'_{ki} a_i$, and $c''_k := \sum_i m''_{ki} a_i$. These elements are in G and we have

$$\varphi \equiv \bigwedge_k c_k + n_k y = 0 \wedge \bigwedge_k c'_k + n'_k y < 0 \wedge \bigwedge_k c''_k + n''_k y \neq 0 .$$

If there is some k with $n_k \neq 0$ then

$$\mathfrak{H} \vDash \varphi(\bar{a}, -c_k/n_k)\,.$$

Since $-c_k/n_k \in G$ we are done. Therefore, we may assume that $n_k = 0$, for all k. Then

$$\varphi \equiv \bigwedge_k c'_k + n'_k y < 0 \ \wedge\ \bigwedge_k c''_k + n''_k y \neq 0\,.$$

Suppose that $n'_0, \dots, n'_{s-1} < 0$ and $n'_s, \dots, n'_{t-1} > 0$. Then this formula simplifies to

$$\varphi \equiv \bigwedge_{k=0}^{s-1} y > -c'_k/n'_k \ \wedge\ \bigwedge_{k=s}^{t-1} y < -c'_k/n'_k \ \wedge\ \bigwedge_k y \neq -c''_k/n''_k\,.$$

Setting $d_0 := \max\{\, -c'_k/n'_k \mid k < s \,\}$ and $d_1 := \min\{\, -c'_k/n'_k \mid s \leq k < t \,\}$ we obtain

$$\varphi \equiv y > d_0 \ \wedge\ y < d_1 \ \wedge\ \bigwedge_k y \neq -c''_k/n''_k\,.$$

Since $\mathfrak{H} \vDash \exists y \varphi(\bar{a}, y)$ it follows that $d_0 < d_1$. Hence, $d_0, d_1 \in G$ implies that G contains infinitely many elements b with $d_0 < b < d_1$. Consequently, we can find an element $b \in G$ with $d_0 < b < d_1$ such that $b \neq -c''_k/n''_k$, for all k. It follows that $\mathfrak{G} \vDash \varphi(\bar{a}, b)$. □

Theorem 4.16. DAG *and* ODAG *admit quantifier elimination.*

Proof. This follows from the preceding lemmas by Proposition 2.8. □

Corollary 4.17. DAG *is the existential closure of the theory of torsion-free abelian groups.* ODAG *is the existential closure of the theory of ordered abelian groups.*

5. *Fields*

Further classes with a well-behaved model theory are the class of algebraically closed fields and the class of real closed fields.

Definition 5.1. (a) The axiom system for the theory of *fields* is the set F consisting of all ring axioms together with the formulae

$$0 \neq 1 \quad \text{and} \quad \forall x \exists y [x \neq 0 \to x \cdot y = 1] .$$

(b) The theory ACF of *algebraically closed fields* is obtained from F by adding, for every $1 < n < \omega$, the sentence

$$\forall y_0 \cdots \forall y_{n-1} \exists x [x^n + y_{n-1} \cdot x^{n-1} + \cdots + y_1 \cdot x + y_0 = 0] .$$

(c) For a prime number p, we obtain the theory ACF_p of *algebraically closed fields of characteristic p* by adding to ACF the sentence

$$\underbrace{1 + \cdots + 1}_{p \text{ times}} = 0 .$$

Similarly, the theory ACF_0 of algebraically closed fields of characteristic 0 is obtained by adding all the sentences

$$\underbrace{1 + \cdots + 1}_{n \text{ times}} \neq 0 , \qquad \text{for all } 0 < n < \omega .$$

(d) We denote by RCF the axiom system for the theory of *real closed fields.* It consists of the axioms for an ordered field and the formulae

$$\forall x \exists y [y \cdot y = x \vee y \cdot y = -x] ,$$
$$\forall x_0 \cdots \forall x_{n-1} [x_0 \cdot x_0 + \cdots + x_{n-1} \cdot x_{n-1} + 1 \neq 0] ,$$
$$\forall y_0 \cdots \forall y_{2n} \exists x [x^{2n+1} + y_{2n} \cdot x^{2n} + \cdots + y_1 \cdot x^1 + y_0 = 0] ,$$

for all $n < \omega$.

Remark. (a) If $\mathfrak{K} \vDash F$ is a field then every atomic formula has the form $p(\bar{x}) = q(\bar{x})$ or, equivalently, $p(\bar{x}) - q(\bar{x}) = 0$, for polynomials $p, q \in \mathbb{Z}[\bar{x}]$.

(b) In Theorem B6.5.5 we have seen that $F_\forall^\vDash$ is the theory of integral domains.

Since the axiom systems F, ACF, ACF_p, and RCF consist solely of $\forall\exists$-sentences it follows by Lemma C2.1.8 that their model classes are closed under unions of chains.

Lemma 5.2. *If $(\mathfrak{K}_\alpha)_{\alpha<\kappa}$ is a chain of fields then their union $\bigcup_{\alpha<\kappa} \mathfrak{K}_\alpha$ is also a field. If every $\mathfrak{K}_\alpha$ is algebraically closed then so is the union. The same holds for real closed fields.*

Proposition 5.3. *Let κ be an infinite cardinal and let $\mathfrak{K}$ and $\mathfrak{L}$ be algebraically closed fields of transcendence degree at least κ. If $\mathfrak{K}$ and $\mathfrak{L}$ have the same characteristic then $\mathfrak{K} \cong_0^\kappa \mathfrak{L}$.*

Proof. First, note that $\mathrm{pIso}_\kappa(\mathfrak{K}, \mathfrak{L}) \neq \varnothing$ since it contains $1 \mapsto 1$. By symmetry, we therefore only need to prove the forth property.

Let $\bar{a} \mapsto \bar{b} \in \mathrm{pIso}_\kappa(\mathfrak{K}, \mathfrak{L})$ and $c \in K$. We denote by $\mathfrak{A}$ the subfield of $\mathfrak{K}$ generated by $\bar{a}$ and $\mathfrak{B}$ is the subfield of $\mathfrak{L}$ generated by $\bar{b}$. The partial isomorphism $\bar{a} \mapsto \bar{b}$ extends to an isomorphism $\pi : \mathfrak{A} \to \mathfrak{B}$.

If $c \in A$ then $d := \pi(c) \in B$ and $\bar{a}c \mapsto \bar{b}d \in \mathrm{pIso}_\kappa(\mathfrak{K}, \mathfrak{L})$.

Next we consider the case that c is algebraic over A. Let $p \in A[x]$ be the minimal polynomial of c. Consider the canonical extension $\pi' : \mathfrak{A}[x] \to \mathfrak{B}[x]$ of π and set $q := \pi'(p)$. Since $\mathfrak{L}$ is algebraically closed, q has some root $d \in L$. It follows that

$$\mathfrak{A}(c) \cong \mathfrak{A}[x]/(p) \cong \mathfrak{B}[x]/(q) \cong \mathfrak{B}(d)$$

and, hence, $\bar{a}c \mapsto \bar{b}d \in \mathrm{pIso}_\kappa(\mathfrak{K}, \mathfrak{L})$.

Finally, suppose that c is not algebraic over A. Since $\mathfrak{L}$ has transcendence degree at least κ, there is some element $d \in L$ that is transcendental over B. It follows that $\mathfrak{A}(c) \cong \mathrm{FF}(\mathfrak{A}[x]) \cong \mathrm{FF}(\mathfrak{B}[x]) \cong \mathfrak{B}(d)$. □

Theorem 5.4. ACF *admits quantifier elimination.*

Proof. By Corollary 2.12 and the preceding proposition it is sufficient to show that every algebraically closed field $\mathfrak{K}$ has an elementary extension $\mathfrak{L}$ with infinite transcendence degree. Let Δ be the elementary diagram of $\mathfrak{K}$ and let C be a countable set of new constant symbols. We set

$$\Phi := \{ p[\bar{c}] \neq 0 \mid p \in K[\bar{x}],\ \bar{c} \subseteq C \} .$$

If $\mathfrak{L} \models \Delta \cup \Phi$ then $\mathfrak{L} \succeq \mathfrak{K}$ implies that $\mathfrak{L}$ is an algebraically closed extension of $\mathfrak{K}$. Furthermore, C is an infinite algebraically independent subset of L.

Hence, it remains to prove that $\Delta \cup \Phi$ is satisfiable. By the Compactness Theorem we only have to check that all finite subsets of $\Delta \cup \Phi$ are satisfiable. Let $\Phi_0 \subseteq \Phi$ be finite and let $p_0, \dots, p_{n-1}$ be the polynomials appearing in Φ_0. Suppose that $p_0, \dots, p_{n-1} \in K[x_0, \dots, x_{k-1}]$. By induction on i, we find elements $a_i \in K$ such that $p_l[\bar{a}] \neq 0$, for all l.

Suppose that we have already chosen $a_0, \dots, a_{i-1}$. We partition the polynomials $p_0, \dots, p_{n-1}$ into three classes.

(i) those containing only variables from $x_0, \dots, x_{i-1}$;

(ii) those not in class (i) that contain only variables from $x_0, \dots, x_i$;

(iii) those containing some variable from $x_{i+1}, \dots, x_{k-1}$.

We choose an arbitrary element $a_i \in K$ such that, for every polynomial p_l in class (ii), we have $p_l[a_0, \dots, a_{i-1}, a_i] \neq 0$. This is possible since K is infinite and, for every polynomial $p_l[a_0, \dots, a_{i-1}, x_i]$, there are only finitely many values for x_i that we cannot choose.

Interpreting the constants $\bar{c}$ in Φ by the elements $\bar{a}$ we obtain a model $\langle \mathfrak{K}, \bar{a} \rangle$ of $\Delta \cup \Phi_0$. □

Theorem 5.5. *If p is a prime number or $p = 0$ then the theory ACF_p is complete.*

Proof. Let $\varphi \in \mathrm{FO}$ be a sentence. We have to show that either $\mathrm{ACF}_p \models \varphi$ or $\mathrm{ACF}_p \models \neg\varphi$. Since ACF admits quantifier elimination there exists a quantifier-free sentence ψ such that

$$\mathrm{ACF}_p \models \varphi \leftrightarrow \psi .$$

ψ is a boolean combination of sentences of the form $\vartheta := 1 + \cdots + 1 = 0$. But for each such sentence we either have $\mathrm{ACF}_p \vDash \vartheta$ or $\mathrm{ACF}_p \vDash \neg\vartheta$. □

After having seen that the theory of algebraically closed fields admits quantifier elimination we turn to real closed fields.

Proposition 5.6. $\mathrm{RCF}_\forall$ *is the theory of ordered integral domains.*

Proof. If $\mathfrak{R}$ is a substructure of a real closed field then it is a commutative ring without zero-divisors. Conversely, let $\mathfrak{R}$ be an ordered integral domain. We can order $\mathrm{FF}(\mathfrak{R})$ by

$$a/b > 0 \quad :\text{iff} \quad a, b > 0 \text{ or } a, b < 0\,.$$

By Theorem B6.6.13, we can embed $\mathrm{FF}(\mathfrak{R})$ into a real closed field. □

Proposition 5.7. RCF *has algebraic prime models.*

Proof. Let $\mathfrak{R}$ be an ordered integral domain and let $\mathfrak{K}$ be the real closure of $\mathrm{FF}(\mathfrak{R})$. We claim that $\mathfrak{K}$ is the algebraic prime model of $\mathfrak{R}$.

Fix an arbitrary ordered real closed extension $\mathfrak{L}$ of $\mathfrak{R}$. Then $\mathrm{FF}(\mathfrak{R}) \subseteq \mathfrak{L}$. Let

$$L_0 := \{\, a \in L \mid a \text{ is algebraic over } \mathrm{FF}(R)\,\}\,.$$

By Theorem B6.6.14, it follows that $\mathfrak{L}_0 \subseteq \mathfrak{L}$ is real closed. Since the order of $\mathfrak{L}_0$ extends the order of $\mathrm{FF}(\mathfrak{R})$, we can use Theorem B6.6.22 to find an isomorphism $\mathfrak{L}_0 \to \mathfrak{K}$. □

Lemma 5.8. *If $\mathfrak{K} \subseteq \mathfrak{L}$ are real closed fields then $\mathfrak{K}$ is simply closed in $\mathfrak{L}$.*

Proof. Let $\varphi(x, \bar{y})$ be quantifier-free and suppose that

$$\mathfrak{L} \vDash \varphi(a, \bar{b})\,, \quad \text{for some } a \in L,\ \bar{b} \subseteq K\,.$$

Note that, for a polynomial $p \in \mathbb{Z}[\bar{x}]$,

$$\begin{aligned} p[\bar{c}] \neq 0 &\quad \text{iff} \quad p[\bar{c}] > 0 \vee -p[\bar{c}] > 0\,, \\ p[\bar{c}] \leq 0 &\quad \text{iff} \quad p[\bar{c}] = 0 \vee -p[\bar{c}] > 0\,. \end{aligned}$$

Therefore, we may assume that $\varphi(x,\bar{y}) = \bigvee_{k\leq n} \psi_k(x,\bar{y})$ where each ψ_k is a conjunction of formulae of the form $p[x,\bar{y}] = 0$ or $p[x,\bar{y}] > 0$, for some $p \in \mathbb{Z}[x,\bar{y}]$. Fix some k such that $\mathfrak{L} \vDash \psi_k(a,\bar{b})$ and suppose that

$$\psi_k(x,\bar{b}) = \bigwedge_{i<m} p_i[x] = 0 \wedge \bigwedge_{i<n} q_i[x] > 0\,,$$

for $p_i, q_i \in K[x]$. If any of the p_i is nonzero then $p_i[a] = 0$ implies that a is algebraic over $\mathfrak{K}$. Since $\mathfrak{K}$ is real closed, it has no proper algebraic extension that is real. Therefore, $a \in K$ and we are done.

Hence, we may assume that

$$\psi_k(x,\bar{b}) = \bigwedge_{i<n} q_i[x] > 0\,.$$

The sign of $q_i[x]$ can only change at a root of q_i. As we have just seen each such root is an element of K. Therefore, there are elements $c_i, d_i \in K$ with $c_i < a < d_i$ and $q_i[x] > 0$, for all $x \in (c_i, d_i)$. Set

$$c := \max\{c_0,\dots,c_{n-1}\} \quad \text{and} \quad d := \min\{d_0,\dots,d_{n-1}\}\,.$$

Then $c < a < d$. Setting $a' := (c+d)/2 \in K$ it follows that $q_i[a'] > 0$, for all $i < n$. Hence, $\mathfrak{L} \vDash \psi_k(a',\bar{b})$. ☐

Theorem 5.9. RCF *admits quantifier elimination.*

Proof. We have shown that RCF has algebraic prime models and that real closed subfields are simply closed. Therefore, the claim follows by Proposition 2.8. ☐

Corollary 5.10. $\mathrm{RCF}^{\vDash} = \mathrm{Th}\langle\mathbb{R},+,-,\cdot,0,1,<\rangle$ *is complete and existentially closed.*

Proof. Every theory that admits quantifier elimination is existentially closed. To show that RCF is complete note that every real closed field $\mathfrak{R}$ has characteristic 0. Hence, $\mathbb{Q} \subseteq \mathfrak{R}$. Let $\mathbb{R}_{\mathrm{alg}}$ be the real closure of $\mathbb{Q}$, that is, the field of algebraic real numbers. It follows that $\mathbb{R}_{\mathrm{alg}}$ can be embedded into every real closed field $\mathfrak{R}$. Since RCF is existentially closed this embedding is elementary. Therefore, $\mathfrak{R} \equiv \mathbb{R}_{\mathrm{alg}}$. ☐

D2. *Products and varieties*

1. *Ultraproducts*

In Section D1.1 we have studied operations that preserve various fragments of first-order logic. But we have found no operation so far that preserves all first-order formulae. In this section we will show that ultraproducts have this property.

We generalise the notation of Section B3.2 as follows. Let $(\mathfrak{A}^i)_{i \in I}$ be a sequence of Σ-structures. For every sort s, we set

$$I_s := \{ i \in I \mid A^i_s \neq \varnothing \} .$$

If $\varphi(\bar{x})$ is a formula and $a_k \in \prod_{i \in I_{s_k}} A^i_{s_k}$, for $k < n$, are parameters then we define

$$[\![\varphi(\bar{a})]\!] := \{ i \in I_{s_0} \cap \dots \cap I_{s_{n-1}} \mid \mathfrak{A}^i \models \varphi(\bar{a}^i) \} .$$

Recall that, for a filter $\mathfrak{u}$ on I, we write

$$\bar{a} \sim_{\mathfrak{u}} \bar{b} \quad :\text{iff} \quad [\![\bar{a} = \bar{b}]\!] \in \mathfrak{u} ,$$

and $[\bar{a}]$ denotes the $\sim_{\mathfrak{u}}$-class of $\bar{a}$.

Theorem 1.1 (Łoś). *Let $(\mathfrak{A}^i)_{i \in I}$ be a sequence of Σ-structures and $\mathfrak{u}$ an ultrafilter on I. For every first-order formula $\varphi(\bar{x})$ and all parameters $a_k \in \prod_{i \in I_{s_k}} A^i_{s_k}$ we have*

$$\prod_i \mathfrak{A}^i / \mathfrak{u} \models \varphi([\bar{a}]) \quad \text{iff} \quad [\![\varphi(\bar{a})]\!] \in \mathfrak{u} .$$

Proof. Let $\mathfrak{A} := \prod_i \mathfrak{A}^i$ and $\mathfrak{B} := \prod_i \mathfrak{A}^i/\mathfrak{u}$. We prove the claim by induction on φ. If $\varphi = s = t$ then we have

$$\begin{aligned} \mathfrak{B} \vDash (s = t)([\bar{a}]) \quad &\text{iff} \quad s^{\mathfrak{B}}([\bar{a}]) = t^{\mathfrak{B}}([\bar{b}]) \\ &\text{iff} \quad s^{\mathfrak{A}}(\bar{a}) \sim_{\mathfrak{u}} t^{\mathfrak{A}}(\bar{a}) \\ &\text{iff} \quad [\![s(\bar{a}) = t(\bar{a})]\!] \in \mathfrak{u}\,. \end{aligned}$$

Similarly, if $\varphi = Rt_0 \dots t_{m-1}$ then

$$\begin{aligned} \mathfrak{B} \vDash (R\bar{t})([\bar{a}]) \quad &\text{iff} \quad \big\langle t_0^{\mathfrak{B}}([\bar{a}]), \dots, t_{m-1}^{\mathfrak{B}}([\bar{a}]) \big\rangle \in R^{\mathfrak{B}} \\ &\text{iff} \quad [\![(R\bar{t})(\bar{a})]\!] \in \mathfrak{u}\,. \end{aligned}$$

For the boolean operators, we have, by inductive hypothesis,

$$\begin{aligned} \mathfrak{B} \vDash \neg\varphi([\bar{a}]) \quad &\text{iff} \quad \mathfrak{B} \nvDash \varphi([\bar{a}]) \\ &\text{iff} \quad [\![\varphi(\bar{a})]\!] \notin \mathfrak{u} \\ &\text{iff} \quad [\![\neg\varphi(\bar{a})]\!] \in \mathfrak{u} \end{aligned}$$

and

$$\begin{aligned} \mathfrak{B} \vDash (\varphi \wedge \psi)([\bar{a}]) \quad &\text{iff} \quad \mathfrak{B} \vDash \varphi([\bar{a}]) \text{ and } \mathfrak{B} \vDash \psi([\bar{a}]) \\ &\text{iff} \quad [\![\varphi(\bar{a})]\!] \in \mathfrak{u} \text{ and } [\![\psi(\bar{a})]\!] \in \mathfrak{u} \\ &\text{iff} \quad [\![\varphi(\bar{a})]\!] \cap [\![\psi(\bar{a})]\!] \in \mathfrak{u} \\ &\text{iff} \quad [\![\varphi(\bar{a}) \wedge \psi(\bar{a})]\!] \in \mathfrak{u}\,. \end{aligned}$$

It remains to consider the case that $\varphi = \exists y \psi$. Let s be the sort of y. We have

$$\begin{aligned} &\mathfrak{B} \vDash \exists y \psi([\bar{a}], y) \\ \text{iff} \quad &I_s \in \mathfrak{u} \text{ and there is some } b \in \textstyle\prod_{i \in I_s} A_s^i \text{ such that } \mathfrak{B} \vDash \psi([\bar{a}], [b]) \\ \text{iff} \quad &\text{there is some } b \in \textstyle\prod_{i \in I_s} A_s^i \text{ such that } [\![\psi(\bar{a}, b)]\!] \in \mathfrak{u} \\ \text{iff} \quad &[\![\exists y \psi(\bar{a}, y)]\!] \in \mathfrak{u}\,. \end{aligned}$$

For the last step note that, on the one hand, we have

$$[\![\psi(\bar{a}, b)]\!] \subseteq [\![\exists y \psi(\bar{a}, y)]\!]\,.$$

Conversely, we can fix, for every $i \in [\![\exists y \psi(\bar{a}, y)]\!]$, some $b^i \in A^i_s$ such that $\mathfrak{A}^i \vDash \psi(\bar{a}^i, b^i)$. For $i \in I_s \setminus [\![\exists y \psi(\bar{a}, y)]\!]$, we choose an arbitrary element $b^i \in A^i_s$. With these choices we have

$$[\![\exists y \psi(\bar{a}, y)]\!] \subseteq [\![\psi(\bar{a}, b)]\!] . \qquad \square$$

Corollary 1.2. $\mathfrak{A} \preceq \mathfrak{A}^{\mathfrak{u}}$, *for all structures* $\mathfrak{A}$ *and every ultrafilter* $\mathfrak{u}$.

For the constructions below we frequently need a special kind of ultrafilter.

Definition 1.3. A filter $\mathfrak{u}$ on a set I is *regular* if there exists a sequence $(s_i)_{i \in I}$ of sets $s_i \in \mathfrak{u}$ such that, for every $k \in I$, the set $\{ i \mid k \in s_i \}$ is finite.

Lemma 1.4. *For every infinite set* I*, there exists a regular ultrafilter* $\mathfrak{u}$ *on* I*.*

Proof. Let $J := \{ s \subseteq I \mid |s| < \aleph_0 \}$. As I is infinite we have $|J| = |I|$ and there exists a bijection $f : J \to I$. Therefore, it is sufficient to construct a regular ultrafilter $\mathfrak{u}$ on J. Its image under f will be the desired regular ultrafilter on I.

For $i \in J$, set $s_i := \{ k \in J \mid i \subseteq k \}$. Since

$$s_i \cap s_j = \{ k \in J \mid i \cup j \subseteq k \} = s_{i \cup j}$$

it follows that $\mathfrak{v} := \{ s_i \mid i \in J \}$ has the finite intersection property. By Corollary B2.4.10, we can therefore find an ultrafilter $\mathfrak{u} \supseteq \mathfrak{v}$. Furthermore, $\mathfrak{u}$ is regular since, for every $k \in J$, the set

$$\{ i \in J \mid k \in s_i \} = \{ i \in J \mid i \subseteq k \}$$

is finite. $\square$

For ultrafilters over countable sets, regularity and non-principality coincide.

Lemma 1.5. *An ultrafilter* $\mathfrak{u}$ *over* ω *is regular if and only if it is non-principal.*

Proof. (⇒) Suppose that $\mathfrak{u}$ is principal, that is, $\mathfrak{u} = \Uparrow\{k\}$, for some k. If $(s_n)_{n<\omega}$ is a sequence of sets $s_n \in \mathfrak{u}$ then we have $k \in s_n$, for all n. Hence, $\mathfrak{u}$ cannot be regular.

(⇐) Suppose that $\mathfrak{u}$ is non-principal. For $n < \omega$, set $s_n := \Uparrow n$. Then we have $s_n \in \mathfrak{u}$ since $\omega \setminus s_n = [n] \notin \mathfrak{u}$. Furthermore, the set

$$\{ n < \omega \mid k \in s_n \} = \{ n < \omega \mid n \leq k \} = [k+1]$$

is finite, for every $k < \omega$. □

We use regular ultrafilters for the following alternative proof of the compactness theorem.

Proposition 1.6. *A set $\Phi \subseteq \mathrm{FO}[\Sigma, X]$ is satisfiable if and only if every finite subset $\Phi_0 \subseteq \Phi$ is satisfiable.*

Proof. Suppose that every finite subset of Φ is satisfiable. By replacing each free variable in Φ by a constant symbol we may assume that every formula in Φ is a sentence. We have to construct a model of Φ.

Let $\mathfrak{u}$ be a regular ultrafilter on Φ and fix a sequence $(s_\varphi)_{\varphi\in\Phi}$ with $s_\varphi \in \mathfrak{U}$ such that the sets

$$\Psi_\psi := \{ \varphi \in \Phi \mid \psi \in s_\varphi \}, \quad \text{for } \psi \in \Phi,$$

are finite. By assumption we can find models $\mathfrak{A}^\psi \vDash \Psi_\psi$, for every $\psi \in \Phi$. We claim that

$$\prod_{\psi\in\Phi} \mathfrak{A}^\psi/\mathfrak{u} \vDash \Phi$$

is the desired model of Φ. Let $\varphi \in \Phi$. Then

$$[\![\varphi]\!] \supseteq \{ \psi \in \Phi \mid \varphi \in \Psi_\psi \} = \{ \psi \in \Phi \mid \psi \in s_\varphi \} = s_\varphi \in \mathfrak{u}.$$

By Łoś' theorem it follows that $\prod_s \mathfrak{A}^s/\mathfrak{u} \vDash \varphi$. □

Lemma 1.7. *Let $\mathfrak{A}$ be a structure, κ an infinite cardinal, and $\mathfrak{u}$ a regular ultrafilter over a set I of size κ. If $\varphi(x)$ is a first-order formula such that $\varphi^{\mathfrak{A}}$ is infinite then*

$$|\varphi^{\mathfrak{A}^{\mathfrak{u}}}| = |\varphi^{\mathfrak{A}}|^{\kappa}.$$

Proof. By the Theorem of Łoś we have

$$\varphi^{\mathfrak{A}^{\mathfrak{u}}} = \{ [a] \in A^I/\mathfrak{u} \mid [\![\varphi(a)]\!] \in \mathfrak{u} \}.$$

Since $\varphi^{\mathfrak{A}} \neq \emptyset$, we can fix some element $c \in \varphi^{\mathfrak{A}}$. For every element $[a] \in \varphi^{\mathfrak{A}^{\mathfrak{u}}}$ with $s_a := [\![\varphi(a)]\!] \in \mathfrak{u}$, we define

$$a'_i := \begin{cases} a_i & \text{if } i \in s_a, \\ c & \text{otherwise}. \end{cases}$$

Note that we have $[a'] = [a]$ since $s_a \subseteq [\![a = a']\!] \in \mathfrak{u}$. Furthermore, $[\![\varphi(a')]\!] = I$. Consequently, we can define a function $f : \varphi^{\mathfrak{A}^{\mathfrak{u}}} \to (\varphi^{\mathfrak{A}})^I$ by mapping an element $[a] \in \varphi^{\mathfrak{A}^{\mathfrak{u}}}$ to some representative $a' \in [a]$ with $[\![\varphi(a')]\!] = I$. Note that f is injective since, for $[a] \neq [b]$, $f(a) \in [a]$ and $f(b) \in [b]$ implies that $f(a) \neq f(b)$. Therefore, we have $|\varphi^{\mathfrak{A}^{\mathfrak{u}}}| \leq |\varphi^{\mathfrak{A}}|^{\kappa}$.

It remains to prove that $|\varphi^{\mathfrak{A}^{\mathfrak{u}}}| \geq |\varphi^{\mathfrak{A}}|^{\kappa}$. Since $\mathfrak{u}$ is regular we can find sets $(s_i)_{i \in I}$ in $\mathfrak{u}$ such that the sets

$$w_k := \{ i \in I \mid k \in s_i \}$$

are finite. Since $\varphi^{\mathfrak{A}}$ is infinite we can fix bijections $\mu_k : (\varphi^{\mathfrak{A}})^{w_k} \to \varphi^{\mathfrak{A}}$, for $k \in I$. For $a \in (\varphi^{\mathfrak{A}})^I$, we define a sequence $a^{\mu} \in (\varphi^{\mathfrak{A}})^I$ by

$$a^{\mu}_i := \mu_i(a \upharpoonright w_i), \quad \text{for } i \in I.$$

Then $[\![\varphi(a^{\mu})]\!] = I$ which implies, by the Theorem of Łoś, that $[a^{\mu}] \in \varphi^{\mathfrak{A}^{\mathfrak{u}}}$. To conclude the proof it is therefore sufficient to show that the mapping $a \mapsto [a^{\mu}]$ is injective. If $a \neq b$ then there is some index $i \in I$ with $a_i \neq b_i$. Hence, $a \upharpoonright w_k \neq b \upharpoonright w_k$, for every k with $i \in w_k$, that is, for every $k \in s_i$. Consequently, $s_i \subseteq [\![a^{\mu} \neq b^{\mu}]\!] \in \mathfrak{u}$. □

Corollary 1.8. *Let κ be an infinite cardinal. Every structure $\mathfrak{A}$ has an elementary extension $\mathfrak{B}$ such that, for every first-order formula $\varphi(\bar{x})$, either*

$$\left|\varphi^{\mathfrak{B}}\right| < \aleph_0 \quad or \quad \left|\varphi^{\mathfrak{B}}\right| = \left|\varphi^{\mathfrak{A}}\right|^{\kappa}.$$

Forming an ultraproduct of a sequence of structures corresponds to taking the limit of their theories in the type space.

Lemma 1.9. *Let $T \subseteq \mathrm{FO}$ and $X \subseteq S^{\bar{s}}(T)$ a set of $\bar{s}$-types. For every accumulation point $\mathfrak{p}$ of X, there exist an ultrafilter $\mathfrak{u}$ on I, a sequence of structures $(\mathfrak{A}_i)_{i\in I}$, and parameters $\bar{a}^i \subseteq A_i$, $i \in I$, with $\mathrm{tp}(\bar{a}^i/\mathfrak{A}_i) \in X$ such that*

$$\mathfrak{p} = \mathrm{tp}\big([(\bar{a}^i)_i] \big/ \textstyle\prod_i \mathfrak{A}_i/\mathfrak{u}\big)\,.$$

Proof. Let $I := \mathfrak{p}$ and fix a regular ultrafilter $\mathfrak{u}$ over $\mathfrak{p}$. Then there exists a sequence $(s_\varphi)_{\varphi\in\mathfrak{p}}$ of sets $s_\varphi \in \mathfrak{u}$ such that, for every $i \in \mathfrak{p}$, the set $\Phi_i := \{\varphi \in \mathfrak{p} \mid i \in s_\varphi\}$ is finite. Since $\mathfrak{p}$ is an accumulation point of X we can find elements $\mathfrak{q}_i \in \langle\Phi_i\rangle \cap X \neq \emptyset$. Fix $\mathfrak{A}_i$ and $\bar{a}^i$ such that $\mathrm{tp}(\bar{a}^i/\mathfrak{A}_i) = \mathfrak{q}_i$, and set $\mathfrak{B} := \prod_{i\in I}\mathfrak{A}_i/\mathfrak{u}$ and $\bar{b} := [(\bar{a}^i)_i]$.

If $i \in s_\varphi$ then $\varphi \in \Phi_i$ which implies $\mathfrak{A}_i \vDash \varphi(\bar{a}^i)$. Therefore, we have $s_\varphi \subseteq [\![\varphi(\bar{a}^i)]\!] \in \mathfrak{u}$, for every $\varphi \in \mathfrak{p}$. By the Theorem of Łoś, it follows that $\mathfrak{B} \vDash \mathfrak{p}(\bar{b})$, that is, $\mathfrak{p} = \mathrm{tp}(\bar{b}/\mathfrak{B})$. □

2. *The theorem of Keisler and Shelah*

According to the Amalgamation Theorem any two elementary equivalent structures have a common elementary extension. In this section we prove the Theorem of Keisler and Shelah, which states that this extension can be taken as an ultrapower with respect to the same ultrafilter.

To construct such an ultrafilter $\mathfrak{u}$, we choose a sufficiently large cardinal λ. Starting with the trivial filter $\{\lambda\}$ on λ, we construct larger and larger filters until we have found the desired ultrafilter. In each step, we have to ensure that the filter we construct is general enough in the sense of being consistent with sufficiently many additional conditions. The precise definition are as follows.

Definition 2.1. Let λ be an infinite cardinal, $P \subseteq \mathscr{P}(\lambda)$, and $C \subseteq \lambda$. Recall that $\mathrm{cl}_\uparrow(P)$ denotes the filter generated by P.

(a) P *forces* C if $C \in \mathrm{cl}_\uparrow(P)$.

(b) P is *consistent* with C if it does not force the complement $\lambda \setminus C$.

(c) P *decides* C if it forces C or $\lambda \setminus C$.

Remark. (a) Note that $\mathrm{cl}_\uparrow(P)$ is an ultrafilter if, and only if, for every set $C \subseteq \lambda$, P forces exactly one of C and $\lambda \setminus C$.

(b) P is not consistent with C if, and only if, there is a finite subset $P_0 \subseteq P$ such that $\bigcap P_0 \cap C = \varnothing$. Hence, P is consistent with C if, and only if, $P \cup \{C\}$ does have the finite intersection property.

Definition 2.2. Let λ be an infinite cardinal and let μ be the least cardinal such that $2^\mu > \lambda$.

(a) We denote by $(<\mu)^\lambda$ the set of all functions $\lambda \to \kappa$ for a cardinal $\kappa < \mu$.

(b) Let $m < \omega$ and $\gamma < \mu$ be ordinals, let $\bar{f} = (f_i)_{i<\gamma}$, $\bar{f}' = (f'_i)_{i<m}$ and $\bar{g} = (g_i)_{i<m}$ be sequences of functions $f_i, f'_i, g_i : \lambda \to \mu$, and let $\bar{\beta} = (\beta_i)_{i<\gamma}$ be a sequence of ordinals $\beta_i < \mu$. A *condition* is a set of the form

$$[\![\bar{f} = \bar{\beta},\ \bar{f}' = \bar{g}]\!] := \{\, \alpha < \lambda \mid f_i(\alpha) = \beta_i\,,\ \text{for all } i < \gamma\,,\ \text{and} \\ f'_i(\alpha) = g_i(\alpha)\,,\ \text{for all } i < m \,\}\,.$$

For $m = 0$, we simply write $[\![\bar{f} = \bar{\beta}]\!]$ instead of $[\![\bar{f} = \bar{\beta},\ \langle\rangle = \langle\rangle]\!]$.

(c) Let $F \subseteq \mu^\lambda$ and $G \subseteq (<\mu)^\lambda$. An (F, G)-*condition* is a condition $[\![\bar{f} = \bar{\beta},\ \bar{f}' = \bar{g}]\!]$ with $\bar{f}, \bar{f}' \subseteq F$ and $\bar{g} \subseteq G$. A set $P \subseteq \mathscr{P}(\lambda)$ is (F, G)-*consistent* if it is consistent with every (F, G)-condition. For $G = \varnothing$, we simply speak of F-*conditions* and F-*consistency.*

Exercise 2.1. Let $P \subseteq \wp(\lambda)$ be F-consistent. Prove that every function $f \in F$ is surjective.

Exercise 2.2. Let $P \subseteq \wp(\lambda)$ be F-consistent. Show that there is no set $C \subseteq \lambda$ such that P forces both C and $\lambda \setminus C$.

Example. Let $\lambda = \aleph_0$ and let P be the set of all cofinite subsets of λ. Then $\mu = \aleph_0$ and a condition $C = [\![\bar{f} = \bar{\beta},\ \bar{f}' = \bar{g}]\!]$ is consistent with P if, and only if, C is infinite. It follows that P is F-consistent, where F is the set of all functions $f : \aleph_0 \to \aleph_0$ such that $f^{-1}(n)$ is infinite, for every $n < \aleph_0$.

Lemma 2.3. *Let $F \subseteq \mu^\lambda$ and $G \subseteq (<\mu)^\lambda$.*

(a) *If A and B are (F, G)-conditions, then $A \cap B$ is also an (F, G)-condition.*

(b) *If $(A_i)_{i<\gamma}$ is a sequence of F-conditions of length $\gamma < \mu$, then the intersection $\bigcap_{i<\gamma} A_i$ is also an F-condition.*

Proof. (a) Suppose that

$$A = [\![\bar{f}_0 = \bar{\beta}_0,\ \bar{f}'_0 = \bar{g}_0]\!] \quad \text{and} \quad B = [\![\bar{f}_1 = \bar{\beta}_1,\ \bar{f}'_1 = \bar{g}_1]\!]\,.$$

Then $A \cap B = [\![\bar{f}_0\bar{f}_1 = \bar{\beta}_0\bar{\beta}_1,\ \bar{f}'_0\bar{f}'_1 = \bar{g}_0\bar{g}_1]\!]$.

(b) follows as in (a) since F-conditions are closed under concatenations of length $\gamma < \mu$. □

Lemma 2.4. *Let I be a directed set and, for $i \in I$, let $P_i \subseteq \wp(\lambda)$, $F_i \subseteq \mu^\lambda$, and $G_i \subseteq (<\mu)^\lambda$ be sets such that $i \leq k$ implies $P_i \subseteq P_k$, $F_i \supseteq F_k$, and $G_i \subseteq G_k$. If P_i is (F_i, G_i)-consistent, for every $i \in I$, then $\bigcup_{i\in I} P_i$ is $(\bigcap_{i\in I} F_i, \bigcup_{i\in I} G_i)$-consistent.*

Proof. Let $C = [\![\bar{f} = \bar{\beta},\ \bar{f}' = \bar{g}]\!]$ be a $(\bigcap_i F_i, \bigcup_i G_i)$-condition. For a contradiction, suppose that $\bigcup_i P_i$ forces $\lambda \setminus C$. Then there exists a finite subset $Q \subseteq \bigcup_i P_i$ such that $\bigcap Q \cap C = \emptyset$. As I is directed, we can fix an index $k \in I$ such that $Q \subseteq P_k$.

Since $\bar{g}$ is a finite tuple, there exists an index $l \in I$ such that $\bar{g} \subseteq G_l$. Consequently, C is an (F_i, G_i)-condition, for all $i \geq l$. Fix $i \in I$ with

$i \geq k, l$. Since $Q \subseteq P_i$, it follows that P_i forces $\lambda \setminus C$. Hence, P_i is not (F_i, G_i)-consistent. A contradiction. □

In the following sequence of lemmas, we will construct larger and larger sets $P \subseteq \wp(\lambda)$ that are (F, G)-consistent, for sufficiently large sets F and G, until we obtain a set P that decides every subset of λ.

Lemma 2.5. *There exists a set $F \subseteq \mu^\lambda$ of size $|F| = 2^\lambda$ such that $\{\lambda\}$ is F-consistent.*

Proof. Let H be the set of all pairs $\langle A, h \rangle$ such that $A \subseteq \lambda$ is a set of size $|A| < \mu$ and $h : S \to \mu$ is a function with domain $S \subseteq \wp(A)$ of size $|S| < \mu$.

Let us first show that $|H| = \lambda$. There are $\lambda^{<\mu} = \lambda$ sets $A \subseteq \lambda$ of size $|A| < \mu$. For each such A, the number of sets $S \subseteq \wp(A)$ of size $|S| < \mu$ is at most

$$(2^{|A|})^{<\mu} \leq (\lambda^{|A|})^{<\mu} = \lambda^{<\mu} = \lambda\,.$$

For each set S, there are $\mu^{|S|} \leq \lambda^{|S|} \leq \lambda^{<\mu} = \lambda$ functions $S \to \mu$. Therefore, $|H| \leq \lambda \otimes \lambda \otimes \lambda = \lambda$. As it is easy to find λ different elements of H, it follows that $|H| = \lambda$.

Fix an enumeration $\langle A_\alpha, h_\alpha \rangle_{\alpha<\lambda}$ of H. For $C \subseteq \lambda$, we define a function $f_C : \lambda \to \mu$ by

$$f_C(\alpha) := \begin{cases} h_\alpha(C \cap A_\alpha) & \text{if } C \cap A_\alpha \in \operatorname{dom} h_\alpha\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

We claim that $F := \{ f_C \mid C \subseteq \lambda \}$ has the desired properties.

To show that $\{\lambda\}$ is F-consistent, consider an F-condition $[\![\bar f = \bar\beta]\!]$ where the sequences $\bar f$ and $\bar\beta$ have length $\gamma < \mu$. Since λ is the only set forced by $\{\lambda\}$, it is sufficient to show that $[\![\bar f = \bar\beta]\!] \neq \varnothing$.

Let $C_i \subseteq \lambda$ be the set such that $f_i = f_{C_i}$, for $i < \gamma$. W.l.o.g. we may assume that $f_i \neq f_k$, for $i \neq k$. Then $C_i \neq C_k$, for $i \neq k$. Hence, there is a set $A \subseteq \lambda$ of size $|A| = |\gamma|$ such that $i \neq k$ implies $C_i \cap A \neq C_k \cap A$.

Set $S := \{ C_i \cap A \mid i < \gamma \}$ and define $h : S \to \mu$ by

$$h(C_i \cap A) := \beta_i .$$

Then $\langle A, h \rangle \in H$. Hence, there is some $\alpha < \lambda$ such that $\langle A, h \rangle = \langle A_\alpha, h_\alpha \rangle$. For each $i < \gamma$, it follows that

$$f_i(\alpha) = f_{C_i}(\alpha) = h_\alpha(C_i \cap A_\alpha) = h(C_i \cap A) = \beta_i .$$

Therefore, $\alpha \in [\![\bar{f} = \bar{\beta}]\!] \neq \emptyset$. □

Lemma 2.6. *Suppose that $P \subseteq \wp(\lambda)$ is F-consistent. For every set $G \subseteq ({<}\mu)^\lambda$, there exists a set $F_o \subseteq F$ of size $|F_o| \leq |G| \otimes |P| \otimes \mu$ such that P is $(F \setminus F_o, G)$-consistent.*

Proof. We shall prove that, for every finite set $G_o \subseteq G$, there is some set $F(G_o) \subseteq F$ of size $|F(G_o)| \leq |P| \oplus \mu$ such that P is $(F \setminus F(G_o), G_o)$-consistent. By Lemma 2.4, it then follows that P is $(F \setminus F_o, G)$-consistent, where

$$F_o := \bigcup \{ F(G_o) \mid G_o \subseteq G \text{ finite} \}$$

has size $|F_o| \leq |G| \otimes \aleph_o \otimes |P| \otimes \mu$.

Fix a finite tuple $\bar{g} \in G^m$, $m < \omega$. By induction on α, we define a sequence of tuples $\bar{f}'_\alpha \in F^m$ as follows. Suppose we have already defined $\bar{f}'_i$, for $i < \alpha$. Set $F_\alpha := \bigcup_{i<\alpha} \bar{f}'_i$. If P is $(F \setminus F_\alpha, \bar{g})$-consistent, we stop. Otherwise, there is some $(F \setminus F_\alpha, \bar{g})$-condition $[\![\bar{f} = \bar{\beta},\ \bar{f}' = \bar{g}]\!]$ that is not consistent with P. We set $\bar{f}'_\alpha := \bar{f}'$.

Let $(\bar{f}'_\alpha)_{\alpha<\gamma}$ be the sequence constructed in this way. Obviously, we have $\gamma < |F|^+$. If $\gamma < \kappa := (|P| \oplus \mu)^+$, we can obtain the desired set as $F(\bar{g}) := \bigcup_{\alpha<\gamma} \bar{f}'_\alpha$.

Hence, assume that $\gamma \geq \kappa$. We will derive a contradiction as follows. For each $\alpha < \kappa$, fix a $(F \setminus F_\alpha, \bar{g})$-condition

$$A_\alpha := [\![\bar{f}_\alpha = \bar{\beta}_\alpha,\ \bar{f}'_\alpha = \bar{g}]\!]$$

such that P forces $\lambda \setminus A_\alpha$. Let P^+ be the closure of P under finite intersections. There are sets $S_\alpha \in P^+$ such that $S_\alpha \cap A_\alpha = \varnothing$. Since $|P^+| \leq |P| \oplus \aleph_0 < \kappa$, we can find a set $I \subseteq \kappa$ of size $|I| = \kappa$ such that

$$S_\alpha = S_{\alpha'}\,, \quad \text{for all } \alpha, \alpha' \in I\,.$$

Let S be the set such that $S = S_\alpha$, for $\alpha \in I$. Since each sequence $\bar{f}_\alpha$ has length less than $\mu < \kappa$, there is a subset $J \subseteq I$ of size $|J| = \kappa$ such that $|\bar{f}_\alpha| = |\bar{f}_{\alpha'}|$, for all $\alpha, \alpha' \in J$.

Set

$$\chi := \sup\big\{\, |g_i(\alpha)|^+ \;\big|\; i < m,\ \alpha < \lambda \,\big\}$$

and let $(\bar{\gamma}_\alpha)_{\alpha<\chi}$ be an enumeration of χ^m. Note that $\chi < \mu$ since $\operatorname{rng} g_i \subseteq \nu_i$, for some $\nu_i < \mu$. Hence,

$$g_i(\alpha) < \nu_i < \mu \quad \text{implies} \quad |g_i(\alpha)|^+ \leq \nu_i < \mu\,.$$

Fix an injective function $h : \chi \to J$ and set

$$A := \bigcap_{i<\chi} [\![\bar{f}_{h(i)} = \bar{\beta}_{h(i)},\ \bar{f}'_{h(i)} = \bar{\gamma}_i]\!]\,.$$

Since $\chi < \mu$ it follows by Lemma 2.3 (b) that A is an F-condition. Hence, the F-consistency of P implies that P does not force $\lambda \setminus A$.

Consequently, $S \cap A \neq \varnothing$ and we can find some $\alpha \in S \cap A$. It follows that

$$\bar{f}_{h(i)}(\alpha) = \bar{\beta}_{h(i)}(\alpha) \quad \text{and} \quad \bar{f}'_{h(i)}(\alpha) = \bar{\gamma}_i(\alpha)\,, \quad \text{for all } i < \chi\,.$$

Fix $i < \chi$ such that $\bar{\gamma}_i = \bar{g}(\alpha)$. Then $\alpha \in A_{h(i)}$. Hence, $S_{h(i)} \cap A_{h(i)} = \varnothing$ implies that $\alpha \notin S_{h(i)} = S$. A contradiction. □

To extend the set P to an ultrafilter, we can use the following lemma and its corollary to ensure that P decides every set.

Lemma 2.7. *Let $P \subseteq \wp(\lambda)$ be (F, G)-consistent. For every set $A \subseteq \lambda$ there is some $F_0 \subseteq F$ of size $|F_0| < \mu$ such that at least one of $P \cup \{A\}$ and $P \cup \{\lambda \setminus A\}$ is $(F \setminus F_0, G)$-consistent.*

Proof. Suppose that $P \cup \{A\}$ is not (F, G)-consistent. Then there is an (F, G)-condition $C_0 := [\![\bar{f}_0 = \bar{\beta}_0,\ \bar{f}'_0 = \bar{g}_0]\!]$ such that $P \cup \{A\}$ forces $\lambda \setminus C_0$. Hence, there is some $S_0 \in \mathrm{cl}_\uparrow(P)$ such that

$$S_0 \cap A \cap C_0 = \varnothing\,.$$

Set $F_0 := \bar{f}_0 \cup \bar{f}'_0$. If $P \cup \{\lambda \setminus A\}$ is $(F \setminus F_0, G)$-consistent, we are done.

Hence, we may assume that this set is not $(F \setminus F_0, G)$-consistent. Then there is an $(F \setminus F_0, G)$-condition $C_1 := [\![\bar{f}_1 = \bar{\beta}_1,\ \bar{f}'_1 = \bar{g}_1]\!]$ such that $P \cup \{\lambda \setminus A\}$ forces $\lambda \setminus C_1$. Hence, there is some set $S_1 \in \mathrm{cl}_\uparrow(P)$ such that

$$S_1 \cap (\lambda \setminus A) \cap C_1 = \varnothing\,.$$

It follows that $S_1 \cap C_1 \subseteq A$, which implies that

$$S_0 \cap S_1 \cap C_0 \cap C_1 \subseteq S_0 \cap C_0 \cap A = \varnothing\,.$$

As $S_0 \cap S_1 \in \mathrm{cl}_\uparrow(P)$, it follows that P forces $\lambda \setminus (C_0 \cap C_1)$. Since $C_0 \cap C_1$ is an (F, G)-condition, P is not (F, G)-consistent. A contradiction. □

Repeating this lemma for each set $A \in H$, we obtain the following statement.

Corollary 2.8. *Let $P \subseteq \wp(\lambda)$ be (F, G)-consistent. For every set $H \subseteq \wp(\lambda)$ there is some $F_0 \subseteq F$ of size $|F_0| \leq |H| \otimes \mu$ and some $Q \subseteq \wp(\lambda)$ of size $|Q| = |H|$ such that $P \cup Q$ is $(F \setminus F_0, G)$-consistent and it decides every set $A \in H$.*

To prove the Theorem of Keisler and Shelah below, we will have to show that $\mathfrak{A}^{\mathfrak{u}} \cong \mathfrak{B}^{\mathfrak{u}}$, for certain structures $\mathfrak{A}$ and $\mathfrak{B}$. This is done via a back-and-forth argument where we construct an increasing chain of partial isomorphisms between the structures $\mathfrak{A}^{\mathfrak{u}}$ and $\mathfrak{B}^{\mathfrak{u}}$. Matters become slightly more complicated since we construct the ultrafilter $\mathfrak{u}$ at the same time. Hence, we do not yet know between which structures we should eventually construct partial isomorphisms. Therefore, we introduce a notion of a partial isomorphism between partially defined ultrapowers.

Definition 2.9. Let $\mathfrak{A}$ and $\mathfrak{B}$ be Σ-structures and let $P \subseteq \wp(\lambda)$ be a set with the finite intersection property. A partial function π from A^λ to B^λ is a *partial isomorphism modulo* P if, for every formula $\varphi(\bar{x}) \in \mathrm{FO}^{<\omega}[\Sigma]$ and every finite mapping $\bar{a} \mapsto \bar{b} \subseteq \pi$,

$$P \quad \text{forces} \quad \{\, k < \lambda \mid \mathfrak{A} \vDash \varphi(\bar{a}(k)) \Leftrightarrow \mathfrak{B} \vDash \varphi(\bar{b}(k)) \,\},$$

$$\text{and} \quad P \quad \text{decides} \quad [\![\mathfrak{A} \vDash \varphi(\bar{a}(k))]\!]_{k<\lambda}\,.$$

Exercise 2.3. Show that every partial isomorphism π from A^λ to B^λ modulo an ultrafilter $\mathfrak{u}$ induces an ordinary partial isomorphism between $\mathfrak{A}^{\mathfrak{u}}$ and $\mathfrak{B}^{\mathfrak{u}}$.

The back-and-forth step of the construction below is contained in the following two lemmas. The first one is a technical result which, intuitively, states that we can realise every partial type.

Lemma 2.10. *Let P be F-consistent, let $\mathfrak{M}$ be a Σ-structure of size $\kappa := |M| < \mu$, and let $\Phi \subseteq \mathrm{FO}^1[\Sigma_{M^\lambda}]$ be a set of first-order formulae over M^λ that is closed under conjunctions.*

If, for every $\varphi(x; \bar{a}) \in \Phi$,

$$P \quad \textit{forces} \quad [\![\mathfrak{M} \vDash \exists x \varphi(x; \bar{a}(\alpha))]\!]_{\alpha<\lambda}\,,$$

there exist a sequence $b \in M^\lambda$ and sets $F_0 \subseteq F$ and $Q \subseteq \wp(\lambda)$ of size

$$|F_0| \leq |P| \oplus |\Phi| \oplus \mu \quad \textit{and} \quad |Q| \leq |\Phi|$$

such that $P \cup Q$ is $(F \setminus F_0)$-consistent and, for every $\varphi(x; \bar{a}) \in \Phi$,

$$P \cup Q \quad \textit{forces} \quad [\![\mathfrak{M} \vDash \varphi(b(\alpha); \bar{a}(\alpha))]\!]_{\alpha<\lambda}\,.$$

Proof. Fix enumerations $(c_i)_{i<\kappa}$ of M and $(\varphi_l(x;\bar{a}_l))_{l<\chi}$ of Φ. For each $l<\chi$, we fix a function $g_l:\lambda\to\kappa$ such that

$$\mathfrak{M}\vDash\exists x\varphi_l(x,\bar{a}_l(\alpha))\quad\text{implies}\quad\mathfrak{M}\vDash\varphi_l(c_{g_l(\alpha)},\bar{a}_l(\alpha))\,.$$

Set $G:=\{\,g_l\mid l<\chi\,\}$. By Lemma 2.6, there is a set $F_1\subseteq F$ of size $|F_1|\leq|P|\oplus\chi\oplus\mu$ such that P is $(F\smallsetminus F_1,G)$-consistent. Fix some $f\in F\smallsetminus F_1$ and set

$$\begin{aligned}F_0&:=F_1\cup\{f\}\,,\\ b(\alpha)&:=\begin{cases}c_{f(\alpha)} & \text{if } f(\alpha)<\kappa\,,\\ c_0 & \text{otherwise}\,,\end{cases}\\ Q&:=\big\{\,[\![\mathfrak{M}\vDash\varphi_l(b(\alpha),\bar{a}_l(\alpha))]\!]_{\alpha<\lambda}\;\big|\;l<\chi\,\big\}\,.\end{aligned}$$

We claim that F_0, Q, and b have the desired properties.

Since

$$[\![\mathfrak{M}\vDash\varphi_l(b(\alpha),\bar{a}_l(\alpha))]\!]_{\alpha<\lambda}\in Q\subseteq\mathrm{cl}_\uparrow(P\cup Q)\,,\qquad\text{for all } l<\chi\,,$$

it remains to show that $P\cup Q$ is $(F\smallsetminus F_0)$-consistent. For a contradiction, suppose otherwise. Then we can find an $(F\smallsetminus F_0)$-condition $C:=[\![\bar{f}=\bar{\beta}]\!]$ such that $P\cup Q$ forces $\lambda\smallsetminus C$. Since Φ is closed under conjunctions, the set Q is closed under finite intersections. Therefore, there are sets $S\in\mathrm{cl}_\uparrow(P)$ and $T\in Q$ such that

$$S\cap T\cap C=\varnothing\,.$$

Let $l<\chi$ be the index such that $T=[\![\mathfrak{M}\vDash\varphi_l(b(\alpha),\bar{a}_l(\alpha))]\!]_{\alpha<\lambda}$. Then

$$S\cap[\![\mathfrak{M}\vDash\varphi_l(b(\alpha),\bar{a}_l(\alpha))]\!]_{\alpha<\lambda}\cap[\![\bar{f}=\bar{\beta}]\!]=\varnothing\,.$$

By choice of g_l, we have

$$\begin{aligned}&[\![\mathfrak{M}\vDash\exists x\varphi_l(x,\bar{a}_l(\alpha))]\!]_{\alpha<\lambda}\cap[\![\langle\rangle=\langle\rangle,\ f=g_l]\!]\\ &\subseteq[\![\mathfrak{M}\vDash\varphi_l(b(\alpha),\bar{a}_l(\alpha))]\!]_{\alpha<\lambda}\,.\end{aligned}$$

Therefore, it follows that

$$S \cap [\![\mathfrak{M} \vDash \exists x \varphi_l(x, \bar{a}_l(\alpha))]\!]_{\alpha<\lambda} \cap [\![\bar{f} = \bar{\beta},\ f = g_l]\!] = \varnothing\,.$$

Hence, P forces $[\![\bar{f} = \bar{\beta},\ f = g_l]\!]$ in contradiction to the $(F \smallsetminus F_1, G)$-consistency of P. □

Lemma 2.11. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be Σ-structures with $|\Sigma| \leq \lambda$, $P \subseteq \wp(\lambda)$ a set that is F-consistent, and π a partial isomorphism from A^λ to B^λ modulo P. For every element $c \in A^\lambda$, there exist an element $d \in B^\lambda$ and sets $Q \subseteq \wp(\lambda)$ and $F_0 \subseteq F$ of size*

$$|Q| \leq |\pi| \oplus \lambda \quad \textit{and} \quad |F_0| \leq |P| \oplus |\pi| \oplus \lambda$$

such that $P \cup Q$ is $(F \smallsetminus F_0)$-consistent and $\pi \cup \{\langle c, d\rangle\}$ is a partial isomorphism modulo $P \cup Q$.

Proof. Note that there are $|\pi|^{<\omega} = |\pi| \oplus \aleph_0$ finite tuples $\bar{a} \subseteq \mathrm{dom}(\pi)$ and there are at most λ formulae $\varphi \in \mathrm{FO}^{<\omega}[\Sigma]$. Hence, we can use Corollary 2.8 to find sets Q_1 and $F_1 \subseteq F$ of size

$$|Q_1| \leq \lambda \oplus |\pi| \quad \text{and} \quad |F_1| \leq \lambda \oplus |\pi| \oplus \mu = \lambda \oplus |\pi|$$

such that $P \cup Q_1$ is $(F \smallsetminus F_1)$-consistent and

$$P \cup Q_1 \quad \text{decides} \quad [\![\mathfrak{A} \vDash \varphi(\bar{a}(\alpha), c(\alpha))]\!]_{\alpha<\lambda}\,,$$

for all $\varphi(\bar{x}, y) \in \mathrm{FO}^{<\omega}[\Sigma]$ and all finite $\bar{a} \subseteq \mathrm{dom}(\pi)$.

Suppose that $\pi = \bar{a} \mapsto \bar{b}$ and set

$$\Phi := \big\{\, \varphi(\bar{x}, y) \bigm| P \cup Q_1 \text{ forces } [\![\mathfrak{A} \vDash \varphi(\bar{a}(\alpha), c(\alpha))]\!]_{\alpha<\lambda} \,\big\}\,.$$

Note that $\varphi \notin \Phi$ implies $\neg\varphi \in \Phi$, by construction of Q_1. Since π is a partial isomorphism modulo P, it follows for $\varphi \in \Phi$ that

$$\begin{aligned}
&[\![\mathfrak{B} \vDash \exists y\varphi(\bar{b}(\alpha), y)]\!]_{\alpha<\lambda} \\
\supseteq\ &[\![\mathfrak{A} \vDash \exists y\varphi(\bar{a}(\alpha), y)]\!]_{\alpha<\lambda} \\
&\cap \big\{ \alpha < \lambda \bigm| \mathfrak{A} \vDash \exists y\varphi(\bar{a}(\alpha), y) \Leftrightarrow \mathfrak{B} \vDash \exists y\varphi(\bar{b}(\alpha), y) \big\} \\
\supseteq\ &[\![\mathfrak{A} \vDash \varphi(\bar{a}(\alpha), c(\alpha))]\!]_{\alpha<\lambda} \\
&\cap \big\{ \alpha < \lambda \bigm| \mathfrak{A} \vDash \exists y\varphi(\bar{a}(\alpha), y) \Leftrightarrow \mathfrak{B} \vDash \exists y\varphi(\bar{b}(\alpha), y) \big\} \\
\in\ &\mathrm{cl}_{\uparrow}(P \cup Q_1)\,.
\end{aligned}$$

Hence, $P \cup Q_1$ forces $[\![\mathfrak{B} \vDash \exists y\varphi(\bar{b}(\alpha), y)]\!]_{\alpha<\lambda}$, for all $\varphi \in \Phi$, and we can use Lemma 2.10 to find an element $d \in B^\lambda$ and sets Q_2 and $F_2 \subseteq F \setminus F_1$ of size

$$|Q_2| \leq |\Phi| = |\Sigma| \oplus |\pi| \oplus \aleph_0 = \lambda \oplus |\pi|$$

$$\text{and} \quad |F_2| \leq |P \cup Q_1| \oplus |\Phi| \oplus \mu = |P| \oplus |\pi| \oplus \lambda$$

such that $P \cup Q_1 \cup Q_2$ is $(F \setminus (F_1 \cup F_2))$-consistent and

$$P \cup Q_1 \cup Q_2 \quad \text{forces} \quad [\![\mathfrak{B} \vDash \varphi(\bar{b}(\alpha), d(\alpha))]\!]_{\alpha<\lambda}\,,$$

for all $\varphi \in \Phi$.

We claim that the extension $\pi \cup \{\langle c, d\rangle\}$ is a partial isomorphism modulo $P \cup Q_1 \cup Q_2$. We have already seen above that $P \cup Q_1$, and hence also $P \cup Q_1 \cup Q_2$, decides every set of the form $[\![\mathfrak{A} \vDash \varphi(\bar{a}(\alpha), c(\alpha))]\!]_{\alpha<\lambda}$ with $\bar{a} \subseteq \mathrm{dom}(\pi)$. To check the remaining condition, we distinguish two cases.

If $\varphi \in \Phi$, the fact that

$$[\![\mathfrak{A} \vDash \varphi(\bar{a}(\alpha), c(\alpha))]\!]_{\alpha<\lambda} \quad \text{and} \quad [\![\mathfrak{B} \vDash \varphi(\bar{b}(\alpha), d(\alpha))]\!]_{\alpha<\lambda}$$

are in $\mathrm{cl}_\uparrow(P \cup Q_1 \cup Q_2)$ implies that

$$\begin{aligned}
&\{\, \alpha < \lambda \mid \mathfrak{A} \vDash \varphi(\bar a(\alpha), c(\alpha)) \Leftrightarrow \mathfrak{B} \vDash \varphi(\bar b(\alpha), d(\alpha)) \,\}\\
\supseteq &\{\, \alpha < \lambda \mid \mathfrak{A} \vDash \varphi(\bar a(\alpha), c(\alpha)) \text{ and } \mathfrak{B} \vDash \varphi(\bar b(\alpha), d(\alpha)) \,\}\\
= &[\![\mathfrak{A} \vDash \varphi(\bar a(\alpha), c(\alpha))]\!]_{\alpha<\lambda} \cap [\![\mathfrak{B} \vDash \varphi(\bar b(\alpha), d(\alpha))]\!]_{\alpha<\lambda}\\
\in &\ \mathrm{cl}_\uparrow(P \cup Q_1 \cup Q_2)\,.
\end{aligned}$$

If $\varphi \notin \Phi$, we have noted above that $\neg\varphi \in \Phi$. Therefore,

$$\begin{aligned}
&\{\, \alpha < \lambda \mid \mathfrak{A} \vDash \varphi(\bar a(\alpha), c(\alpha)) \Leftrightarrow \mathfrak{B} \vDash \varphi(\bar b(\alpha), d(\alpha)) \,\}\\
= &\{\, \alpha < \lambda \mid \mathfrak{A} \vDash \neg\varphi(\bar a(\alpha), c(\alpha)) \Leftrightarrow \mathfrak{B} \vDash \neg\varphi(\bar b(\alpha), d(\alpha)) \,\}\\
\in &\ \mathrm{cl}_\uparrow(P \cup Q_1 \cup Q_2)\,.
\end{aligned}$$

Consequently, $P \cup Q_1 \cup Q_2$ forces

$$\{\, \alpha < \lambda \mid \mathfrak{A} \vDash \neg\varphi(\bar a(\alpha), c(\alpha)) \Leftrightarrow \mathfrak{B} \vDash \neg\varphi(\bar b(\alpha), d(\alpha)) \,\}$$

for all formulae φ. □

Theorem 2.12 (Keisler, Shelah). *Let λ be an infinite cardinal and let μ be the least cardinal such that $2^\mu > \lambda$. There exists an ultrafilter $\mathfrak{u}$ on λ such that*

$$\mathfrak{A} \equiv \mathfrak{B} \quad \textit{implies} \quad \mathfrak{A}^{\mathfrak{u}} \cong \mathfrak{B}^{\mathfrak{u}}\,,$$

for all structures $\mathfrak{A}$ and $\mathfrak{B}$ of size $|A|, |B| < \mu$.

Proof. Note that every Σ-structure $\mathfrak{M}$ of size $\kappa := |M| < \mu$ is interdefinable with a reduct $\mathfrak{M}|_{\Sigma_0}$ for some $\Sigma_0 \subseteq \Sigma$ of size $|\Sigma_0| \leq 2^\kappa \leq \lambda$ since there are only 2^κ distinct relations and functions on M. We may therefore w.l.o.g. assume that the signature Σ of every structure is contained in a fixed signature Σ_+ of size λ consisting, for all finite sequences $\bar s t$ of sorts, of λ relation symbols of type $\bar s$ and λ function symbols of type $\bar s \to t$. Furthermore, we may assume that all structures have universe κ, for some cardinal $\kappa < \mu$. Note that, by Lemma B1.1.5, there are, up to isomorphism, at most $2^{|\Sigma| \oplus \kappa} \leq 2^\lambda$ such Σ-structures.

Therefore, we can fix an enumeration $\langle \mathfrak{A}_i, \mathfrak{B}_i \rangle_{i<2^\lambda}$ of all pairs of Σ_i-structures with $\Sigma_i \subseteq \Sigma_+$ where the universe is some cardinal less than μ and such that $\mathfrak{A}_i \equiv \mathfrak{B}_i$. We also fix a surjective function

$$R : 2^\lambda \to [3] \times 2^\lambda \times 2^\lambda$$

and enumerations $(u_\alpha)_{\alpha<2^\lambda}$ of μ^λ and $(S_\alpha)_{\alpha<2^\lambda}$ of $\mathscr{P}(\lambda)$.

We will construct an ultrafilter $\mathfrak{u}$ such that $\mathfrak{A}_i^\mathfrak{u} \cong \mathfrak{B}_i^\mathfrak{u}$, for all i. By induction on $\gamma < 2^\lambda$, we construct

- an increasing sequence $(P_\gamma)_{\gamma<2^\lambda}$ of sets $P_\gamma \subseteq \mathscr{P}(\lambda)$,
- a decreasing sequence $(F_\gamma)_{\gamma<2^\alpha}$ of sets $F_\gamma \subseteq \mu^\lambda$, and
- for each $i < 2^\lambda$, an increasing sequence $(\pi^i_\gamma)_{\gamma<2^\lambda}$ of partial functions π^i_γ from $A_i^\lambda \subseteq (<\mu)^\lambda$ to $B_i^\lambda \subseteq (<\mu)^\lambda$

satisfying the following conditions:

(1) P_γ is F_γ-consistent;

(2) each π^i_γ is a partial isomorphism from A_i^λ to B_i^λ modulo P_γ;

(3)
$$\begin{aligned} \big|\textstyle\bigcup_{i<2^\lambda} \operatorname{dom}(\pi^i_\gamma)\big| &\leq |\gamma|\,, \\ |P_\gamma| &\leq \lambda \oplus |\gamma|\,, \\ |F_0| &= 2^\lambda\,, \\ |F_0 \setminus F_\gamma| &\leq \lambda \oplus |\gamma|\,; \end{aligned}$$

(4) if $R(\gamma) = \langle 0, i, \alpha \rangle$ and $u_\alpha \in A_i^\lambda$, then $u_\alpha \in \operatorname{dom}(\pi^i_{\gamma+1})$;

(5) if $R(\gamma) = \langle 1, i, \beta \rangle$ and $u_\beta \in B_i^\lambda$, then $u_\beta \in \operatorname{rng}(\pi^i_{\gamma+1})$;

(6) if $R(\gamma) = \langle 2, \alpha, \beta \rangle$, then $P_{\gamma+1}$ decides S_α.

First, let us show that, after having performed this construction, the limit $\mathfrak{u} := \bigcup_{\gamma<2^\lambda} P_\gamma$ is the desired ultrafilter. By (6) and the surjectivity of R, $\mathfrak{u}$ is an ultrafilter. Furthermore, by (2) $\pi^i := \bigcup_{\gamma<2^\lambda} \pi^i_\gamma$ is a partial isomorphism between $\mathfrak{A}_i^{\mathfrak{u}}$ and $\mathfrak{B}_i^{\mathfrak{u}}$. Finally, by (4), (5), and the surjectivity of R, π^i is bijective.

Hence, it remains to do the induction. We start with $P_0 := \{\lambda\}$ and $\pi_0^i := \langle\rangle \mapsto \langle\rangle$, for all $i < 2^\lambda$. According to Lemma 2.5, there exists a set F_0 of size $|F_0| = 2^\lambda$ such that P_0 is F_0-consistent. Note that Condition (2) is satisfied, since $\mathfrak{A}_i \equiv \mathfrak{B}_i$, while all other conditions are satisfied trivially.

For limit ordinals δ, we set

$$P_\delta := \bigcup_{\gamma<\delta} P_\gamma\,, \quad F_\delta := \bigcap_{\gamma<\delta} F_\gamma\,, \quad \text{and} \quad \pi_\delta^i := \bigcup_{\gamma<\delta} \pi_\gamma^i\,.$$

Then Condition (1) follows by Lemma 2.4, while Conditions (2)–(6) follow immediately from the inductive hypothesis.

For the successor step, suppose that we have already defined P_γ, F_γ, and π_γ^i. Depending on the value of $R(\gamma)$, we distinguish three cases. First, suppose that $R(\gamma) = \langle 0, i, \alpha\rangle$, for some $i, \alpha < 2^\lambda$. If $u_\alpha \notin A_i^\alpha$, we simply set $P_{\gamma+1} := P_\gamma$, $F_{\gamma+1} := F_\gamma$, and $\pi_{\gamma+1}^k := \pi_\gamma^k$, for all k. Hence, suppose that $u_\alpha \in A_i^\lambda$. By Lemma 2.11, there exist an element $v \in B_i^\lambda$ and sets Q' and $F' \subseteq F_\gamma$ of size

$$|Q'|, |F'| \leq \lambda \oplus |\gamma|$$

such that $P \cup Q'$ is $(F_\gamma \setminus F')$-consistent and $\pi_\gamma^i \cup \{\langle u_\alpha, v\rangle\}$ is a partial isomorphism modulo $P \cup Q'$.

We set $P_{\gamma+1} := P_\gamma \cup Q'$, $F_{\gamma+1} := F_\gamma \setminus F'$, and $\pi_{\gamma+1}^i := \pi_\gamma^i \cup \{\langle u_\alpha, v\rangle\}$. By construction, $\pi_{\gamma+1}^i$ satisfies Conditions (1), (2), and (4). Conditions (3), (5), and (6) are also satisfied.

If $R(\gamma) = \langle 1, i, \beta\rangle$, for some $i, \beta < 2^\lambda$, we proceed analogously to the first case applying Lemma 2.11 to $(\pi_\gamma^i)^{-1}$.

Finally, if $R(\gamma) = \langle 2, \alpha, \beta\rangle$, we use Corollary 2.8 to find sets Q' and $F' \subseteq F_\gamma$ of size $|Q'| = 1$ and $|F'| \leq \mu \leq \lambda$ such that $P_{\gamma+1} := P_\gamma \cup Q'$ is $(F \setminus F')$-consistent and decides S_α. We set $F_{\gamma+1} := F_\gamma \setminus F'$ and $\pi_{\gamma+1}^i := \pi_\gamma^i$, for all i. □

Corollary 2.13. *Let* $\mathfrak{A}$ *and* $\mathfrak{B}$ *be* Σ*-structures. We have*

$$\mathfrak{A} \equiv \mathfrak{B} \quad \text{iff} \quad \mathfrak{A}^{\mathfrak{u}} \cong \mathfrak{B}^{\mathfrak{u}} \quad \textit{for some ultrafilter } \mathfrak{u}\,.$$

The Theorem of Keisler and Shelah can be used to characterise first-order axiomatisable classes via their closure properties.

Definition 2.14. We say that a class $\mathcal{K}$ is *closed under reverse ultrapowers* if $\mathfrak{A}^{\mathfrak{u}} \in \mathcal{K}$ implies $\mathfrak{A} \in \mathcal{K}$, for every structure $\mathfrak{A}$ and all ultrafilters $\mathfrak{u}$.

Theorem 2.15. *A class* $\mathcal{K}$ *of* Σ*-structures is first-order axiomatisable if and only if* $\mathcal{K}$ *is closed under isomorphisms, ultraproducts, and reverse ultrapowers.*

Proof. One direction follows immediately from Corollary 1.2. For the other one, let $\Phi := \mathrm{Th}(\mathcal{K})$. We claim that $\mathrm{Mod}(\Phi) = \mathcal{K}$. Suppose otherwise. Then there exists a model $\mathfrak{B} \vDash \Phi$ such that $\mathfrak{B} \notin \mathcal{K}$. If we can show that $T := \mathrm{Th}(\mathfrak{B})$ is an accumulation point of the set $X := \{\, \mathrm{Th}(\mathfrak{A}) \mid \mathfrak{A} \in \mathcal{K} \,\}$ then we can apply Lemma 1.9 to find an ultraproduct $\mathfrak{C} := \prod_{i \in I} \mathfrak{A}_i/\mathfrak{u}$ of structures $\mathfrak{A}_i \in \mathcal{K}$ such that $\mathrm{Th}(\mathfrak{C}) = T = \mathrm{Th}(\mathfrak{B})$. Hence, by Corollary 2.13, there exists an ultrafilter $\mathfrak{v}$ such that $\mathfrak{B}^{\mathfrak{v}} \cong \mathfrak{C}^{\mathfrak{v}}$. But $\mathfrak{C} \in \mathcal{K}$ implies $\mathfrak{C}^{\mathfrak{v}} \in \mathcal{K}$ while $\mathfrak{B} \notin \mathcal{K}$ implies $\mathfrak{B}^{\mathfrak{v}} \notin \mathcal{K}$. Contradiction.

It remains to show that T is an accumulation point of X. Let $T \in \langle \varphi \rangle$. Then $\neg\varphi \notin \Phi \subseteq T$ and, by definition of Φ, there exists some structure $\mathfrak{A} \in \mathcal{K}$ such that $\mathfrak{A} \nvDash \neg\varphi$. Hence, $\mathrm{Th}(\mathfrak{A}) \in \langle \varphi \rangle \cap X \neq \varnothing$. ☐

3. *Reduced products and Horn formulae*

In this section we study classes that are closed under arbitrary products and formulae that are preserved in products.

Definition 3.1. A formula φ is *preserved in reduced products* if, for every family $(\mathfrak{A}^i)_{i \in I}$ of structures and every filter $\mathfrak{u}$ over I, we have

$$\mathfrak{A}^i \vDash \varphi \text{ for all } i \quad \text{implies} \quad \prod_{i \in I} \mathfrak{A}^i/\mathfrak{u} \vDash \varphi\,.$$

If this holds only for $\mathfrak{u} = \{I\}$ then φ is *preserved in products.* Finally, we say that φ is *preserved in nonempty products* if the above is true only for $\mathfrak{u} = \{I\}$ and $I \neq \varnothing$.

Definition 3.2. (a) A *basic Horn formula* is a formula of the form

$$\varphi := \bigwedge \Phi \to \psi\,,$$

where ψ is an atomic formula or the formula false and Φ is a set (possibly empty) of atomic formulae. If $\psi \neq$ false then we say that φ is *strict.*

(b) A *Horn formula* is a formula of the form

$$\varphi = Q_0\bar{x}_0 \cdots Q_{n-1}\bar{x}_{n-1} \bigwedge \Phi$$

where Φ is a set of basic Horn formulae and the $Q_i \in \{\exists, \forall\}$ are quantifiers. We allow both Φ and the sequences $\bar{x}_i$ to be infinite. We call φ *strict* if Φ only contains strict basic Horn formulae. A Horn formula is *universal* if it is of the form $\forall\bar{x}\psi$ where ψ is a single basic Horn formula.

We denote the set of all Horn formulae by $\mathrm{HO}_\infty[\Sigma, X]$. $\mathrm{SH}_\infty[\Sigma, X]$ is the subset of all strict Horn formulae. The set of all universal (strict) Horn formulae is denoted by $\mathrm{H}\forall_\infty[\Sigma, X]$ and $\mathrm{SH}\forall_\infty[\Sigma, X]$, respectively. We write $\mathrm{HO}[\Sigma, X]$, $\mathrm{SH}[\Sigma, X]$, $\mathrm{H}\forall[\Sigma, X]$, and $\mathrm{SH}\forall[\Sigma, X]$, for the corresponding fragments of first-order logic.

(c) A formula is *positive primitive* if it is obtained from atomic formulae by (possibly infinite) conjunctions and existential quantifications. Again we allow quantifiers of the form $\exists\bar{x}$ where $\bar{x}$ is a possibly infinite sequence of variables.

Lemma 3.3. *Suppose that $\varphi(\bar{x})$ is a positive primitive formula, $(\mathfrak{A}^i)_{i\in I}$ a nonempty sequence of structures, and $\bar{a} \subseteq \prod_i A^i$. Then we have*

$$\textstyle\prod_i \mathfrak{A}^i \vDash \varphi(\bar{a}) \quad \text{iff} \quad \mathfrak{A}^i \vDash \varphi(\bar{a}^i)\,, \quad \textit{for all } i \in I\,.$$

Proof. W.l.o.g. we may assume that φ is term-reduced. We prove the claim by induction on φ. For atomic formulae φ, the claim holds by definition of a direct product. If φ is a conjunction then the claim follows immediately from the inductive hypothesis. Hence, we may assume that $\varphi(\bar{x}) = \exists \bar{y} \psi(\bar{x}, \bar{y})$.

If $\prod_i \mathfrak{A}^i \vDash \varphi(\bar{a})$ then there exists a sequence $\bar{b} \subseteq \prod_i A^i$ such that $\prod_i \mathfrak{A}^i \vDash \psi(\bar{a}, \bar{b})$. By inductive hypothesis, we therefore have

$$\mathfrak{A}^i \vDash \psi(\bar{a}^i, \bar{b}^i), \quad \text{for all } i,$$

and it follows that $\mathfrak{A}^i \vDash \exists \bar{y} \psi(\bar{a}^i, \bar{y})$.

Conversely, suppose that $\mathfrak{A}^i \vDash \varphi(\bar{a}^i)$, for all i. Choose sequences $\bar{b}^i \subseteq A^i$ such that $\mathfrak{A}^i \vDash \psi(\bar{a}^i, \bar{b}^i)$. By inductive hypothesis, it follows that $\prod_i \mathfrak{A}^i \vDash \psi(\bar{a}, \bar{b})$. This implies that $\prod_i \mathfrak{A}^i \vDash \varphi(\bar{a})$. □

Theorem 3.4. *Let $\varphi(\bar{x})$ be a Horn formula, $(\mathfrak{A}^i)_{i \in I}$ a nonempty sequence of structures, and $\bar{a} \subseteq \prod_i A^i$. Then*

$$\mathfrak{A}^i \vDash \varphi(\bar{a}^i), \textit{ for all } i, \quad \textit{implies} \quad \prod_i \mathfrak{A}^i \vDash \varphi(\bar{a}).$$

Proof. We prove the claim by induction on φ. Suppose that $\mathfrak{A}^i \vDash \varphi(\bar{a}^i)$, for all i. First, we consider the case that $\varphi = \bigwedge \Phi \to \psi$ is a basic Horn formula. If $\prod_i \mathfrak{A}^i \nvDash \Phi(\bar{a})$ then we are done. Hence we may assume that $\prod_i \mathfrak{A}^i \vDash \Phi(\bar{a})$. By Lemma 3.3, it follows that $\mathfrak{A}^i \vDash \Phi(\bar{a}^i)$, for all i. Since $\mathfrak{A}^i \vDash \varphi(\bar{a}^i)$ this implies that $\mathfrak{A}^i \vDash \psi(\bar{a}^i)$. In this case ψ cannot be false and we can use Lemma 3.3 to conclude that $\prod_i \mathfrak{A}^i \vDash \psi(\bar{a})$, as desired.

If φ is a conjunction then the claim follows immediately by inductive hypothesis. For $\varphi = \exists \bar{y} \psi(\bar{x}, \bar{y})$ we can argue in the same way as in the proof of Lemma 3.3. Finally, assume that $\varphi = \forall \bar{y} \psi(\bar{x}, \bar{y})$. Let $\bar{b} \subseteq \prod_i A^i$. Since $\mathfrak{A}^i \vDash \varphi(\bar{a}^i)$ we have $\mathfrak{A}^i \vDash \psi(\bar{a}^i, \bar{b}^i)$. By inductive hypothesis, this implies that $\prod_i \mathfrak{A}^i \vDash \psi(\bar{a}, \bar{b})$. Since $\bar{b}$ was arbitrary it follows that $\prod_i \mathfrak{A}^i \vDash \forall \bar{y} \psi(\bar{a}, \bar{y})$. □

For first-order formulae we can generalise these results to reduced products.

Lemma 3.5. *Suppose that $\varphi(\bar{x})$ is a positive primitive first-order formula, $(\mathfrak{A}^i)_{i \in I}$ a nonempty sequence of structures, $\mathfrak{u}$ a filter over I, and $[\bar{a}]$ a tuple* *in* $\prod_i \mathfrak{A}^i/\mathfrak{u}$. *Then we have*

$$\prod_i \mathfrak{A}^i/\mathfrak{u} \vDash \varphi([\bar{a}]) \quad \text{iff} \quad [\![\varphi(\bar{a})]\!] \in \mathfrak{u}\,.$$

Proof. The proof is analogous to those of Lemma 3.3 and Theorem 1.1. We assume that φ is term-reduced and we prove the claim by induction on φ.

For atomic formulae φ, the claim holds by the definition of a reduced product. If φ is a conjunction then the claim follows immediately from the inductive hypothesis and the fact that filters are closed under finite intersections. Hence, we may assume that $\varphi(\bar{x}) = \exists \bar{y} \psi(\bar{x}, \bar{y})$.

If $\prod_i \mathfrak{A}^i/\mathfrak{u} \vDash \varphi([\bar{a}])$ then there exists a sequence $\bar{b} \subseteq \prod_i A^i$ such that $\prod_i \mathfrak{A}^i/\mathfrak{u} \vDash \psi([\bar{a}], [\bar{b}])$. By inductive hypothesis, we therefore have

$$[\![\psi(\bar{a}, \bar{b})]\!] \in \mathfrak{u}\,.$$

Since $[\![\psi(\bar{a}, \bar{b})]\!] \subseteq [\![\exists \bar{y} \psi(\bar{a}, \bar{y})]\!]$ it follows that

$$[\![\exists \bar{y} \psi(\bar{a}, \bar{b})]\!] \in \mathfrak{u}\,.$$

Conversely, suppose that $s := [\![\exists \bar{y} \psi(\bar{a}, \bar{b})]\!] \in \mathfrak{u}$. For every $i \in s$, we choose sequences $\bar{b}^i \subseteq A^i$ such that $\mathfrak{A}^i \vDash \psi(\bar{a}^i, \bar{b}^i)$. For $i \in [\![\exists \bar{y}\, \text{true}]\!] \setminus s$, we take an arbitrary tuple $\bar{b}^i \subseteq A^i$. Then $[\![\psi(\bar{a}, \bar{b})]\!] = s \in \mathfrak{u}$ which implies that $\prod_i \mathfrak{A}^i/\mathfrak{u} \vDash \psi([\bar{a}], [\bar{b}])$. Consequently, we have $\prod_i \mathfrak{A}^i/\mathfrak{u} \vDash \varphi([\bar{a}])$. □

For first-order Horn formulae we can extend the Theorem of Łoś to arbitrary filters.

Theorem 3.6. *Suppose that* $\varphi(\bar{x})$ *is a first-order Horn formula,* $(\mathfrak{A}^i)_{i \in I}$ *a nonempty sequence of structures,* $\mathfrak{u}$ *a filter on I, and* $[\bar{a}]$ *a tuple in* $\prod_i \mathfrak{A}^i/\mathfrak{u}$. *Then*

$$[\![\varphi(\bar{a})]\!] \in \mathfrak{u} \quad \textit{implies} \quad \prod_i \mathfrak{A}^i/\mathfrak{u} \vDash \varphi([\bar{a}])\,.$$

Proof. We prove the claim by induction on φ. Let $s := [\![\varphi(\bar{a})]\!] \in \mathfrak{u}$. First, we consider the case that $\varphi = \bigwedge \Phi \to \psi$ is a basic Horn formula. If $\prod_i \mathfrak{A}^i/\mathfrak{u} \not\models \Phi([\bar{a}])$ then we are done. Hence we may assume that $\prod_i \mathfrak{A}^i/\mathfrak{u} \models \Phi([\bar{a}])$. By Lemma 3.5, it follows that $w := [\![\bigwedge \Phi(\bar{a})]\!] \in \mathfrak{u}$. Consequently, we have $s \cap w \subseteq [\![\psi(\bar{a})]\!] \in \mathfrak{u}$. In this case ψ cannot be false and we can use Lemma 3.3 to conclude that $\prod_i \mathfrak{A}^i/\mathfrak{u} \models \psi([\bar{a}])$, as desired.

If φ is a conjunction then the claim follows immediately by inductive hypothesis. For $\varphi = \exists \bar{y}\psi(\bar{x}, \bar{y})$ we can argue in the same way as in the proof of Lemma 3.5. Finally, assume that $\varphi = \forall \bar{y}\psi(\bar{x}, \bar{y})$. Let $b_k \in \prod_{i \in I_{s_k}} A^i_{s_k}$. Then $s \subseteq [\![\psi(\bar{a}, \bar{b})]\!] \in \mathfrak{u}$. By inductive hypothesis, this implies that $\prod_i \mathfrak{A}^i/\mathfrak{u} \models \psi([\bar{a}], [\bar{b}])$. Since $\bar{b}$ was arbitrary it follows that $\prod_i \mathfrak{A}^i/\mathfrak{u} \models \forall \bar{y}\psi([\bar{a}], [\bar{y}])$. □

Corollary 3.7. *Let Σ be a signature and X a set of variables.*

(a) $\mathrm{HO}_\infty[\Sigma, X]$*-formulae are preserved in nonempty products.*

(b) $\mathrm{SH}_\infty[\Sigma, X]$*-formulae are preserved in products.*

(c) $\mathrm{HO}[\Sigma, X]$*-formulae are preserved in nonempty reduced products.*

(d) $\mathrm{SH}[\Sigma, X]$*-formulae are preserved in reduced products.*

Proof. (a) and (c) follow immediately from Theorem 3.4 and 3.6, respectively. For (b) and (d) it is sufficient to note that in the empty product 1 every n-ary relation contains the tuple $\langle\langle\rangle, \dots, \langle\rangle\rangle$. Hence, we have

$$1 \models \varphi(\langle\rangle, \dots, \langle\rangle),$$

for every atomic formula φ. □

Example. Groups, rings, and modules are SH-axiomatisable. Hence, theses classes are closed under reduced products.

Lemma 3.8 (McKinsey). *Let $\mathcal{K}$ be a class of structures that is closed under nonempty products. If Φ is a set of Horn formulae and Ψ a nonempty set of atomic formulae (possibly including the formula* false*) such that*

$$\mathfrak{A} \models \forall \bar{x}\Big(\bigwedge \Phi \to \bigvee \Psi\Big), \quad \text{for all } \mathfrak{A} \in \mathcal{K},$$

then there is some formula $\psi \in \Psi$ such that

$$\mathfrak{A} \vDash \forall \bar{x}\Big(\bigwedge \Phi \to \psi\Big), \quad \textit{for all } \mathfrak{A} \in \mathcal{K}.$$

Proof. For a contradiction, suppose that, for every formula $\psi \in \Psi$ there are a structure $\mathfrak{A}^\psi \in \mathcal{K}$ and parameters $\bar{a}^\psi \subseteq A^\psi$ such that

$$\mathfrak{A}^\psi \vDash \bigwedge \Phi(\bar{a}^\psi) \wedge \neg\psi(\bar{a}^\psi).$$

Set $\mathfrak{B} := \prod_{\psi \in \Psi} \mathfrak{A}^\psi$ and $\bar{b} := (\bar{a}^\psi)_\psi \subseteq B$. Since $\Psi \neq \varnothing$ we have $\mathfrak{B} \in \mathcal{K}$. Furthermore, it follows by Theorem 3.4 that

$$\mathfrak{B} \vDash \bigwedge \Phi(\bar{b}).$$

Hence, there is some $\psi \in \Psi$ such that $\mathfrak{B} \vDash \psi(\bar{b})$. By Lemma 3.3, this implies that $\mathfrak{A}^\psi \vDash \psi(\bar{a}^\psi)$. Contradiction. □

The converse of Corollary 3.7 is given by the following preservation theorem.

Theorem 3.9. *A first-order sentence φ is preserved in nonempty reduced products if and only if it is equivalent to a first-order Horn sentence.*

4. Quasivarieties

Classes that are axiomatised by universal Horn formulae admit a nice algebraic characterisation.

Definition 4.1. Let $\mathcal{K}$ be a class of Σ-structures.

(a) A *$\mathcal{K}$-presentation* is a pair $\langle C; \Phi\rangle$ consisting of a set C of constant symbols disjoint from Σ and a set Φ of atomic sentences over the signature $\Sigma_C = \Sigma \cup C$. The constants in C are called the *generators* of the presentation.

(b) A *model* of a $\mathcal{K}$-presentation $\langle C; \Phi\rangle$ is a Σ_C-structure $\mathfrak{A}$ such that

$$\mathfrak{A} \vDash \Phi \quad \text{and} \quad \mathfrak{A}|_\Sigma \in \mathcal{K}.$$

(c) A model $\mathfrak{A}$ of a $\mathcal{K}$-presentation $\langle C; \Phi\rangle$ is *free* if

- $\mathfrak{A}$ is generated by the constants in C and
- for every model $\mathfrak{B}$ of $\langle C; \Phi \rangle$ there is a homomorphism $\mathfrak{A} \to \mathfrak{B}$.

(d) We say that $\mathcal{K}$ *has free models* if every $\mathcal{K}$-presentation has a free model.

Remark. Note that the homomorphism $h : \mathfrak{A} \to \mathfrak{B}$ in (c) maps $c^{\mathfrak{A}}$ to $c^{\mathfrak{B}}$, for every $c \in C$. Since $\mathfrak{A}$ is generated by C it follows that h is unique.

Example. Let $\mathcal{K}$ be the class of all groups, $C := \{a, b\}$, and let Φ be the set consisting of the single formula $a \cdot b = b \cdot a$. Then $\langle C; \Phi \rangle$ is a $\mathcal{K}$-presentation. Its free model consists of the direct product

$$\langle \mathbb{Z}, +, -, 0 \rangle \times \langle \mathbb{Z}, +, -, 0 \rangle$$

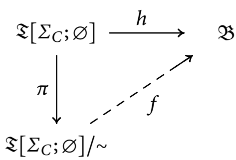
with additional constants $a = \langle 0, 1 \rangle$ and $b = \langle 1, 0 \rangle$.

Example. Suppose that Σ is a signature without relation symbols. The class $\mathrm{Str}[\Sigma]$ of all Σ-structures has free models. Let $\langle C; \Phi \rangle$ be a $\mathrm{Str}[\Sigma]$-presentation. W.l.o.g. we may assume that Φ is closed under entailment. In particular, it is =-closed and, as in Lemma C2.4.4, we obtain a Herbrand model $\mathfrak{H}$ of Φ that is of the form $\mathfrak{H} = \mathfrak{T}[\Sigma_C; \varnothing]/{\sim}$ where

$$s \sim t \quad \text{iff} \quad s = t \in \Phi\,.$$

We claim that $\mathfrak{H}$ is a free model of $\langle C; \Phi \rangle$.

Suppose that $\mathfrak{B}$ is a model of $\langle C; \Phi \rangle$. We have to find a homomorphism $f : \mathfrak{H} \to \mathfrak{B}$. Let π be the canonical projection $\mathfrak{T}[\Sigma_C; \varnothing] \to \mathfrak{H}$. By Theorem B3.1.9, there exists a unique homomorphism $h : \mathfrak{T}[\Sigma_C; \varnothing] \to \mathfrak{B}$.



Since $\mathfrak{B}$ is a model of Φ it follows that $\ker \pi = \sim \subseteq \ker h$. Hence, we can use the Factorisation Lemma to find the desired homomorphism $f : \mathfrak{H} \to \mathfrak{B}$.

We start by giving conditions ensuring that $\mathcal{K}$ has free models.

Lemma 4.2. *Let $\langle C; \Phi \rangle$ be a $\mathcal{K}$-presentation and $\mathfrak{A}$ a Σ_C-structure with $\mathfrak{A}|_\Sigma \in \mathcal{K}$. Then $\mathfrak{A}$ is a free model of $\langle C; \Phi \rangle$ if and only if*

- *C generates $\mathfrak{A}$ and*
- *for every atomic formula φ over Σ_C, we have*

$$(*) \quad \mathfrak{A} \vDash \varphi \quad \text{iff} \quad \mathrm{Th}(\mathcal{K}) \vDash \bigwedge \Phi \to \varphi\,.$$

Proof. Let $\varphi_o(\bar{x}) \in \mathrm{FO}[\Sigma, X]$ and $\Phi_o(\bar{x}) \subseteq \mathrm{FO}[\Sigma, X]$ be the formulae obtained from φ and Φ by replacing the constant symbols in C by variables.

($\Rightarrow$) If every structure in $\mathcal{K}$ satisfies the sentence

$$\forall \bar{x}\big[\bigwedge \Phi_o(\bar{x}) \to \varphi_o(\bar{x})\big]$$

then, in particular, so does $\mathfrak{A}|_\Sigma$. Hence, $\mathfrak{A} \vDash \varphi$.

Conversely, suppose that $\mathfrak{A} \vDash \varphi_o(\bar{c})$ and let $\mathfrak{B} \in \mathcal{K}$ be a structure with $\mathfrak{B} \vDash \Phi_o(\bar{b})$, for some $\bar{b} \subseteq B$. Since $\mathfrak{A}$ is free there exists a homomorphism $h : \mathfrak{A} \to \langle \mathfrak{B}, \bar{b} \rangle$. Since $h(\bar{c}) = \bar{b}$ and atomic formulae are preserved under homomorphisms it follows that

$$\mathfrak{A} \vDash \varphi_o(\bar{c}) \quad \text{implies} \quad \mathfrak{B} \vDash \varphi_o(\bar{b})\,,$$

as desired.

($\Leftarrow$) For every $\varphi \in \Phi$ we have $\Phi \vDash \varphi$. By $(*)$, this implies that $\mathfrak{A} \vDash \varphi$. Consequently, $\mathfrak{A}$ is a model of $\langle C; \Phi \rangle$. If $\mathfrak{B}$ is another model of $\langle C; \Phi \rangle$ then we have $\mathfrak{B}|_\Sigma \in \mathcal{K}$ and $\mathfrak{B}|_\Sigma \vDash \Phi_o(\bar{c}^{\mathfrak{B}})$. By $(*)$ it follows that $\mathfrak{B} \vDash \psi$, for every atomic formula ψ with $\mathfrak{A} \vDash \psi$. Consequently, $\mathfrak{B}$ satisfies the atomic diagram of $\mathfrak{A}$ and we can use Corollary C2.2.4 to find a homomorphism $\mathfrak{A} \to \mathfrak{B}$. □

Theorem 4.3. *Let $\mathcal{K}$ be a class of Σ-structures that is closed under isomorphic copies. The following statements are equivalent:*

(1) *Every $\mathcal{K}$-presentation with a model has a free model.*

(2) *$\mathcal{K}$ is closed under nonempty products and substructures.*

(3) *$\mathcal{K}$ is $\mathrm{H}\forall_\infty$-axiomatisable.*

Proof. (3) ⇒ (2) follows from Corollary 3.7 and Lemma C2.1.6.

(2) ⇒ (1) Let $\langle C; \Phi\rangle$ be a $\mathcal{K}$-presentation with a model and let Ψ be the set of all atomic formulae $\psi(\bar{x})$ (including false) such that

$$\mathrm{Th}(\mathcal{K}) \not\models \bigwedge \Phi \to \psi(\bar{c}) .$$

If every model of $\langle C; \Phi\rangle$ would satisfy $\bigvee \Psi$ then it would follow by Lemma 3.8 that

$$\mathrm{Th}(\mathcal{K}) \models \bigwedge \Phi \to \psi(\bar{c}) ,$$

for some $\psi \in \Psi$. By choice of Ψ we can therefore find some structure $\mathfrak{A} \in \mathcal{K}$ and elements $\bar{c} \subseteq A$ such that

$$\langle \mathfrak{A}, \bar{c}\rangle \models \bigwedge \Phi \wedge \neg \bigvee \Psi .$$

It follows that

$$\langle \mathfrak{A}, \bar{c}\rangle \models \varphi \quad \text{iff} \quad \mathrm{Th}(\mathcal{K}) \models \bigwedge \Phi \to \varphi ,$$

for every atomic formula φ over Σ_C. Setting $\mathfrak{A}_0 := \langle\!\langle \bar{c} \rangle\!\rangle_{\mathfrak{A}}$ we still have

$$\langle \mathfrak{A}_0, \bar{c}\rangle \models \varphi \quad \text{iff} \quad \mathrm{Th}(\mathcal{K}) \models \bigwedge \Phi \to \varphi ,$$

for all such φ. Since $\mathcal{K}$ is closed under substructures we have $\mathfrak{A}_0 \in \mathcal{K}$. Hence, Lemma 4.2 implies that $\mathfrak{A}_0$ is a free model of $\langle C; \Phi\rangle$.

(1) ⇒ (3) Set $T := \mathrm{Th}_{\mathrm{H}\forall_\infty}(\mathcal{K})$ and suppose that $\mathfrak{B}$ is a model of T. Let Φ be the atomic diagram of $\mathfrak{B}$. Because $\mathfrak{B}$ is a model of the $\mathcal{K}$-presentation $\langle B; \Phi\rangle$ there exists, by (1), a free model $\mathfrak{A}$ of $\langle B; \Phi\rangle$. By

Corollary c2.2.4 there exists a homomorphism $h : \mathfrak{B} \to \mathfrak{A}$. Since B generates $\mathfrak{A}$ this homomorphism is surjective. If we can show that it is an embedding then it follows that $\mathfrak{B} \cong \mathfrak{A}|_\Sigma$ and, since $\mathcal{K}$ is closed under isomorphic copies, we have $\mathfrak{B} \in \mathcal{K}$, as desired.

Let $\Phi_0(\bar{x})$ be the set of formulae obtained from Φ by replacing the constants in B by variables. Let $\psi(\bar{x})$ be an atomic formula over Σ with $\mathfrak{A} \vDash \psi(\bar{b})$, for some $\bar{b} \subseteq B$. By Lemma 4.2, T contains the formula

$$\forall \bar{x}\Big(\bigwedge \Phi_0(\bar{x}) \to \psi(\bar{x})\Big).$$

Since $\mathfrak{B} \vDash T$ we have $\mathfrak{B} \vDash \bigwedge \Phi_0(\bar{b}) \to \psi(\bar{b})$. By definition of Φ this implies that $\mathfrak{B} \vDash \psi(\bar{b})$. Consequently, h is an embedding. □

Theorem 4.4. *Let $\mathcal{K}$ be a class of Σ-structures that is closed under isomorphic copies. The following statements are equivalent:*

(1) *$\mathcal{K}$ has free models.*

(2) *$\mathcal{K}$ is closed under products and substructures.*

(3) *$\mathcal{K}$ is* SH$\forall_\infty$*-axiomatisable.*

Proof. (3) ⇒ (2) follows from Corollary 3.7 and Lemma c2.1.6.

(2) ⇒ (1) Note that the empty product is a model of every $\mathcal{K}$-presentation. Hence, the claim follows from Theorem 4.3.

(1) ⇒ (3) By Theorem 4.3, we know that $\mathcal{K}$ has an H$\forall_\infty$-axiomatisation T. We claim that $T \subseteq \text{SH}\forall_\infty$. Suppose otherwise. Then T contains a formula of the form $\forall \bar{x}(\bigwedge \Phi \to \text{false})$. Let $X := \bar{x}$ be the set of variables appearing in Φ. The $\mathcal{K}$-presentation $\langle X; \Phi\rangle$ has a free model $\langle \mathfrak{A}, \bar{c}\rangle$, by (1). But then $\mathfrak{A} \vDash \Phi(\bar{c})$ would imply that $\mathfrak{A} \vDash \text{false}$. A contradiction. □

Definition 4.5. Let $\mathcal{K}$ be a class of Σ-structures.

(a) $\mathcal{K}$ is a *quasivariety* if it is SH$\forall$-axiomatisable.

(b) $\mathcal{K}$ is a *variety* if it can be axiomatised by a set of formulae of the form $\forall \bar{x}\varphi$ where φ is an atomic formula.

Example. The classes of all groups, all rings, and all modules are varieties. The class of lattices (with signature $\sqcap$, $\sqcup$, $\sqsubseteq$) is a quasivariety, but not a variety. If we omit $\sqsubseteq$ then the class becomes a variety. The class of all fields is not a quasivariety.

Definition 4.6. Let $\mathcal{K}$ be a class of Σ-structures. We define the following operations.

(a) $\mathrm{Prod}(\mathcal{K})$ is the class of all nonempty products of structures in $\mathcal{K}$.

(b) $\mathrm{Sub}(\mathcal{K})$ is the class of all substructures of structures in $\mathcal{K}$.

(c) $\mathrm{Iso}(\mathcal{K})$ is the class of all structures isomorphic to one in $\mathcal{K}$.

(d) $\mathrm{Hom}(\mathcal{K})$ is the class of all weak homomorphic images of structures in $\mathcal{K}$.

(e) $\mathrm{ERP}(\mathcal{K})$ is the class of all structures that can be embedded into a reduced product of structures in $\mathcal{K}$.

(f) Finally, we define the abbreviations

$$\mathrm{QV} := \mathrm{Iso} \circ \mathrm{Sub} \circ \mathrm{Prod}\,,$$
$$\mathrm{Var} := \mathrm{Hom} \circ \mathrm{Sub} \circ \mathrm{Prod}\,.$$

Theorem 4.7. *Let $\mathcal{K}$ be a class of Σ-structures.*

(a) $\mathrm{QV}(\mathcal{K})$ *is the smallest class of Σ-structures containing $\mathcal{K}$ that is closed under products, substructures, and isomorphic copies.*

(b) $\mathrm{QV}(\mathcal{K}) = \mathrm{Mod}(\mathrm{Th}_{\mathrm{SH}\forall_\infty}(\mathcal{K}))$.

(c) *If $\mathcal{K}$ or $\mathrm{QV}(\mathcal{K})$ is first-order axiomatisable then $\mathrm{QV}(\mathcal{K})$ is a quasivariety.*

Proof. Let $T := \mathrm{SH}\forall_\infty(\mathcal{K})$.

(a) and (b) Let $\mathcal{H}$ be the smallest class of Σ-structures containing $\mathcal{K}$ that is closed under products, substructures, and isomorphic copies. Then we have $\mathrm{QV}(\mathcal{K}) = (\mathrm{Iso} \circ \mathrm{Sub} \circ \mathrm{Prod})(\mathcal{K}) \subseteq \mathcal{H}$. Furthermore, by

Lemma c2.1.6 and Corollary 3.7 it follows that every structure in $\mathcal{H}$ is a model of T. Consequently, we have

$$\mathrm{QV}(\mathcal{K}) \subseteq \mathcal{H} \subseteq \mathrm{Mod}(T)\,,$$

and it remains to prove that $\mathrm{Mod}(T) \subseteq \mathrm{QV}(\mathcal{K})$.

Suppose that $\mathfrak{A} \vDash T$ and fix an enumeration $\bar{a}$ of A without repetitions. Let $\Phi(\bar{x})$ be the set of all atomic formulae $\varphi(\bar{x})$ with $\mathfrak{A} \vDash \varphi(\bar{a})$ and let $\Psi(\bar{x})$ be the set of all atomic formulae $\varphi(\bar{x})$ (including false) with $\mathfrak{A} \nvDash \varphi(\bar{a})$. Consider a formula $\psi \in \Psi$. Since $\mathfrak{A}$ is a model of T we have $\forall\bar{x}(\bigwedge\Phi \to \psi) \notin T$. Therefore, we can find a structure $\mathfrak{B}^\psi \in \mathcal{K}$ and parameters $\bar{b}^\psi \subseteq B$ such that

$$\mathfrak{B}^\psi \vDash \bigwedge \Phi(\bar{b}^\psi) \wedge \neg\psi(\bar{b}^\psi)\,.$$

Let $\langle \mathfrak{C}, \bar{c}\rangle := \prod_{\psi\in\Psi}\langle\mathfrak{B}^\psi, \bar{b}^\psi\rangle$. Since the algebraic diagrams of $\langle\!\langle \bar{c}\rangle\!\rangle_{\langle\mathfrak{C},\bar{c}\rangle}$ and $\langle\mathfrak{A}, \bar{a}\rangle$ coincide we can use Corollary c2.2.4 to find an embedding $h : \mathfrak{A} \to \mathfrak{C}$ with $h(\bar{a}) = \bar{c}$. Hence, $\mathfrak{A}$ is isomorphic to a substructure of a product of structures in $\mathcal{K}$, i.e.,

$$\mathfrak{A} \in (\mathrm{Iso} \circ \mathrm{Sub} \circ \mathrm{Prod})(\mathcal{K}) = \mathrm{QV}(\mathcal{K})\,.$$

(c) Let T_0 be an axiomatisation of either $\mathcal{K}$ or $\mathrm{QV}(\mathcal{K})$. Note that in both cases we have $T = (T_0)^{\vDash}_{\mathrm{SH}\forall_\infty}$. For every formula $\varphi \in T$, we will construct a first-order formula $\varphi' \in T$ with $\varphi' \vDash \varphi$. This implies that $T \cap \mathrm{FO} \vDash T$. It follows that $\mathrm{Mod}(T \cap \mathrm{FO}) = \mathrm{QV}(\mathcal{K})$, as desired.

It remains to find φ'. Let $\forall\bar{x}(\bigwedge\Phi \to \psi) \in T$. Then $T_0 \cup \Phi \vDash \psi$. By the Compactness Theorem, we can find a finite subset $\Phi_0 \subseteq \Phi$ such that $T_0 \cup \Phi_0 \vDash \psi$. Setting $\varphi' := \forall\bar{x}(\bigwedge\Phi_0 \to \psi)$ it follows that $T_0 \vDash \varphi'$. Furthermore, since φ' is a universal strict Horn formula we have $\varphi' \in T$, as desired. □

Corollary 4.8. *A class $\mathcal{K}$ is a quasivariety if and only if it is first-order axiomatisable and closed under products, substructures, and isomorphic copies.*

Lemma 4.9. *Let $\mathfrak{A}$ be a Σ-structure and $\mathcal{K}$ a nonempty class of Σ-structures.*

$$\mathfrak{A} \in \mathrm{ERP}(\mathcal{K}) \quad \text{iff} \quad \mathrm{Th}_{\mathrm{H}\forall}(\mathcal{K}) \subseteq \mathrm{Th}_{\mathrm{H}\forall}(\mathfrak{A}) .$$

Proof. ($\Rightarrow$) follows from the preservation properties of universal Horn formulae. For ($\Leftarrow$), suppose that $\mathrm{Th}_{\mathrm{H}\forall}(\mathcal{K}) \subseteq \mathrm{Th}_{\mathrm{H}\forall}(\mathfrak{A})$. Let Δ_+ be the set of all atomic first-order formulae and Δ_- the set of all negations of atomic first-order formulae. We set $\Phi_+ := \mathrm{Th}_{\Delta_+}(\mathfrak{A}_A)$ and $\Phi_- := \mathrm{Th}_{\Delta_-}(\mathfrak{A}_A)$.

First, we show that, for every finite subset $\Psi \subseteq \Phi_+$, there exists a structure $\mathfrak{B}^\Psi \in \mathrm{Prod}(\mathcal{K})$ and parameters $\bar{b}^\Psi \subseteq B$ such that

$$\langle \mathfrak{B}^\Psi, \bar{b}^\Psi \rangle \vDash \Psi \cup \Phi_- .$$

Suppose that $\Psi = \{\psi_0(\bar{a}), \dots, \psi_n(\bar{a})\}$. For every $\neg\varphi(\bar{a}) \in \Phi_-$, we have

$$\mathfrak{A} \vDash \psi_0(\bar{a}) \wedge \dots \wedge \psi_n(\bar{a}) \wedge \neg\varphi(\bar{a}) .$$

It follows that $\mathfrak{A} \nvDash \psi_0(\bar{a}) \wedge \dots \wedge \psi_n(\bar{a}) \rightarrow \varphi(\bar{a})$. By assumption this implies that

$$\forall \bar{x}[\psi_0(\bar{x}) \wedge \dots \wedge \psi_n(\bar{x}) \rightarrow \varphi(\bar{x})] \notin \mathrm{Th}_{\mathrm{H}\forall}(\mathcal{K}) .$$

Consequently, there is a structure $\mathfrak{C}^\varphi \in \mathcal{K}$ and elements $\bar{c}^\varphi \subseteq C$ such that

$$\mathfrak{C}^\varphi \vDash \psi_0(\bar{c}^\varphi) \wedge \dots \wedge \psi_n(\bar{c}^\varphi) \wedge \neg\varphi(\bar{c}^\varphi) .$$

Similarly, we have

$$\forall \bar{x}[\psi_0(\bar{x}) \wedge \dots \wedge \psi_n(\bar{x}) \rightarrow \text{false}] \notin \mathrm{Th}_{\mathrm{H}\forall}(\mathcal{K}) ,$$

and there is a structure $\mathfrak{C}^\perp \in \mathcal{K}$ and elements $\bar{c}^\perp \subseteq C$ such that

$$\mathfrak{C}^\perp \vDash \psi_0(\bar{c}^\perp) \wedge \dots \wedge \psi_n(\bar{c}^\perp) .$$

We form the product

$$\langle \mathfrak{B}, \bar{b} \rangle := \prod_{\varphi \in \Phi_- \cup \{\perp\}} \langle \mathfrak{C}^\varphi, \bar{c}^\varphi \rangle .$$

By Lemma 3.3 it follows that

$$\begin{aligned} &\mathfrak{B} \vDash \psi_i(\bar{b})\,, && \text{for all } i\,, \\ \text{and}\quad &\mathfrak{B} \vDash \neg\varphi(\bar{b})\,, && \text{for all } \neg\varphi \in \Phi_-\,. \end{aligned}$$

Furthermore,

$$\mathfrak{B} = \prod_{\varphi \in \Phi_- \cup \{\bot\}} \mathfrak{C}^\varphi \in \mathrm{Prod}(\mathcal{K})\,,$$

as desired.

It remains to construct a model $\langle \mathfrak{D}, \bar{d} \rangle$ of $\Phi_+ \cup \Phi_-$ that is a reduced product of structures in $\mathcal{K}$. By the Diagram Lemma, this implies that $\mathfrak{A}$ can be embedded into the product $\mathfrak{D}$.

If Φ_+ is finite we can use the structure $\langle \mathfrak{B}^{\Phi_+}, \bar{b}^{\Phi_+} \rangle$. Hence, we may assume that Φ_+ is infinite. Let $\mathfrak{u}$ be a regular ultrafilter over Φ_+ and let $(s_\varphi)_{\varphi \in \Phi_+}$ be the corresponding sequence of sets $s_\varphi \in \mathfrak{u}$ such that, for every $i \in \Phi_+$, the set

$$w_i := \{\, \varphi \in \Phi_+ \mid i \in s_\varphi \,\}$$

is finite. We claim that the reduced product

$$\langle \mathfrak{D}, \bar{d} \rangle := \prod_{i \in \Phi_+} \langle \mathfrak{B}^{w_i}, \bar{b}^{w_i} \rangle / \mathfrak{u}$$

is the desired model of $\Phi_+ \cup \Phi_-$.

First consider $\varphi(\bar{a}) \in \Phi_+$. For every $i \in s_\varphi$, we have $\mathfrak{B}^{w_i} \vDash \varphi(\bar{b}^{w_i})$. Therefore, $s_\varphi \subseteq [\![\varphi(\bar{d})]\!] \in \mathfrak{u}$ and it follows that $\mathfrak{D} \vDash \varphi(\bar{d})$. Furthermore, we have $\langle \mathfrak{D}, \bar{d} \rangle \vDash \Phi_-$, since $\langle \mathfrak{B}^{w_i}, \bar{b}^{w_i} \rangle \vDash \Phi_-$, for all i. Finally, note that $\mathfrak{D}$ is a reduced product of structures in $\mathrm{Prod}(\mathcal{K})$. Therefore, it can be written as a reduced product of structures in $\mathcal{K}$. □

Theorem 4.10. *Let $\mathcal{K}$ be a class of Σ-structures. The following statements are equivalent:*

(1) *$\mathcal{K}$ is closed under substructures, reduced products, and isomorphic copies.*

(2) $\mathcal{K}$ *is* $\mathrm{H}\forall$*-axiomatisable.*

Proof. (2) ⇒ (1) follows from the preservation properties of universal Horn formulae. For (1) ⇒ (2), let $T := \mathrm{Th}_{\mathrm{H}\forall}(\mathcal{K})$. By Lemma 4.9, we have

$$\mathrm{Mod}(T) \subseteq \mathrm{ERP}(\mathcal{K}) = \mathcal{K} \subseteq \mathrm{Mod}(T)\,,$$

as desired. □

Corollary 4.11. *Let T be a $\mathrm{H}\forall[\Sigma]$-theory and $\varphi \in \mathrm{FO}[\Sigma]$ a first-order formula. The following statements are equivalent:*

(1) *We have $\mathfrak{A} \vDash \varphi$, for every structure $\mathfrak{A} \in \mathrm{ERP}(\mathrm{Mod}(T \cup \{\varphi\}))$.*

(2) *φ is equivalent modulo T to a finite conjunction of $\mathrm{H}\forall[\Sigma]$-formulae.*

Proof. (2) ⇒ (1) follows from the preservation properties of universal Horn formulae. For (1) ⇒ (2), let $\Phi := (T \cup \{\varphi\})^{\vDash}_{\mathrm{H}\forall}$. Clearly, $T \cup \{\varphi\} \vDash \Phi$. If we can show that $\Phi \vDash T \cup \{\varphi\}$ then the claim follows by compactness.

Suppose that $\mathfrak{A} \vDash \Phi$. By Lemma 4.9, we have

$$\mathfrak{A} \in \mathrm{ERP}(\mathrm{Mod}(T \cup \{\varphi\}))\,,$$

which, by (1), implies that $\mathfrak{A} \vDash \varphi$. Furthermore, we have $\mathfrak{A} \vDash T$ since $T \subseteq \Phi$. □

Theorem 4.12. *Let $\mathcal{K}$ be a class of Σ-structures containing the empty product and set $T := \mathrm{Th}_{\mathrm{H}\forall}(\mathcal{K})$. Then*

$$\mathrm{QV}(\mathcal{K}) = \mathrm{ERP}(\mathcal{K}) = \mathrm{Mod}(T)\,.$$

Proof. Let $\mathcal{Q}$ be the class of all structures that can be embedded into a reduced product of structures in $\mathcal{K}$. Any quasivariety containing $\mathcal{K}$ must contain $\mathcal{Q}$. Hence, it is sufficient to show that $\mathcal{Q}$ is a quasivariety.

By Lemma 4.9, we have $\mathcal{Q} = \mathrm{Mod}(T)$. Every Horn formula in T is strict since $\mathcal{K}$ contains the empty product. Consequently, $T \subseteq \mathrm{SH}\forall[\Sigma]$ and $\mathcal{Q} = \mathrm{Mod}(T)$ is a quasivariety. □

We conclude this section with a analogous characterisations of varieties.

Definition 4.13. Let $\mathcal{K}$ be a class of structures. A element $\mathfrak{A} \in \mathcal{K}$ is *free* (in $\mathcal{K}$) if there exists a subset $C \subseteq A$ such that $\mathfrak{A}_C$ is a free model of $\langle C; \varnothing \rangle$. In this case we also say that $\mathfrak{A}$ is *freely generated by C*.

We can use Lemma 4.2 to obtain a characterisation of free structures.

Lemma 4.14. *Let $\mathcal{K}$ be a class of structures, $\mathfrak{A} \in \mathcal{K}$, and $C \subseteq A$. Then $\mathfrak{A}$ is freely generated by C if and only if $\mathfrak{A}$ is generated by C and, for every tuple $\bar{a} \subseteq C$ of distinct elements and each atomic formula $\varphi(\bar{x})$ with $\mathfrak{A} \vDash \varphi(\bar{a})$, we have*

$$\mathfrak{B} \vDash \forall \bar{x} \varphi\,, \quad \text{for all } \mathfrak{B} \in \mathcal{K}\,.$$

Lemma 4.15. *Let $\mathcal{K}$ be a class of structures and $\mathfrak{A}$ and $\mathfrak{B}$ structures in $\mathcal{K}$ freely generated by, respectively, C and D. If $|C| = |D|$ then every bijection $C \to D$ extends to an isomorphism $\mathfrak{A} \cong \mathfrak{B}$.*

Proof. Let $f : C \to D$ be a bijection. By definition of a free model, we can extend f to a homomorphism $g : \mathfrak{A} \to \mathfrak{B}$ and f^{-1} to a homomorphism $h : \mathfrak{B} \to \mathfrak{A}$. Since $h \circ g$ is a homomorphism $\mathfrak{A} \to \mathfrak{A}$ with $(h \circ g) \upharpoonright C = \mathrm{id}_C$ it follows by uniqueness that $h \circ g = \mathrm{id}_A$. Similarly, we have $g \circ h = \mathrm{id}_B$. Hence, $g : \mathfrak{A} \to \mathfrak{B}$ is the desired isomorphism. □

Lemma 4.16. *Let $\mathcal{K}$ be a class of structures that is closed under nonempty products, substructures, and isomorphic copies.*

(a) *If a structure $\mathfrak{A} \in \mathcal{K}$ is generated by a set X of size $|X| = \kappa$ then $\mathcal{K}$ contains a structure $\mathfrak{F}_\kappa \in \mathcal{K}$ that is freely generated by a set of size κ. Furthermore, there exists a surjective homomorphism $\mathfrak{F}_\kappa \to \mathfrak{A}$.*

(b) *If $\mathcal{K}$ contains a structure with at least 2 elements then $\mathcal{K}$ contains, for every cardinal κ, a structure that is freely generated by a set of size κ.*

Proof. (a) Let C be a set of κ constant symbols. By Theorem 4.3, the $\mathcal{K}$-presentation $\langle C; \varnothing \rangle$ has a free model $\mathfrak{F}$. If we can show that $c^{\mathfrak{F}} \neq d^{\mathfrak{F}}$, for all distinct constants $c, d \in C$, then it follows that $\mathfrak{F}$ is freely generated by C.

For a contradiction, suppose that there are $c \neq d$ with $c^{\mathfrak{F}} = d^{\mathfrak{F}}$. By Lemma 4.14 it follows that every structure in $\mathcal{K}$ satisfies $\forall x \forall y (x = y)$. Hence, every structure in $\mathcal{K}$ has at most 1 element. This contradicts the fact that $\mathfrak{A}$ contains a subset $X \subseteq A$ of size κ.

Finally, note that we can extend any bijection $C \to X$ to a homomorphism $\mathfrak{F} \to \mathfrak{A}$. Since $\mathfrak{A}$ is generated by X this homomorphism is surjective and $\mathfrak{A}$ is a weak homomorphic image of $\mathfrak{F}$.

(b) follows from (a). If $\mathcal{K}$ contains a structure with at least 2 elements then $\mathcal{K}$ contains arbitrarily large structures since it is closed under products. □

Theorem 4.17 (Birkhoff). *Let $\mathcal{K}$ be a class of Σ-structures. The following statements are equivalent:*

(1) *$\mathcal{K}$ is closed under nonempty products, substructures, and weak homomorphic images.*

(2) $\mathcal{K} = \mathrm{Var}(\mathcal{K})$

(3) *$\mathcal{K}$ is axiomatised by a set of formulae of the form $\forall \bar{x} \varphi$ where φ is an atomic formula.*

Proof. It is easy to see that (1) and (2) are equivalent. The implication (3) ⇒ (1) follows from Lemmas C2.1.6 and C2.1.3 (a), and Corollary 3.7. Hence, it remains to prove that (1) implies (3).

Set $\mathcal{H} := \mathrm{Mod}(T)$ where T is the set of all sentences $\forall \bar{x} \varphi \in \mathrm{Th}(\mathcal{K})$ where φ is an atomic formula. We have to show that $\mathcal{H} \subseteq \mathcal{K}$.

First, we consider the case that $\mathcal{K}$ contains a structure with at least 2 elements. Then $\mathcal{K}$ has arbitrarily large free structures $\mathfrak{F}$, by Lemma 4.16. Hence, $\mathfrak{F} \in \mathcal{K} \subseteq \mathcal{H}$. But the class $\mathcal{H}$ is closed under nonempty products, substructures, and weak homomorphic images, by (3) ⇒ (1). By Lemma 4.14 it follows that $\mathfrak{F}$ is also a free structure of $\mathcal{H}$. Since free structures are uniquely determined by the cardinality of their set of generators we can conclude that $\mathcal{K}$ contains all free structures of $\mathcal{H}$. Since, by

Lemma 4.16 (a), every structure of $\mathcal{H}$ is a weak homomorphic image of a free structure and $\mathcal{K}$ is closed under weak homomorphic images it follows that $\mathcal{H} \subseteq \mathcal{K}$.

It remains to consider the case that $\mathcal{K}$ only contains structures with at most 1 element. Then $\forall x \forall y(x = y) \in T$ and all structures of $\mathcal{H}$ contain at most 1 element. Since each such structure can be described up to isomorphism by formulae of the form $\forall \bar{x} \varphi$ it follows that $\mathcal{H} = \mathcal{K}$. □

5. The Theorem of Feferman and Vaught

In general, first-order formulae are not preserved in products. Nevertheless the first-order theories of products are well behaved. We will prove below that the first-order theory of a product can be computed from the first-order theories of its factors. In fact, this result holds not only for ordinary direct products, but it can be extended to a quite general notion of a product.

Definition 5.1. Let S and T be disjoint sets of sorts, Σ an S-sorted signature, Γ a T-sorted one, and $\iota \in T$ a sort of T. Suppose that $(\mathfrak{A}^i)_{i \in I}$ is a sequence of Σ-structures and $\mathfrak{J}$ a Γ-structure whose domain of sort ι is $J_\iota = \wp(I)$. For $s \in S$, let $I_s := \{ i \in I \mid A^i_s \neq \varnothing \}$.

The *generalised product* of $(\mathfrak{A}^i)_i$ over $\mathfrak{J}$ is the structure

$$\prod_{i \in \mathfrak{J}} \mathfrak{A}^i := \left\langle U, \subseteq, E_=, (\zeta^{\mathfrak{J}})_{\zeta \in \Gamma}, (\xi')_{\xi \in \Sigma} \right\rangle,$$

with domains

$$U_s := \begin{cases} \prod_{i \in I_s} A^i_s & \text{for } s \in S\,, \\ J_s & \text{for } s \in T\,. \end{cases}$$

The relations and functions $\zeta^{\mathfrak{J}}$, for $\zeta \in \Gamma$, are taken from $\mathfrak{J}$, while the relations R', for $R \in \Sigma$, are defined by

$$R' := \left\{ \langle w, a_0, \dots, a_{n-1} \rangle \in \wp(I) \times U^n \;\middle|\; w = [\![R\bar{a}^i]\!]_{i \in I} \right\}.$$

As usual the functions f', for $f \in \Sigma$, are defined component wise

$$f'(\bar{a}) := (f^{\mathfrak{A}_i}(\bar{a}^i))_i \,.$$

Finally, $\subseteq$ is the subset relation on $J_\iota = \mathcal{P}(I)$ and

$$E_= := \big\{ \langle w, a, b \rangle \in \mathcal{P}(I) \times U^2 \bigm| w = [\![a^i = b^i]\!]_{i \in I} \big\} \,.$$

Example. (a) Let $(\mathfrak{A}^i)_{i \in I}$ be a sequence of structures and $\mathfrak{u}$ a filter on I. The reduced product $\prod_i \mathfrak{A}^i / \mathfrak{u}$ can be interpreted in the generalised product $\prod_{i \in \mathfrak{I}} \mathfrak{A}^i$ with index structure $\mathfrak{I} := \langle \mathcal{P}(I), \mathfrak{u} \rangle$. A relation R of $\prod_i \mathfrak{A}^i / \mathfrak{u}$ can be defined by the formula

$$\varphi_R(\bar{x}) := \exists z (Rz\bar{x} \wedge \mathfrak{u}z) \,.$$

(b) Suppose that $\mathfrak{G}_i = \langle V_i, E_i \rangle$, $i < 2$, are two directed graphs. Their *asynchronous product* is the graph $\mathfrak{H} = \langle V, E \rangle$ with universe $V := V_0 \times V_1$ and edge relation

$$E := (\mathrm{id}_{V_0} \times E_1) \cup (E_0 \times \mathrm{id}_{V_1}) \,.$$

We can interpret $\mathfrak{H}$ in the generalised product over the index structure $\mathfrak{I} := \langle \mathcal{P}[2] \rangle$ by the formula

$$\varphi_E(x, y) := \exists u \exists v [u \nsubseteq v \wedge v \nsubseteq u \wedge E_= uxy \wedge Evxy] \,,$$

which states that, for $x = \langle x_0, x_1 \rangle$ and $y = \langle y_0, y_1 \rangle$, there are sets $u = \{i\}$ and $v = \{k\}$ with $i \neq k$ such that $x_i = y_i$ and $\langle x_k, y_k \rangle \in E_k$.

Theorem 5.2 (Feferman-Vaught). *For every first-order formula $\varphi(\bar{x}, \bar{y})$, there exist a finite number of first-order formulae $\chi_0(\bar{x}), \dots, \chi_{m-1}(\bar{x})$ and $\psi(\bar{y}, \bar{z})$ such that,*

$$\prod_{i \in \mathfrak{I}} \mathfrak{A}^i \vDash \varphi(\bar{w}, \bar{a})$$

$$\text{iff} \quad \langle \mathfrak{I}, \subseteq \rangle \vDash \psi\big(\bar{w}, [\![\chi_0(\bar{a}^i)]\!]_i, \dots, [\![\chi_{m-1}(\bar{a}^i)]\!]_i\big) \,,$$

for all sequences $(\mathfrak{A}^i)_{i \in I}$, index structures $\mathfrak{I}$, and tuples $\bar{a} \subseteq \prod_i A^i$ and $\bar{w} \subseteq J$.

Proof. We construct the formulae χ_i and ψ by induction on φ. If φ is an atomic formula whose free variables all range over J then we have

$$\prod_{i \in \mathfrak{I}} \mathfrak{A}^i \vDash \varphi(\bar{w}) \quad \text{iff} \quad \langle \mathfrak{I}, \subseteq \rangle \vDash \varphi(\bar{w}) .$$

If $\varphi = Rst_0 \dots t_{n-1}$ where $R \in \Sigma$ then we have

$$\prod_{i \in \mathfrak{I}} \mathfrak{A}^i \vDash \varphi(\bar{w}, \bar{a}) \quad \text{iff} \quad [\![Rt_0 \dots t_{n-1}[\bar{a}^i]]\!]_i = s^{\mathfrak{I}}[\bar{w}] .$$

Hence, we can set $\chi_0 := Rt_0 \dots t_{n-1}$ and $\psi := z_0 = s$.

Similarly, if $\varphi = E_= st_0 t_1$ then we define $\chi_0 := t_0 = t_1$ and $\psi := z_0 = s$. If φ is a boolean combination then we can take the corresponding boolean combination of the formulae obtained by inductive hypothesis.

Hence, it remains to consider the case that $\varphi = \exists z \varphi'(\bar{x}, \bar{y}, z)$. Let $\chi'_0, \dots, \chi'_{m-1}$ and ψ' be the formulae for φ' obtained from the inductive hypothesis. If z ranges over J then we have

$$\prod_{i \in \mathfrak{I}} \mathfrak{A}^i \vDash \varphi(\bar{w}, \bar{a})$$

$$\begin{aligned} \text{iff} \quad & \text{there is some } w' \in J \text{ with} \\ & \langle \mathfrak{I}, \subseteq \rangle \vDash \psi'\big(\bar{w}, w', [\![\chi'_0(\bar{a}^i)]\!]_i, \dots, [\![\chi'_{m-1}(\bar{a}^i)]\!]_i\big) \\ \text{iff} \quad & \langle \mathfrak{I}, \subseteq \rangle \vDash \exists z' \psi'\big(\bar{w}, z', [\![\chi'_0(\bar{a}^i)]\!]_i, \dots, [\![\chi'_{m-1}(\bar{a}^i)]\!]_i\big) . \end{aligned}$$

If, on the other hand, z ranges over sequences in $\prod_i A^i$ then we proceed as follows. As φ only mentions finitely many symbols of the signature we may assume that the signature is finite. Therefore, every first-order formula can be written as a finite disjunction of Hintikka-formulae. Let r be the maximal quantifier rank of the formulae χ'_l, $l < m$, and let $\chi''_0, \dots, \chi''_{p-1}$ be an enumeration of all Hintikka-formulae of this quantifier rank. We can find a formula ψ'' such that

$$\begin{aligned} & \langle \mathfrak{I}, \subseteq \rangle \vDash \psi'\big(\bar{w}, [\![\chi'_0(\bar{a}^i, b^i)]\!]_i, \dots, [\![\chi'_{m-1}(\bar{a}^i, b^i)]\!]_i\big) \\ \text{iff} \quad & \langle \mathfrak{I}, \subseteq \rangle \vDash \psi''\big(\bar{w}, [\![\chi''_0(\bar{a}^i, b^i)]\!]_i, \dots, [\![\chi''_{p-1}(\bar{a}^i, b^i)]\!]_i\big) . \end{aligned}$$

Therefore, we may w.l.o.g. assume that, for all elements $\bar{a}$ and b, the sets

$$[\![\chi'_0(\bar{a}^i, b^i)]\!]_i, \dots, [\![\chi'_{m-1}(\bar{a}^i, b^i)]\!]_i$$

form a partition of I. For $s \subseteq [m]$, let

$$\chi_s(\bar{x}) := \bigwedge_{l \in s} \exists z \chi'_l(\bar{x}, z) \wedge \forall z \bigvee_{l \in s} \chi'_l(\bar{x}, z),$$

and define

$$\psi(\bar{y}, \bar{z}) := \exists u_0 \cdots \exists u_{m-1} \Big(\text{“}u_0, \dots, u_{m-1} \text{ form a partition of } I\text{”} \\ \wedge \psi'(\bar{y}, \bar{u}) \wedge \bigwedge_{l<m} u_l \subseteq \bigcup_{s \ni l} z_s \Big).$$

We claim that the formulae ψ and χ_s, for $s \subseteq [m]$, have the desired properties. Note that

$$k \in [\![\chi_s(\bar{a}^i)]\!]_i \quad \text{iff} \quad s = \big\{ l < m \bigm| k \in [\![\exists z \chi'_l(\bar{a}^i, z)]\!]_i \big\},$$

which implies that

$$\begin{aligned} k \in \bigcup_{s \ni l} [\![\chi_s(\bar{a}^i)]\!]_i \quad &\text{iff} \quad k \in [\![\chi_s(\bar{a}^i)]\!]_i \text{ for some } s \ni l \\ &\text{iff} \quad l \in \big\{ l < m \bigm| k \in [\![\exists z \chi'_l(\bar{a}^i, z)]\!]_i \big\} \\ &\text{iff} \quad k \in [\![\exists z \chi'_l(\bar{a}^i, z)]\!]_i . \end{aligned}$$

First, suppose that there is some $b \in \prod_i A^i$ with

$$\prod_{i \in \mathfrak{I}} \mathfrak{A}^i \vDash \varphi'(\bar{a}, \bar{w}, b).$$

Setting $u_l := [\![\chi'_l(\bar{a}^i, b^i)]\!]_i$ it follows by inductive hypothesis that

$$\langle \mathfrak{I}, \subseteq \rangle \vDash \psi'(\bar{w}, u_0, \dots, u_{m-1}).$$

Furthermore, $u_l \subseteq [\![\exists z \chi'_l(\bar{a}^i, z)]\!]_i$ which, by the above remark, implies that $u_l \subseteq \bigcup_{s \ni l} [\![\chi_s(\bar{a}^i)]\!]_i$. Since, by assumption, $u_0, \dots, u_{m-1}$ form a partition of I, it follows that

$$\langle \mathfrak{I}, \subseteq \rangle \vDash \psi\big(\bar{w}, ([\![\chi_s(\bar{a}^i)]\!]_i)_{s \subseteq [m]}\big)\,.$$

Conversely, suppose that

$$\langle \mathfrak{I}, \subseteq \rangle \vDash \psi\big(\bar{w}, ([\![\chi_s(\bar{a}^i)]\!]_i)_{s \subseteq [m]}\big)\,.$$

Then there are sets $u_l \subseteq [\![\exists z \chi'_l(\bar{a}^i, z)]\!]_i$, $l < m$, forming a partition of I such that

$$\langle \mathfrak{I}, \subseteq \rangle \vDash \psi'(\bar{w}, u_0, \dots, u_{m-1})\,.$$

For each $i \in u_l$, fix some element $b^i \in A^i$ with $\mathfrak{A}^i \vDash \chi'_l(\bar{a}^i, b^i)$. Since the u_l form a partition of I this defines an element $b \in \prod_i A^i$. By inductive hypothesis, we have

$$\prod_{i \in \mathfrak{I}} \mathfrak{A}^i \vDash \varphi'(\bar{a}, \bar{w}, b)\,.$$

□

Corollary 5.3. *Let $(\mathfrak{A}^i)_{i \in I}$ and $(\mathfrak{B}^i)_{i \in I}$ be two sequences of structures and suppose that $\mathfrak{I}$ is a suitable index structure.*

$$\mathfrak{A}^i \equiv \mathfrak{B}^i, \quad \textit{for all } i \in I, \quad \textit{implies} \quad \prod_{i \in \mathfrak{I}} \mathfrak{A}^i \equiv \prod_{i \in \mathfrak{I}} \mathfrak{B}^i\,.$$

D3. O-minimal structures

1. Ordered topological structures

In this chapter we study ordered algebraic structures where the definable relations have similar properties as those in real closed fields. We start with some general remarks concerning densely ordered structures and the order topology.

Definition 1.1. Let $\langle A, < \rangle$ be an open dense linear order.

(a) For convenience, we add to A a least element $-\infty$ and a greatest one $+\infty$. Let A_∞ denote the resulting order.

(b) An *interval* is a nonempty set of the form

$$\begin{aligned} & (a,b) := {\uparrow} a \cap {\downarrow} b\,, && [a,b) := {\Uparrow} a \cap {\downarrow} b\,, \\ \text{or} \quad & (a,b] := {\uparrow} a \cap {\Downarrow} b\,, && [a,b] := {\Uparrow} a \cap {\Downarrow} b\,, \end{aligned}$$

with $a, b \in A_\infty$. Intervals of the form (a,b) are called *open,* those of the form $[a,b]$ *closed.*

(c) For functions $f, g : D \to A_\infty$ with $D \subseteq A$, we define

$$\begin{aligned} f < g \quad & :\text{iff} \quad f(c) < g(c) \quad \text{for all } c \in D\,, \\ f \leq g \quad & :\text{iff} \quad f(c) \leq g(c) \quad \text{for all } c \in D\,, \end{aligned}$$

and we set

$$\begin{aligned} (f,g) &:= \{ \langle c, a \rangle \in D \times A \mid f(c) < a < g(c) \}\,, \\ [f,g] &:= \{ \langle c, a \rangle \in D \times A_\infty \mid f(c) \leq a \leq g(c) \}\,. \end{aligned}$$

(d) We equip A with the order topology and each product A^n with the corresponding product topology. For $\bar{a}, \bar{b} \in A^n$, we define

$$B(\bar{a}, \bar{b}) := (a_0, b_0) \times \cdots \times (a_{n-1}, b_{n-1}) \subseteq A^n.$$

Sets of this form are called *boxes.* Recall that the topological closure of a set $U \subseteq A$ is denoted by $\mathrm{cl}(U)$, its interior by $\mathrm{int}(U)$, and the boundary by ∂U.

Remark. For every $n < \omega$, the set of boxes forms an open base for the topology on A^n. This topology is Hausdorff.

Definition 1.2. A function $f : A \to B$ between linear orders is *monotone* if it is either increasing or decreasing. It is *strictly monotone* if it is strictly increasing or strictly decreasing.

The following lemma gives a criterion for a function defined on a direct product to be continuous. It will be used in Section 3.

Lemma 1.3. *Let $\mathfrak{X}$ be a topological space, $\langle A, < \rangle$ and $\langle B, < \rangle$ open dense linear orders, and $f : X \times A \to B$ a function such that*

(1) *for each $x \in X$, the function $f(x, \cdot) : A \to B$ is continuous and monotone, and*

(2) *for each $a \in A$, the function $f(\cdot, a) : X \to B$ is continuous.*

Then f is continuous.

Proof. Let $J \subseteq B$ be an open interval. To prove that $f^{-1}[J]$ is open we show that, for every pair $\langle x, a \rangle \in f^{-1}[J]$, there are open sets $O \subseteq X$ and $I \subseteq A$ with $\langle x, a \rangle \in O \times I$ and $f[O \times I] \subseteq J$.

By (1) there is an open interval $(b_0, b_1) \subseteq A$ with $a \in (b_0, b_1)$ such that $f[\{x\} \times (b_0, b_1)] \subseteq J$. We use (2) to obtain open sets $O_0, O_1 \subseteq X$ such that $f[O_i \times \{b_i\}] \subseteq J$, for $i < 2$. Let $O := O_0 \cap O_1$. We claim that $f[O \times (b_0, b_1)] \subseteq J$.

Let $y \in O$ and $b_0 < c < b_1$. By symmetry, we assume that the function $f(y, \cdot) : A \to B$ is increasing. Then $f(y, b_0) \leq f(y, c) \leq f(y, b_1)$. Since $f(y, b_0), f(y, b_1) \in J$, this implies that $f(y, c) \in J$. ☐

We investigate the structure of definable relations in ordered structures. Throughout this chapter we will work with definitions with parameters.

Definition 1.4. Let $\mathfrak{A}$ be a structure.

(a) A relation $R \subseteq A^n$ is *parameter-definable* if there exists a first-order formula $\varphi(\bar{x};\bar{y})$ and parameters $\bar{c} \subseteq A$ such that $R = \varphi(\bar{x};\bar{c})^{\mathfrak{A}}$.

(b) A topology $\mathcal{C}$ on $\mathfrak{A}$ is *definable* if there exists a first-order formula $\varphi(x,\bar{y};\bar{z})$ and parameters $\bar{c} \subseteq A$ such that the family $(\varphi(x,\bar{a};\bar{c})^{\mathfrak{A}})_{\bar{a}\subseteq A}$ is a base of $\mathcal{C}$.

Lemma 1.5. *Let $\mathfrak{A} = \langle A, < \rangle$ be an open dense linear order and $n < \omega$.*

(a) *There exists a formula $\beta(\bar{x};\bar{y},\bar{z})$ such that*

$$\mathfrak{A} \vDash \beta(\bar{c};\bar{a},\bar{b}) \quad \text{iff} \quad \bar{c} \in B(\bar{a},\bar{b})\,.$$

(b) *If $X \subseteq A^n$ is parameter-definable then so are* $\mathrm{cl}(X)$ *and* $\mathrm{int}(X)$.

(c) *If $X \subseteq Y \subseteq A^n$ are parameter-definable sets and X is open in Y then there exists a parameter-definable open set O such that $X = Y \cap O$.*

Proof. (a) Set

$$\beta(\bar{x};\bar{y},\bar{z}) := \bigwedge_{i<n} (y_i < x_i \wedge x_i < z_i)\,.$$

(b) Let $\varphi(\bar{x})$ be the formula defining X. By (a), there exists a formula expressing that $\bar{c} \in B(\bar{a},\bar{b})$. We can define $\mathrm{cl}(X)$ by the formula

$$\psi(\bar{x}) := \forall \bar{y}\bar{z}\big[\bar{x} \in B(\bar{y},\bar{z}) \rightarrow (\exists \bar{u} \in B(\bar{y},\bar{z}))\varphi(\bar{u})\big]\,,$$

which expresses that every neighbourhood of $\bar{x}$ contains a point of X. Similarly, we can define $\mathrm{int}(X)$ by

$$\vartheta(\bar{x}) := \exists \bar{y}\bar{z}\big[\bar{x} \in B(\bar{y},\bar{z}) \wedge (\forall \bar{u} \in B(\bar{y},\bar{z}))\varphi(\bar{u})\big]\,.$$

(c) Let $\varphi(\bar{x})$ and $\psi(\bar{x})$ be the formulae defining X and Y, respectively and set

$$O := \bigcup \{ B(\bar{a}, \bar{b}) \mid B(\bar{a}, \bar{b}) \cap Y \subseteq X \} .$$

Then O is an open set with $Y \cap O = X$. It can be defined by the formula

$$\vartheta(\bar{x}) := \exists \bar{y}\bar{z}[\bar{x} \in B(\bar{y}, \bar{z}) \wedge (\forall \bar{u} \in B(\bar{y}, \bar{z}))(\psi(\bar{u}) \rightarrow \varphi(\bar{u}))] . \quad \square$$

We have seen that every parameter-definable relation in a real closed field is given by a boolean combination of polynomial equations and inequalities. As a consequence these relations are structurally quite tame. The next definition isolates the combinatorial core responsible for this simplicity.

Definition 1.6. A structure $\mathfrak{A}$ is *o-minimal* if there exists a parameter-definable open dense linear order $<$ on A such that every parameter-definable subset $X \subseteq A$ is a finite union of singletons $\{a\}$ and open intervals (a, b) with $a, b \in A_\infty$.

In this chapter $<$ will always denote the order with respect to which the given structure is o-minimal.

Example. (a) Every open dense linear order $\langle A, < \rangle$ is o-minimal since these structures admit quantifier elimination.

(b) As already mentioned above, real closed fields are another prominent example of o-minimal structures. Because of quantifier elimination each parameter-definable set in such a field is a boolean combination of sets defined by polynomial inequalities. To see that a real closed field is o-minimal it is therefore sufficient to note that every inequality $p[x] > 0$ defines a finite union of open intervals.

Lemma 1.7. *Let $\mathfrak{A}$ be an o-minimal structure and $X \subseteq A$ parameter-definable.*

(a) $\inf X$ *and* $\sup X$ *exist in* A_∞.

(b) *∂X is finite. Let $a_1 < \cdots < a_{n-1}$ be an increasing enumeration of ∂X and set $a_0 := -\infty$ and $a_n := \infty$. Each interval (a_i, a_{i+1}), $0 \leq i < n$, is either contained in X or disjoint from X.*

Proof. By definition of o-minimality, X is of the form

$$X = (a_0, b_0) \cup \cdots \cup (a_{n-1}, b_{n-1}) \cup \{c_0, \dots, c_{m-1}\}.$$

Consequently,

$$\begin{aligned} &\sup X = \max\{b_0, \dots, b_{n-1}, c_0, \dots, c_{m-1}\} \\ \text{and}\quad &\inf X = \min\{a_0, \dots, a_{n-1}, c_0, \dots, c_{m-1}\} \end{aligned}$$

exist. For the second claim, note that

$$\partial X \subseteq \{a_0, \dots, a_{n-1}, b_0, \dots, b_{n-1}, c_0, \dots, c_{m-1}\}$$

is finite. W.l.o.g. we may assume that the decomposition of X has been chosen such that

$$\begin{aligned} &(a_i, b_i) \cap (a_k, b_k) = \varnothing, &&\text{for } i \neq k, \\ \text{and}\quad &c_i \notin (a_k, b_k), &&\text{for all } i, k. \end{aligned}$$

If $d < e$ are consecutive elements of an increasing enumeration of X then we either have

$$\begin{aligned} &d = a_i, && e = b_i, && \text{and} && (d, e) = (a_i, b_i) \subseteq X, \\ &d = b_i, && e = a_{i+1}, && \text{and} && (d, e) \cap X = \varnothing, \\ &d = b_i, && e = c_k, && \text{and} && (d, e) \cap X = \varnothing, \\ &d = c_i, && e = a_k, && \text{and} && (d, e) \cap X = \varnothing, \\ \text{or}\quad &d = c_i, && e = c_k, && \text{and} && (d, e) \cap X = \varnothing. \end{aligned}$$

□

Definition 1.8. Let $\langle A, < \rangle$ be an open dense linear order and $n < \omega$. A set $X \subseteq A^n$ is *definably connected* if it is parameter-definable and there is no partition $X = Y_0 \cup Y_1$ of X into two disjoint nonempty parameter-definable subsets $Y_0, Y_1 \subseteq X$ that are open in X.

Lemma 1.9. *Let $\mathfrak{A}$ be an o-minimal structure.*

(a) *A subset $X \subseteq A$ is definably connected if and only if it is either empty or a single interval.*

(b) *The image of a definably connected set $X \subseteq A^m$ under a continuous parameter-definable function $f : X \to A^n$ is definably connected.*

(c) *Let $X, Y \subseteq A^n$ be parameter-definable. If $X \subseteq Y \subseteq \mathrm{cl}(X)$ and X is definably connected then so is Y.*

(d) *If $X, Y \subseteq A^n$ are definably connected and $X \cap Y \neq \emptyset$ then $X \cup Y$ is definably connected.*

Proof. (a) By definition of o-minimality, X is of the form

$$X = (a_0, b_0) \cup \dots \cup (a_{n-1}, b_{n-1}) \cup \{c_0, \dots, c_{m-1}\},$$

where we assume that n and m are chosen minimal. If $n > 1$, or $n = 1$ and $m > 0$, then we can decompose X into the sets

$$Y_0 := (a_0, b_0)$$
$$Y_1 := (a_1, b_1) \cup \dots \cup (a_{n-1}, b_{n-1}) \cup \{c_0, \dots, c_{m-1}\}.$$

Similarly, if $n = 0$ and $m > 1$ then we can set $Y_0 := \{c_0\}$ and $Y_1 := \{c_1, \dots, c_{m-1}\}$. Consequently, the pair $\langle n, m\rangle$ can only take the values $\langle 0, 0\rangle$, $\langle 0, 1\rangle$, or $\langle 1, 0\rangle$. In the first case $X = \emptyset$ and, otherwise, X is an interval.

(b) Suppose that $f[X]$ is not definably connected. Let $Y_0 \cup Y_1 = f[X]$ be the corresponding decomposition. Then we obtain a decomposition $f^{-1}[Y_0] \cup f^{-1}[Y_1] = X$ of X into two disjoint nonempty parameter-definable open subsets. Hence, X is not definably connected.

(c) Suppose that Y is not definably connected and let $Z_0 \cup Z_1 = Y$ be the corresponding decomposition. The sets $Z_0 \cap X$ and $Z_1 \cap X$ are disjoint, parameter-definable, and open in X. If we can show that they are nonempty then the result follows. Fix $a \in Z_i \subseteq \mathrm{cl}(X)$ and an open set O such that $O \cap Y \subseteq Z_i$. Since $a \in \mathrm{cl}(X)$ it follows that $O \cap X \neq \emptyset$. Hence, there is some element $b \in O \cap X \subseteq (O \cap Y) \cap X \subseteq Z_i \cap X$.

(d) Suppose that $X \cup Y$ is not definably connected and let $Z_0 \cup Z_1 = X \cup Y$ be a corresponding decomposition. If $Z_0 \cap X \neq \varnothing$ and $Z_1 \cap X \neq \varnothing$ then $Z_0 \cap X$ and $Z_1 \cap X$ witness the fact that X is not definably connected. Suppose that $Z_0 \cap X = \varnothing$, i.e., $X \subseteq Z_1$. Then we have $Y \cap Z_1 \supseteq (X \cap Y) \cap Z_1 = X \cap Y \neq \varnothing$ and $Y \cap Z_0 = Z_0 \neq \varnothing$. Consequently, Y is not definably connected. □

Corollary 1.10. *Let $\mathfrak{A}$ be an o-minimal structure and $f : [a, b] \to A$ parameter-definable and continuous. Then* $\operatorname{rng} f$ *contains every element between $f(a)$ and $f(b)$.*

2. *O-minimal groups and rings*

Before continuing to develop the theory of o-minimal structures let us give examples of o-minimal structures from algebra. We consider groups and rings.

Proposition 2.1. *Let $\mathfrak{M}$ be an o-minimal structure and suppose that $\cdot$ is a parameter-definable operation such that $\mathfrak{G} := \langle M, \cdot, < \rangle$ forms an ordered group.*

(a) *The only parameter-definable subgroups of $\mathfrak{G}$ are $\{e\}$ and M.*

(b) *$\mathfrak{G}$ is abelian, divisible, and torsion-free.*

Proof. (a) Let $H \subset M$ be a parameter-definable proper subgroup of $\mathfrak{G}$. First, we show that H is convex. For a contradiction, suppose otherwise. Then there are elements $h \in H$ and $a \in M \setminus H$ with $e < a < h$. This implies that $h^n < ah^n < h^{n+1}$, for all n. Consequently, we obtain a strictly increasing sequence

$$e < a < h < ah < h^2 < ah^2 < h^3 < \dots$$

where every second element belongs to H while the other elements belong to $M \setminus H$. Hence, H cannot be written as a finite union of intervals. A contradiction.

By Lemma 1.7, the supremum $c := \sup H$ exists. Because H is convex it follows that $(e, c) \subseteq H$. Suppose that $c > e$ and let $h \in (e, c)$. Then $h < c$ implies $e < h^{-1}c$ and $e < h$ implies $h^{-1} < e$ and $h^{-1}c < c$. Hence, $h^{-1}c \in (e, c) \subseteq H$ and it follows that $c = hh^{-1}c \in H$. Thus, we have $c < ch \in H$, in contradiction to the choice of c. Consequently, we have $c = e$ and $H = \{e\}$.

(b) We have already shown in Lemma D1.4.5 that all ordered groups are torsion-free.

For every $a \in M$, the centraliser $C(a) := \{ x \in M \mid ax = xa \}$ is a parameter-definable subgroup of $\mathfrak{G}$. Since $a \in C(a)$ it follows by (a) that $C(a) = M$. Consequently, every element a commutes with all other elements and $\mathfrak{G}$ is abelian.

Analogously, for $1 < n < \omega$, we can consider the non-trivial parameter-definable subgroup $D_n := \{ a^n \mid a \in M \}$. By (a), it follows that $D_n = M$. Hence, for every $a \in M$ there is some $b \in M$ with $a = b^n$. Consequently, $\mathfrak{G}$ is divisible. □

Theorem 2.2. *An ordered group $\mathfrak{G}$ is o-minimal if and only if it is abelian, divisible, and torsion-free.*

Proof. $(\Rightarrow)$ was already shown in Proposition 2.1. For $(\Leftarrow)$, suppose that $\mathfrak{G} = \langle G, +, -, 0, < \rangle$ is a model of ODAG. We have seen in Theorem D1.4.16 that this theory admits quantifier elimination. Hence, every parameter-definable subset $X \subseteq G$ is given as a boolean combination of inequalities $x < a$, for $a \in G$. It follows that X can be written as a finite union of intervals. □

Theorem 2.3. *Let $\mathfrak{A}$ be an o-minimal structure and suppose that $+$ and $\cdot$ are parameter-definable operations such that $\langle A, +, \cdot, < \rangle$ forms an ordered ring. Then $\langle A, +, \cdot, < \rangle$ is a real closed field.*

Proof. For every $a \in A$, there exists the parameter-definable additive subgroup $aA := \{ ax \mid x \in A \}$. If $a \neq 0$ then $a \in aA$ implies, by Proposition 2.1 (a), that $aA = A$. In particular, there is some element $b \in A$ with $ab = 1$.

Let $P := \{ a \in A \mid a > 0 \}$. Then P is closed under multiplication and, hence, forms an ordered group $\langle P, \cdot, < \rangle$. By Proposition 2.1 (b), it follows that this group is abelian. Since, for every element $a \in A$, we have $a \in P$, or $a = 0$, or $-a \in P$, it follows that $\cdot$ is commutative, for all elements of A. Consequently, $\langle A, +, \cdot, < \rangle$ is an ordered field.

It remains to prove that it is real closed. We use the characterisation of Proposition B6.6.17. Let $p \in A[x]$ be a polynomial over A. The corresponding polynomial function $A \to A : a \mapsto p[a]$ is parameter-definable. Suppose that $a < b$ are elements with $p[a] < 0 < p[b]$. By Corollary 1.10, there exists an element $c \in (a, b)$ with $p[c] = 0$. □

Corollary 2.4. *An ordered ring is o-minimal if and only if it is a real closed field.*

Besides real closed fields and models of ODAG, let us also mention the following example of an o-minimal structure.

Theorem 2.5 (Wilkie). *The structure* $\langle \mathbb{R}, +, \cdot, 0, 1, \exp \rangle$ *is o-minimal where* $\exp(x) := e^x$ *is the exponential function.*

3. *Cell decompositions*

In this section we prove an important structure result on parameter-definable relations in o-minimal structures. We will show that each such relation can be decomposed into finitely many 'simple' parts.

We start by considering binary relations $R \subseteq M^2$. The general theorem below will then follow by induction on the arity.

Lemma 3.1. *Let* $\mathfrak{M}$ *be o-minimal and* $f : (a, b) \to M$ *parameter-definable.*

(a) *There exist elements* $a \leq c < d \leq b$ *such that* $f \restriction (c, d)$ *is either constant or injective.*

(b) *If* f *is injective then there are elements* $a \leq c < d \leq b$ *such that* $f \restriction (c, d)$ *is strictly monotone.*

(c) *If f is strictly monotone then there are elements $a \leq c < d \leq b$ such that $f \upharpoonright (c, d)$ is continuous.*

Proof. (a) If there is some $x \in M$ such that $f^{-1}(x)$ is infinite then, being parameter-definable, $f^{-1}(x)$ contains an open interval (c, d). Hence, $f \upharpoonright (c, d)$ is constant.

It remains to consider the case that all sets $f^{-1}(x)$, $x \in M$, are finite. Then $f[(a, b)]$ is an infinite parameter-definable subset of M. Hence, it contains some interval I. We define a function $g : I \to (a, b)$ by

$$g(z) := \min \{ c \mid f(c) = z \} .$$

The function g is injective since it has a left-inverse f. As above, we can conclude that $g[I]$ is infinite and it contains an interval (c, d). Setting $J := f[(c, d)]$ it follows that the restriction $g \upharpoonright J : J \to (c, d)$ is surjective. Consequently, $g \upharpoonright J$ is a bijection between J and (c, d) and f is its inverse. In particular, $f \upharpoonright (c, d)$ is injective.

(b) Let $x \in (a, b)$. Since f is injective, we have a partition

$$\begin{aligned}(a, x) = \{ y \in (a, x) \mid f(y) < f(x) \} \\ \cup \{ y \in (a, x) \mid f(y) > f(x) \} .\end{aligned}$$

One of these two sets must contain an interval (c, x), for some $a < c < x$. The same holds for the interval (x, b). For $\sigma, \rho \in \{+, -\}$, define

$$\begin{aligned}\varphi_{\sigma\rho}(x) := \exists y \exists z \big[a < y < x < z < b \\ \wedge \forall u [y < u < x \to f(x) <^{\sigma} f(u)] \\ \wedge \forall u [x < u < z \to f(x) <^{\rho} f(u)] \big] ,\end{aligned}$$

where $<^+ := <$ and $<^- := >$. It follows that every $x \in (a, b)$ satisfies exactly one of the formulae φ_{++}, φ_{+-}, φ_{-+}, φ_{--}.

Consequently, (a, b) contains an open interval all elements of which satisfy the same formula. Replacing (a, b) by this interval we may assume that all elements of (a, b) satisfy the same formula. By symmetry, we may further assume that this formula is either φ_{-+} or φ_{++}.

First, suppose that all elements in (a,b) satisfy φ_{-+}. For $x \in (a,b)$, let

$$s(x) := \sup\left\{ s \in (x,b) \;\middle|\; f(x) < f(z) \text{ for all } z \in (x,s] \right\}.$$

Then we have $s(x) = b$ since $s(x) < b$ would contradict $\varphi_{-+}(s(x))$. Consequently, f is strictly increasing.

It remains to consider the case that all elements in (a,b) satisfy φ_{++}. Set

$$B := \{ x \in (a,b) \mid f(x) < f(z) \text{ for all } z \in (x,b) \}.$$

If B is infinite then it contains an open interval I. Hence, f is strictly increasing on I and we are done. Consequently, let us assume that B is finite. Replacing a by $\sup B$ we may assume that,

$(*)$ for every $x \in (a,b)$, there is some $x < y < b$ with $f(y) < f(x)$.

Fix $c \in (a,b)$. We claim that, for all sufficiently large elements $y \in (c,b)$, we have $f(y) < f(c)$. Otherwise, we would have $f(y) > f(c)$, for all sufficiently large $y \in (c,b)$. Let $d \in [c,b)$ be the minimal element such that $f(y) > f(c)$ for all $y \in (d,b)$. If $f(d) > f(c)$ then d would not be minimal since $\varphi_{++}(d)$ holds. Hence, $f(d) < f(c)$ and, by $(*)$, there is some $d < e < b$ such that $f(e) < f(d) < f(c)$. Contradiction.

Consequently, we have $f(y) < f(c)$, for all sufficiently large y. Set

$$y(c) := \inf\left\{ y \in [c,b) \;\middle|\; f(z) < f(c) \text{ for all } z \in (y,b) \right\}.$$

Then $\varphi_{++}(c)$ implies that $c < y(c)$ and $f(y(c)) < f(c)$. Minimality of $y(c)$ implies that $y(c)$ satisfies the following formula:

$$\begin{aligned} \psi_{\searrow}(y) := \exists uv\big[& a < u < y < v < b \\ & \wedge \forall st[u < s < y < t < v \rightarrow f(s) > f(t)]\big]. \end{aligned}$$

Since c was arbitrary it follows that, for every element $c \in (a,b)$, there is some $y \in (c,b)$ satisfying $\psi_{\searrow}$.

Therefore, there is an interval $(d, b) \subseteq (a, b)$ such that $\psi_{\searrow}$ holds for all $y \in (d, b)$. Replacing a by d we may assume that all elements of (a, b) satisfy this formula.

Let $\psi_{\nearrow}$ be the formula obtained from $\psi_{\searrow}$ by replacing the inequality $f(s) > f(t)$ by $f(s) < f(t)$. An analogous argument shows that we may assume that every element of (a, b) satisfies $\psi_{\nearrow}$. But no element can simultaneously satisfy $\psi_{\searrow}$ and $\psi_{\nearrow}$. Contradiction.

(c) By symmetry, we may assume that f is strictly increasing. Since $\operatorname{rng} f$ is infinite it contains an open interval $I \subseteq \operatorname{rng} f$. Choose elements $x < y$ in I and set $c := f^{-1}(x)$ and $d := f^{-1}(y)$. Then f induces an order-preserving bijection $(c, d) \to (x, y)$. Every order-isomorphism is continuous since the topology is defined in terms of the order. Consequently, $f \upharpoonright (c, d)$ is continuous. $\square$

Theorem 3.2 (Monotonicity Theorem). *Let $\mathfrak{M}$ be o-minimal and $f : (a, b) \to M$ parameter-definable. There exist elements*

$$a = a_0 < a_1 < \dots < a_n = b$$

such that, for every $i < n$, the restriction $f \upharpoonright (a_i, a_{i+1})$ is either constant, or strictly monotone and continuous.

Proof. Let X be the set of all elements $x \in (a, b)$ such that, for some $a \leq c < x < d \leq b$, the restriction $f \upharpoonright (c, d)$ is either constant, or strictly monotone and continuous. Note that $(a, b) \setminus X$ is finite since, otherwise, it would contain some interval I and we could use Lemma 3.1 to find an interval $I_0 \subseteq I$ such that $f \upharpoonright I_0$ is either constant, or strictly monotone and continuous. This would imply $I_0 \subseteq X$. A contradiction.

Let $b_1 < \dots < b_{m-1}$ be an enumeration of $(a, b) \setminus X$ and set $b_0 := a$ and $b_m := b$. It is sufficient to prove the theorem for $f \upharpoonright (b_i, b_{i+1})$. Hence, we may w.l.o.g. assume that $X = (a, b)$. There exist finitely many elements $a = a_0 < a_1 < \dots < a_{n-1} < a_n = b$ such that, for each interval (a_i, a_{i+1}), one of the following cases occurs:

(1) For all $x \in (a_i, a_{i+1})$, f is constant on some neighbourhood of x.

(2) For all $x \in (a_i, a_{i+1})$, f is strictly increasing on some neighbourhood of x.

(3) For all $x \in (a_i, a_{i+1})$, f is strictly decreasing on some neighbourhood of x.

We consider each case in turn.

(1) Fix some element $x \in (a_i, a_{i+1})$ and set

$$s := \sup \{ y \in (x, a_{i+1}) \mid f \text{ is constant on } [x, y) \} .$$

Then we have $s = a_{i+1}$ since, if $s < a_{i+1}$, then $s \in (a_i, a_{i+1})$ and f would be constant on some neighbourhood of s. A contradiction. Therefore, f is constant on $[x, a_{i+1})$. In the same way we can show that f is constant on $(a_i, x]$. Hence, it is constant on the whole interval (a_i, a_{i+1}).

(2) Fix some $x \in (a_i, a_{i+1})$ and set

$$s := \sup \{ y \in (x, a_{i+1}) \mid f \text{ is strictly increasing on } [x, y) \} .$$

As above, we have $s = a_{i+1}$ and f is strictly increasing on $[x, a_{i+1})$. Similarly, it is strictly increasing on $(a_i, x]$.

(3) This case follows in the same way as (2). □

Corollary 3.3. *Let $\mathfrak{M}$ be o-minimal and $f : (a, b) \to M$ parameter-definable.*

(a) *For every $c \in [a, b)$, the right sided limit* $\lim_{x \downarrow c} f(x)$ *exist in M_∞.*

(b) *For every $c \in (a, b]$, the left sided limit* $\lim_{x \uparrow c} f(x)$ *exist in M_∞.*

Corollary 3.4. *Let $\mathfrak{M}$ be o-minimal and $f : [a, b] \to M$ parameter-definable. Then f takes a maximum and a minimum value on $[a, b]$.*

The Cell Decomposition Theorem below is proved by an induction on the dimension. For the base case of this induction, we will need the following technical result.

Theorem 3.5. *Let $\mathfrak{M}$ be o-minimal and suppose that $R \subseteq M^2$ is a parameter-definable relation such that, for every $a \in M$, the fibre*

$$R_a := \{ b \in M \mid \langle a, b \rangle \in R \}$$

is finite. Then there is a constant $n < \omega$ such that $|R_a| \leq n$, for all $a \in M$.

Proof. We call a pair $\langle a, b \rangle \in M^2_\infty$ *generic* if there exist open intervals $I, J \subseteq M_\infty$ with $\langle a, b \rangle \in I \times J$ such that either

- $R \cap I \times J = \emptyset$, or
- $\langle a, b \rangle \in R$ and $R \cap I \times J$ is the graph of a continuous function $I \to M$.

(In this definition we consider intervals of the form $(c, \infty]$ and $[-\infty, c)$ as open.) Note that the sets

$$G_0 := \{ \langle a, b \rangle \in M^2 \mid \langle a, b \rangle \text{ is generic} \},$$
$$G_+ := \{ a \in M \mid \langle a, \infty \rangle \text{ is generic} \},$$
$$G_- := \{ a \in M \mid \langle -\infty, b \rangle \text{ is generic} \}$$

are parameter-definable. For $n < \omega$, let s_n be the (parameter-definable) function with

$$\operatorname{dom} s_n = \{ a \in M \mid |R_a| \geq n \}$$

such that $s_n(a) := b_n$ where $b_0 < b_1 < \cdots < b_n < \cdots$ is an enumeration of R_a.

For an element $a \in M$, let n be the maximal number such that the functions $s_0, \ldots, s_{n-1}$ are defined and continuous on some neighbourhood of a. We call *a normal* if $a \notin \operatorname{cl}(\operatorname{dom} s_n)$. Otherwise, a is *special.* Let N be the set of normal points and S the set of special ones. Note that, if a is normal and n is the number from above then there is some open neighbourhood U of a such that $\operatorname{dom} s_n$ is disjoint from U. This implies that

$$|R_x| = n, \text{ for all } x \in U, \text{ and } \langle a, b \rangle \text{ is generic, for all } b \in M_\infty.$$

We claim that N and S are definable. It is sufficient to show that, for every special element a, there is some $b \in M_\infty$ such that $\langle a, b \rangle$ is not generic. Let $a \in S$ and let n be the number from above. We define

$$\lambda_-(a) := \begin{cases} \lim_{x \uparrow a} s_n(x) & \text{if } (t, a) \subseteq \operatorname{dom} s_n\,, \text{ for some } t\,, \\ \infty & \text{otherwise}\,, \end{cases}$$

$$\lambda_0(a) := \begin{cases} s_n(x) & \text{if } a \in \operatorname{dom} s_n\,, \\ \infty & \text{otherwise}\,, \end{cases}$$

$$\lambda_+(a) := \begin{cases} \lim_{x \downarrow a} s_n(x) & \text{if } (a, t) \subseteq \operatorname{dom} s_n\,, \text{ for some } t\,, \\ \infty & \text{otherwise}\,, \end{cases}$$

and $\quad \beta(a) := \min \{ \lambda_-(a),\ \lambda_0(a),\ \lambda_+(a) \}\,.$

It follows that $\beta(a)$ is the least element $b \in M_\infty$ such that $\langle a, b \rangle$ is not generic.

To conclude the proof of the theorem we distinguish two cases. First, suppose that S is finite. Let $a_1 < \dots < a_{k-1}$ be an enumeration of S and set $a_0 := -\infty$ and $a_k := \infty$. We claim that $|R_x|$ is constant on each interval (a_i, a_{i+1}). Let

$$F_n := \{ x \in (a_i, a_{i+1}) \mid |R_x| = n \}\,.$$

Since $|R_x|$ is constant on an open neighbourhood of each element $a \in N$ it follows that the sets F_n are open. As (a_i, a_{i+1}) is connected this implies that there is some n such that $F_n = (a_i, a_{i+1})$.

It remains to consider the case that S is infinite. Let

$$\begin{aligned} S_- &:= \{ a \in S \mid \langle a, b \rangle \in R \text{ for some } b < \beta(a) \}\,, \\ S_+ &:= \{ a \in S \mid \langle a, b \rangle \in R \text{ for some } b > \beta(a) \}\,, \\ \beta_-(a) &:= \max \{ b \in R_a \mid b < \beta(a) \}\,, \\ \beta_+(a) &:= \max \{ b \in R_a \mid b > \beta(a) \}\,. \end{aligned}$$

At least one of the sets $S_- \cap S_+$, $S_- \setminus S_+$, $S_+ \setminus S_-$, $S \setminus (S_- \cup S_+)$ is infinite.

Let us consider the case that $S_- \cap S_+$ is infinite. As β_-, β, β_+ are parameter-definable we can use the Monotonicity Theorem to find an open interval $I \subseteq S_- \cap S_+$ on which each of these functions is continuous. Note that $\beta_- < \beta < \beta_+$. We can partition I as

$$I = \{ a \in I \mid \langle a, \beta(a) \rangle \in R \} \cup \{ a \in I \mid \langle a, \beta(a) \rangle \notin R \} .$$

One of these two sets contains an open interval I_0. Hence, we have either $\beta \restriction I_0 \subseteq R$ or $\beta \restriction I_0 \cap R = \emptyset$. In both cases it follows that $\beta \restriction I_0 \subseteq G_0$ since $\beta_- \restriction I_0$, $\beta \restriction I_0$, and $\beta_+ \restriction I_0$ are continuous. But $\langle a, \beta(a) \rangle$ is never generic. Contradiction.

In a similar way one can show that the remaining three cases also lead to contradictions. □

In the preceding proof we have used the observation that the elements of a fibre R_a depend continuously on a. This is a consequence of the Monotonicity Theorem. Since this situation will occur several times in the following, we introduce some terminology.

Definition 3.6. Let $\mathfrak{M}$ be an ordered structure.

(a) For $D \subseteq M^n$, we denote by $\mathrm{Cn}(D)$ the set of all parameter-definable continuous functions $D \to M$. Furthermore, we set

$$\mathrm{Cn}_\infty(D) := \mathrm{Cn}(D) \cup \{-\infty, \infty\} ,$$

where we regard $-\infty$ and ∞ as the constant functions with the respective value.

(b) Let $R \subseteq M^{n+1}$ be a relation and suppose that $D \subseteq M^n$ is a set such that every fibre $R_{\bar{a}}$ with $\bar{a} \in D$ contains exactly k elements. We say that a family of parameter-definable functions $s_0, \dots, s_{k-1} : D \to M$ is a *local enumeration* of R over D if

$$s_0 < \dots < s_{k-1} \quad \text{and} \quad R_{\bar{a}} = \{s_0(\bar{a}), \dots, s_{k-1}(\bar{a})\} , \quad \text{for } \bar{a} \in D .$$

Note that we can write the last condition also as

$$R \cap (D \times M) = s_0 \cup \dots \cup s_{k-1} .$$

A local enumeration $s_0, \dots, s_{k-1}$ is *continuous* if every s_i is continuous.

Corollary 3.7. *Let $R \subseteq M^2$ be a parameter-definable relation such that each fibre R_a, $a \in M$, is finite. There are finitely many elements*

$$-\infty = a_0 < a_1 < \cdots < a_{m-1} < a_m = \infty$$

such that over every interval (a_i, a_{i+1}) there exists a continuous local enumeration of R.

Proof. This follows immediately from the Monotonicity Theorem and Theorem 3.5. □

After having dealt with the case of binary relations, we turn to relations of larger arity. First, we define the 'simple parts' we want to decompose our relation into. These are generalisations of the notion of an interval to higher dimensions.

Definition 3.8. Let $\mathfrak{M}$ be an ordered structure.

(a) Let $\bar{\delta} \in [2]^n$. A *$\bar{\delta}$-cell* is a subset $C \subseteq M^n$ defined inductively as follows.

- The set M^0 is the unique $\langle\rangle$-cell.
- A $\bar{\delta}0$-cell is the graph of a function $f \in \mathrm{Cn}(D)$ where D is a $\bar{\delta}$-cell.
- A $\bar{\delta}1$-cell is a set of the form (f, g) where D is a $\bar{\delta}$-cell and $f, g \in \mathrm{Cn}_\infty(D)$ are functions with $f < g$.

A *cell* is a set that is a $\bar{\delta}$-cell for some $\bar{\delta}$. A cell is *open* if it is a $\langle 1, \dots, 1\rangle$-cell. (We also consider the $\langle\rangle$-cell as open.)

(b) The *dimension* of a $\bar{\delta}$-cell C is is the number

$$\dim C := \delta_0 + \cdots + \delta_{n-1}\,.$$

Lemma 3.9. *Let $C \subseteq M^n$ be a cell.*

(a) *If C is not open then it has empty interior.*

(b) *C is locally closed, i.e., there is an open set O with $C = \mathrm{cl}(C) \cap O$.*

(c) *C is homeomorphic to an open cell $D \subseteq M^{\dim C}$ via a parameter-definable homeomorphism $p : C \to D$.*

(d) *C is definably connected.*

Proof. (a) If $\mathrm{int}(C) \neq \varnothing$ then there is some box B with $B \subseteq C$. This implies that C is a $\langle 1, \dots, 1 \rangle$-cell.

(b) We prove the claim by induction on n. For $n = 0$, $C = M^0$ is clopen. Suppose that $n > 0$ and let $D := \pi(C) \subseteq M^{n-1}$ be the projection of C to M^{n-1}. By inductive hypothesis, D is locally closed. Hence, $\mathrm{cl}(D) \setminus D$ is a closed set. If C is the graph of a function $f \in \mathrm{Cn}(D)$ then

$$\mathrm{cl}(C) \setminus C \subseteq (\mathrm{cl}(D) \setminus D) \times M \,.$$

Hence, C is open in the closed set $C \cup (\mathrm{cl}(D) \setminus D) \times M$.

If $C = (f, g)$, for $f, g \in \mathrm{Cn}(D)$, then

$$\mathrm{cl}(C) \setminus C \subseteq f \cup g \cup (\mathrm{cl}(D) \setminus D) \times M \,.$$

As above it follows that C is locally closed.

The cases that $f = -\infty$ or $g = \infty$ follow analogously.

(c) Suppose that C is a $\bar{\delta}$-cell and let $i_0 < \dots < i_{k-1}$ be an enumeration of all indices i with $\delta_i = 1$. We define a map $p : M^n \to M^{\dim C}$ by

$$p(\bar{a}) := \langle a_{i_0}, \dots, a_{i_{k-1}} \rangle \,.$$

By induction on n, we prove that that p is a homeomorphism from C to an open cell $p[C] \subseteq M^{\dim C}$.

If C is open then $p = \mathrm{id}_C$ and there is nothing to do. Hence, suppose that C is not open. Then $n > 0$ and we can distinguish two cases.

If C is the graph of some function $f \in \mathrm{Cn}(D)$ then we can use the inductive hypothesis to obtain a homeomorphism $q : D \to q[D]$ from D to an open cell $q[D]$. Let $\pi : M^n \to M^{n-1}$ be the projection to the first $n-1$ coordinates. Then $\pi \restriction C : C \to D$ is a homeomorphism. Hence, so is $p = q \circ \pi \restriction C : C \to q[D]$.

It remains to consider the case that $C = (f, g)$, for $f, g \in \mathrm{Cn}_\infty(D)$. Then $p(\bar{a}b) = \langle q(\bar{a}), b \rangle$ where $q : D \to q[D]$ is the homeomorphism from the inductive hypothesis. Set $f' := f \circ q^{-1}$ and $g' := g \circ q^{-1}$. Then $f', g' \in \mathrm{Cn}_\infty(q[D])$ and $p : C \to (f', g')$ is a homeomorphism.

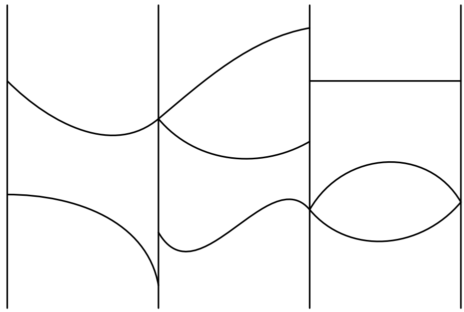


Figure 1.. A cell decomposition of $\mathbb{R}^2$.

(d) We proceed by induction on n. Clearly, M^0 is definably connected. Suppose that $n > 0$. By inductive hypothesis, the projection D of C to M^{n-1} is definably connected. For a contradiction, suppose that $C = O_0 \cup O_1$ where O_0 and O_1 are disjoint parameter-definable open sets. Since each fibre $\pi^{-1}(a) \cap C$ is definably connected we have $\pi^{-1}(a) \subseteq O_i$, for some i. Hence, there are sets $U_0, U_1 \subseteq M^{n-1}$ such that $O_i = \pi^{-1}[U_i] \cap C$. Clearly, U_0 and U_1 are open and parameter-definable. Since D is definably connected it follows that one of them is empty. □

We will show below that we can partition every definable relation into disjoint cells. In the same way we defined the notion of a cell by induction on the dimension, we also construct these partitions inductively.

Definition 3.10. (a) A *cell decomposition* of M^n is a partition $\mathcal{D}$ of M^n into finitely many pairwise disjoint cells where, for $n > 1$, we further require that the projection $\pi[\mathcal{D}]$ of $\mathcal{D}$ onto the first $n - 1$ components is a cell decomposition of M^{n-1}.

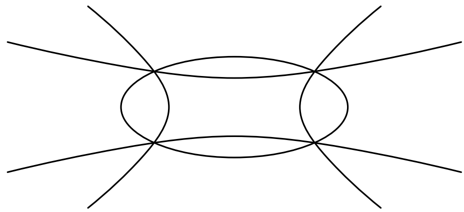
(b) A cell decomposition $\mathcal{D}$ *partitions* a relation $R \subseteq M^n$ if we have $R = C_0 \cup \dots \cup C_{k-1}$, for some cells $C_0, \dots, C_{k-1} \in \mathcal{D}$.

(c) A relation $R \subseteq M^{n+1}$ is *finite* over M^n if every fibre

$$R_{\bar{a}} := \{ b \in M \mid \bar{a}b \in R \}$$

is finite. We call R *uniformly finite* over M^n if there is a number $k < \omega$ such that $|R_{\bar{a}}| < k$, for all $\bar{a} \in M^n$.

Exercise 3.1. Find a cell decomposition of $\mathbb{R}^2$ partitioning the relation



which consists of all pairs $\langle x, y \rangle \in \mathbb{R}^2$ such that

$$\tfrac{4}{9}x^2 + \tfrac{9}{4}y^2 = 1 \quad \text{or} \quad \tfrac{4}{3}x^2 - \tfrac{9}{4}y^2 = 1 \quad \text{or} \quad \tfrac{27}{4}y^2 - \tfrac{4}{9}x^2 = 1\,.$$

Theorem 3.11 (Cell Decomposition Theorem). *Let $\mathfrak{M}$ be an o-minimal structure.*

(a) *For every finite family $R_0, \dots, R_{t-1} \subseteq M^n$ of parameter-definable relations there is a cell decomposition of M^n simultaneously partitioning each R_i.*

(b) *For every parameter-definable function $f : S \to M$ with $S \subseteq M^n$, there is a cell decomposition $\mathcal{D}$ of M^n partitioning S such that, for each cell $C \in \mathcal{D}$, the restriction $f \restriction C : C \to M$ is continuous.*

(c) *Every parameter-definable relation $R \subseteq M^n$ that is finite over M^{n-1} is uniformly finite.*

Proof. We prove all statements simultaneously by induction on n. Note that, for $n = 1$, (a) holds since $\mathfrak{M}$ is o-minimal, (b) follows from the Monotonicity Theorem, and (c) holds trivially.

For the inductive step, suppose that $n > 1$ and we have proved (a), (b), and (c) already for subsets of M^{n-1}.

We start by proving (c). We call a box $B \subseteq M^{n-1}$ *R-normal* if, for every point $\bar{a}b \in R$ with $\bar{a} \in B$, there exists an open interval I with $b \in I$ such that $R \cap (B \times I)$ is the graph of some continuous function

$f : B \to M$. (Note that this function f is then necessarily parameter-definable.) A point $\bar{a} \in M^{n-1}$ is called R-normal if it is contained in some R-normal box. Below we will establish the following claims.

(1) If B is R-normal then there exists a continuous local enumeration of R over B.

(2) If $S \subseteq M^{n-1}$ is definably connected and all elements of S are R-normal then there exists a continuous local enumeration of R over S.

(3) Every open cell $C \subseteq M^n$ contains an R-normal point.

First, let us show how (c) follows from (1)–(3). By inductive hypothesis, there exists a cell decomposition $\mathcal{D}$ of M^{n-1} partitioning the set of R-normal points. If a cell $C \in \mathcal{D}$ is open then, by (3), it contains an R-normal point. Hence, all points of C are R-normal and, by (2), there is a number $k(C)$ such that $|R_{\bar{a}}| < k(C)$, for all $\bar{a} \in C$. For cells $C \in \mathcal{D}$ that are not open, we can use Lemma 3.9 (c) to obtain similar bounds $k(C)$. Setting $k := \max \{ k(C) \mid C \in \mathcal{D} \}$ we obtain the desired bound on the size of $R_{\bar{a}}$. Hence, it remains to prove the claims.

(1) Fix $\bar{a} \in B$ and suppose that $b_0 < \dots < b_{k-1}$ is an enumeration of $R_{\bar{a}}$. Since B is R-normal we can find open intervals $I_0, \dots, I_{k-1}$ with $b_i \in I_i$ and continuous functions $s_0, \dots, s_{k-1} \in \mathrm{Cn}(B)$ such that

$$R \cap (B \times I_i) = s_i\,, \quad \text{for all } i < k\,.$$

We claim that $s_0, \dots, s_{k-1}$ is a local enumeration of R over B.

First, let us show that $s_0 < \dots < s_{k-1}$. For a contradiction, suppose that $s_i \nless s_{i+1}$. Since s_i and s_{i+1} are continuous this implies that there is some point $\bar{c} \in B$ with $s_i(\bar{c}) = s_{i+1}(\bar{c})$. In particular, $s_{i+1}(\bar{c}) \in I_i$. As s_{i+1} is continuous, there is a neighbourhood $U \subseteq B$ of $\bar{c}$ such that $s_{i+1}[U] \subseteq I_i$. Since $R \cap (B \times I_i) = s_i$, it follows that $s_{i+1} \upharpoonright U = s_i \upharpoonright U$. Thus, the set $\{ \bar{c} \in B \mid s_i(\bar{c}) = s_{i+1}(\bar{c}) \}$ is open. Since

$$\{ \bar{c} \in B \mid s_i(\bar{c}) < s_{i+1}(\bar{c}) \} \quad \text{and} \quad \{ \bar{c} \in B \mid s_i(\bar{c}) > s_{i+1}(\bar{c}) \}$$

are also open and B is definably connected it follows that $s_i = s_{i+1}$. But $s_i(\bar{a}) < s_{i+1}(\bar{a})$. A contradiction.

It remains to prove that $R \cap (B \times M) = s_0 \cup \cdots \cup s_{k-1}$. Let $\bar{b}c \in R \cap (B \times M)$. There exists a continuous function $f \in \mathrm{Cn}(B)$ with $f(\bar{b}) = c$ and $f \subseteq R$. In particular, $\langle \bar{a}, f(\bar{a}) \rangle \in R$. Hence, there is some index $i < k$ such that $f(\bar{a}) = b_i = s_i(\bar{a})$. As above, it follows that $f = s_i$.

(2) If S is empty there is nothing to do. Hence, we may assume that there is some $\bar{a} \in S$. Let $k := |R_{\bar{a}}|$. By (1), the set $\{ \bar{b} \in S \mid |R_{\bar{b}}| = k \}$ is clopen in S. This implies that $|R_{\bar{b}}| = k$, for all $\bar{b} \in S$. Consequently, we can find functions $s_0 < \cdots < s_{k-1}$ such that

$$R_{\bar{b}} = \{ s_0(\bar{b}), \dots, s_{k-1}(\bar{b}) \}, \quad \text{for } \bar{b} \in S.$$

It follows from (1) that each s_i is continuous.

(3) Let $B \subseteq C$ be a box. We will show that B contains an R-normal point. Suppose that $B = B_0 \times I$, for a box $B_0 \subseteq M^{n-2}$ and an open interval $I \subseteq M$. For $\bar{a} \in B_0$, we define

$$R(\bar{a}) := \{ \langle b, c \rangle \mid b \in I \text{ and } \bar{a}bc \in R \}.$$

Then $R(\bar{a})$ is finite over M. By Corollary 3.7 it follows that the set

$$\{ c \in M \mid c \text{ is not } R(\bar{a})\text{-normal} \}$$

is finite. Consequently, the set

$$S_B(R) := \{ \langle \bar{a}, b \rangle \in B \mid b \text{ is not } R(\bar{a})\text{-normal} \}$$

has empty interior. By inductive hypothesis, we can find a cell decomposition $\mathcal{D}$ of M^{n-1} partitioning B and $S_B(R)$. Let $C \in \mathcal{D}$ be an open cell with $C \subseteq B$. Then $C \cap S_B(R) = \emptyset$. Replacing B by a box contained in C we may assume that $S_B(R) = \emptyset$. We can apply (2) to $R(\bar{a})$ to find numbers $k(\bar{a}) < \omega$, for $\bar{a} \in B_0$, such that $|R_{\bar{a}b}| = k(\bar{a})$, for all $b \in I$.

We claim that there exists a bound k with $k(\bar{a}) \leq k$, for all $\bar{a}$. Fix $c \in I$ and define

$$R^c := \{ \langle \bar{a}, b \rangle \mid \langle \bar{a}, c, b \rangle \in R \}.$$

This set is finite over M^{n-2}. By inductive hypothesis, there exists a number m such that $|R^c_{\bar a}| < m$, for all $\bar a \in B_0$. Since $R^c_{\bar a} = R_{\bar a c}$ it follows that $|R_{\bar a c}| < m$. Consequently, we have $k(\bar a) \leq m$, for all $\bar a \in B_0$, which implies that $|R_{\bar a b}| < m$, for all $\bar a b \in B$, as desired.

We still have to find an R-normal element in B. For $k < m$, set

$$B_k := \{ \bar a \in B \mid |R_{\bar a}| = k \},$$

and let $s^k_0, \dots, s^k_{k-1} : B_k \to M$ be a local enumeration of $R_{\bar a}$ over B_k. By inductive hypothesis, we can find a cell decomposition $\mathcal{D}$ partitioning each set B_k such that, for every $C \in \mathcal{D}$, all restrictions $s^k_i \upharpoonright C$ are continuous. Since B is open and partitioned by $\mathcal{D}$ there exists an open cell $C \in \mathcal{D}$ with $C \subseteq B$. Fix k such that $C \subseteq B_k$. The functions $s^k_0, \dots, s^k_{k-1}$ are continuous on C. Consequently, each point of C is R-normal.

We prove (a) next. Let $R_0, \dots, R_{t-1} \subseteq M^n$ be parameter-definable and set

$$B := \partial_{n-1} R_0 \cup \dots \cup \partial_{n-1} R_{t-1},$$

where

$$\partial_{n-1} R := \{ \bar a b \in M^n \mid b \in \partial R_{\bar a} \}.$$

Note that B is finite over M^{n-1}. By (c), it follows that there is some bound $m < \omega$ such that $|B_{\bar a}| < m$, for all $\bar a \in M^{n-1}$. For $k < m$, let

$$B^k := \{ \bar a \mid |B_{\bar a}| = k \},$$

and let $s^k_1, \dots, s^k_k : B^k \to M$ be a local enumeration of $B_{\bar a}$ over B^k. We set $s^k_0 := -\infty$ and $s^k_{k+1} := \infty$. Finally, let

$$C_{lki} := \{ \bar a \in B^k \mid s^k_i(\bar a) \in (R_l)_{\bar a} \},$$
$$D_{lki} := \{ \bar a \in B^k \mid (s^k_i(\bar a), s^k_{i+1}(\bar a)) \subseteq (R_l)_{\bar a} \},$$

for $l < t$ and $0 \leq i \leq k \leq m$. By inductive hypothesis, there exists a cell decomposition $\mathcal{C}_0$ of M^{n-1} simultaneously partitioning the sets B^k, C_{lki}, and D_{lki}. By (b) we can choose a suitable refinement $\mathcal{C}$ of $\mathcal{C}_0$ such that, for every $C \in \mathcal{C}$ with $C \subseteq B^k$, the functions $s^k_1 \restriction C, \dots, s^k_k \restriction C$ are continuous.

For $C \in \mathcal{C}$ with $C \subseteq B^k$, we define a partition of $C \times M$ by

$$\mathcal{D}_C := \big\{ (s^k_i \restriction C, s^k_{i+1} \restriction C) \bigm| 0 \leq i < m \big\} \cup \big\{ s^k_i \restriction C \bigm| 0 < i < m \big\} .$$

The union $\mathcal{D} := \bigcup_{C \in \mathcal{C}} \mathcal{D}_C$ is the desired cell decomposition of M^n.

It remains to prove (b). Let $f : S \to M$ be parameter-definable with domain $S \subseteq M^n$. By (a), it is sufficient to show that we can find a partition $S = R_0 \cup \dots \cup R_{k-1}$ where each R_i is a parameter-definable set such that $f \restriction R_i$ is continuous. First, we can use (a) to partition S into finitely many cells. To find the desired partition of S it is sufficient to consider each of these cells separately. Hence, we may assume that S is a single cell.

If S is not open then we can use the definable homeomorphism $p : S \to p[S] \subseteq M^{\dim S}$ from Lemma 3.9 (c). By inductive hypothesis, we know that the image $p[S]$ can be partitioned into parameter-definable subsets $C_0, \dots, C_{k-1}$ such that all restrictions $(f \circ p^{-1}) \restriction C_i$ are continuous. Consequently, we can set $R_i := p^{-1}[C_i]$ to obtain the desired partition of S.

It remains to consider the case that S is an open cell. We call a point $\langle \bar{a}, b \rangle \in S$ *regular* if there exists a box $B \subseteq M^{n-1}$ and an open interval $I \subseteq M$ such that

(1) $\langle \bar{a}, b \rangle \in B \times I \subseteq S$,

(2) for every $\bar{c} \in B$, the function $f(\bar{c}, \cdot)$ is continuous and monotone on I,

(3) the function $f(\cdot, b)$ is continuous at $\bar{a}$.

Let $S_{\mathrm{reg}} \subseteq S$ be the set of all regular points. Note that S_{reg} is parameter-definable.

First, we prove that S_{reg} is dense in S. Let $B \subseteq M^{n-1}$ be a box and $I = (c, d) \subseteq M$ an interval such that $B \times I \subseteq S$. We have to show that $(B \times I) \cap S_{\mathrm{reg}} \neq \emptyset$. By the Monotonicity Theorem, we can find, for every $\bar{a} \in B$, a greatest element $\lambda(\bar{a}) \in (c, d]$ such that the function $f(\bar{a}, \cdot)$ is continuous and monotone on $(c, \lambda(\bar{a}))$. Since $\lambda : B \to M$ is parameter-definable we can use the inductive hypothesis to find a box $C_0 \subseteq B$ such that $\lambda \upharpoonright C_0$ is continuous. Fix elements $c < e < b < d$. We can find a cell $C_1 \subseteq C_0$ such that $\lambda(\bar{a}) \geq b$, for all $\bar{a} \in C_1$. By inductive hypothesis, there is a cell $C_2 \subseteq C_1$ such that $f(\,\cdot\,, e)$ is continuous on C_2. It follows that every point of $C_2 \times \{e\}$ is regular. Hence, $C_2 \times \{e\} \subseteq (B \times I) \cap S_{\mathrm{reg}} \neq \emptyset$, as desired.

By (a), we obtain a cell decomposition $\mathcal{D}$ partitioning both S and S_{reg}. We claim that $f \upharpoonright C$ is continuous, for every $C \in \mathcal{D}$ with $C \subseteq S$. Since S_{reg} is dense in S we have $S_{\mathrm{reg}} \cap C \neq \emptyset$, for such a cell C. This implies that $C \subseteq S_{\mathrm{reg}}$. Consequently, for each $\bar{a}b \in C$, the function $f(\,\cdot\,, b)$ is continuous at $\bar{a}$. It follows that C can be written as a union of boxes $B \times I$ that, for every $\langle \bar{a}, b \rangle \in B \times I$, satisfy conditions (1)–(3) above. Consequently, we can use Lemma 1.3 to conclude that f is continuous on each box $B \times I$. This implies that f is continuous on C. □

The Cell Decomposition Theorem has a number of important corollaries.

Proposition 3.12. *Let $R \subseteq M^m$ be a nonempty parameter-definable relation. Then R has only finitely many definably connected components. These components form a partition of R and each of them is clopen in R.*

Proof. Let $\mathcal{D}$ be a cell decomposition partitioning R and set

$$\mathcal{D}_0 := \{\, C \in \mathcal{D} \mid C \subseteq R \,\}.$$

Let $\mathcal{C}$ be a maximal subset of $\mathcal{D}_0$ such that $C := \bigcup \mathcal{C}$ is definably connected. We claim that every definably connected subset $S \subseteq R$ with $C \cap S \neq \emptyset$ is contained in C.

Let $\mathcal{D}_S := \{\, D \in \mathcal{D}_0 \mid D \cap S \neq \emptyset \,\}$. Then $S \subseteq \bigcup \mathcal{D}_S$. Since every cell is definably connected it follows that $\bigcup \mathcal{D}_S$ is definably connected. Furthermore, we have $C \cap \bigcup \mathcal{D}_S \supseteq C \cap S \neq \emptyset$. Hence, $C \cup \bigcup \mathcal{D}_S$ is also definably connected. By choice of $\mathcal{C}$ it follows that $\mathcal{D}_S \subseteq \mathcal{C}$. Hence, $S \subseteq \bigcup \mathcal{D}_S \subseteq C$, as desired.

We have shown that C is a definably connected component of R. It follows that we can partition R into definably connected components of the form $\bigcup \mathcal{C}$, for $\mathcal{C} \subseteq \mathcal{D}_0$. Since $\mathcal{D}_0$ is finite there are only finitely many such components.

Finally, note that the closure of a definably connected subset of R is also definably connected. Therefore, each definably connected component of R is closed in R. Since its complement is a finite union of closed sets it follows that each component is also open. □

Proposition 3.13. *Let $\mathfrak{M}$ be o-minimal and let $\pi : M^{m+n} \to M^m$ be the projection to the first m coordinates.*

(a) *For every cell $C \subseteq M^{m+n}$ and every point $\bar{a} \in \pi(C)$, the fibre $C_{\bar{a}}$ is a cell in M^n.*

(b) *For every cell decomposition $\mathcal{D}$ of M^{m+n} and every $\bar{a} \in M^m$, we obtain a cell decomposition*

$$\mathcal{D}_{\bar{a}} := \{ C_{\bar{a}} \mid C \in \mathcal{D},\ \bar{a} \in \pi(C) \}$$

of M^n.

Proof. (a) For $n = 1$, the fibre $C_{\bar{a}}$ is either a single point of an open interval. Hence, it is a cell. Suppose we have proved the claim already for $n - 1$ and let $C \subseteq M^{m+n}$. For $f \in \mathrm{Cn}(D)$, let $f_{\bar{a}} \in \mathrm{Cn}(D_{\bar{a}})$ be the function defined by $f_{\bar{a}}(x) := f(\bar{a}, x)$.

If C is the graph of a function $f \in \mathrm{Cn}(D)$ then $C_{\bar{a}}$ is the graph of $f_{\bar{a}}$. Similarly, if $C = (f, g)$, for $f, g \in \mathrm{Cn}_\infty(D)$, then $C_{\bar{a}} = (f_{\bar{a}}, g_{\bar{a}})$. Hence, $C_{\bar{a}}$ is again a cell.

(b) Clearly, $\mathcal{D}_{\bar{a}}$ is a finite partition of M^n. Therefore, the claim follows by (a). □

Corollary 3.14. *Let $R \subseteq M^m \times M^n$ be parameter-definable.*

(a) *There exists a number $k < \omega$ such that, for every $\bar{a} \in M^m$, the fibre $R_{\bar{a}} \subseteq M^n$ has a partition into at most k cells. In particular, each fibre $R_{\bar{a}}$ has at most k definably connected components.*

(b) *There exists a number $k < \omega$ such that, for every $\bar{a} \in M^m$, the fibre $R_{\bar{a}} \subseteq M^n$ has at most k isolated points. In particular, the size of every finite fibre $R_{\bar{a}}$ is bounded by k.*

Proof. (a) Let $\mathcal{D}$ be a cell decomposition of M^{m+n} partitioning R. For every $\bar{a} \in M^m$, the induced cell decomposition $\mathcal{D}_{\bar{a}}$ of M^n partitions R and it contains at most $|\mathcal{D}|$ cells. Hence, we can set $k := |\mathcal{D}|$.

(b) follows immediately from (a). □

Corollary 3.15. *Every o-minimal theory is graduated and, hence, admits elimination of $\exists^{\aleph_0}$.*

Proof. This follows by Theorem D1.2.15. □

An important consequence of the Cell Decomposition Theorem is the fact that whether a structure is o-minimal only depends on its first-order theory.

Theorem 3.16. *Let $\mathfrak{M}$ be an o-minimal structure. If $\mathfrak{N} \equiv \mathfrak{M}$ then $\mathfrak{N}$ is also o-minimal.*

Proof. Let $\varphi(x; \bar{y})$ be a first-order formula. We have to show that, for every choice of paramters $\bar{a} \subseteq N$, the set $\varphi(x; \bar{a})^{\mathfrak{N}}$ can be written as a finite union of intervals.

For $n < \omega$, let ψ_n be the first-order sentence stating that there are elements $\bar{a}$ such that $\varphi(x; \bar{a})$ is not a union of at most n intervals. By Theorem 3.11, there exists a number $m < \omega$ such that $\mathfrak{M} \nvDash \psi_m$. Hence, $\mathfrak{N} \nvDash \psi_m$ and every set of the form $\varphi(x; \bar{a})^{\mathfrak{N}}$ with $\bar{a} \subseteq N$ can be written as a union of at most m intervals. □

Part E.

Classical Model Theory

E1. Saturation

1. Homogeneous structures

Recall the relations $\sqsubseteq^\kappa_{\mathrm{FO}}$ introduced in Section C4.4. We have seen that, in general, they are not reflexive. In this section we will take a closer look at those structures $\mathfrak{A}$ that satisfy $\mathfrak{A} \cong^\kappa_{\mathrm{FO}} \mathfrak{A}$.

Definition 1.1. Let $\mathfrak{A}$ be a Σ-structure and κ a cardinal.

(a) $\mathfrak{A}$ is *κ-homogeneous* if $\mathfrak{A} \cong^\kappa_{\mathrm{FO}} \mathfrak{A}$, that is, whenever $\bar a, \bar b \in A^{<\kappa}$ are sequences of length less than κ with $\langle \mathfrak{A}, \bar a\rangle \equiv \langle \mathfrak{A}, \bar b\rangle$ and $c \in A$ is another element, then there exists an element $d \in A$ such that $\langle \mathfrak{A}, \bar a c\rangle \equiv \langle \mathfrak{A}, \bar b d\rangle$. We call $\mathfrak{A}$ *homogeneous* if it is $|A|$-homogeneous.

(b) $\mathfrak{A}$ is *strongly κ-homogeneous* if, whenever $\bar a, \bar b \in A^{<\kappa}$ are sequences of length less than κ with $\langle \mathfrak{A}, \bar a\rangle \equiv \langle \mathfrak{A}, \bar b\rangle$ then there exists an automorphism π of $\mathfrak{A}$ such that $\pi(\bar a) = \bar b$. We call $\mathfrak{A}$ *strongly homogeneous* if it is strongly $|A|$-homogeneous.

Example. (a) The structures $\langle \mathbb{Z}, <\rangle$ and $\langle \mathbb{Q}, <\rangle$ are strongly homogeneous.

(b) The theory of $\langle \omega, \leq\rangle$ has exactly three countable (strongly) homogeneous models whose order types are ω, $\omega + \zeta$, and $\omega + \zeta \cdot \eta$, respectively, where ζ is the order type of the integers and η is the order type of the rationals.

Exercise 1.1. Show that $\langle \mathbb{R}, +\rangle$ is strongly $\aleph_0$-homogeneous.

Lemma 1.2. *Every strongly κ-homogeneous structure is κ-homogeneous.*

Proof. Let $\mathfrak{A}$ be strongly κ-homogeneous. Suppose that $\bar a, \bar b \in A^{<\kappa}$ are sequences with $\langle \mathfrak{A}, \bar a\rangle \equiv \langle \mathfrak{A}, \bar b\rangle$ and let $c \in A$. By assumption, there exists an isomorphism $\pi : \langle \mathfrak{A}, \bar{a} \rangle \to \langle \mathfrak{A}, \bar{b} \rangle$. If we set $d := \pi(c)$ then we have

$$\pi : \langle \mathfrak{A}, \bar{a}c \rangle \cong \langle \mathfrak{A}, \bar{b}d \rangle .$$

This implies $\langle \mathfrak{A}, \bar{a}c \rangle \equiv \langle \mathfrak{A}, \bar{b}d \rangle$. □

Lemma 1.3. *Every homogeneous structure is strongly homogeneous.*

Proof. Let $\mathfrak{A}$ be a homogeneous structure of size $\kappa := |A|$. If $\bar{a}, \bar{b} \in A^{<\kappa}$ are sequences with $\langle \mathfrak{A}, \bar{a} \rangle \equiv \langle \mathfrak{A}, \bar{b} \rangle$ then $\mathfrak{A} \cong^{\kappa}_{\mathrm{FO}} \mathfrak{A}$ implies, by definition of $\cong^{\kappa}_{\mathrm{FO}}$, that

$$\langle \mathfrak{A}, \bar{a} \rangle \cong^{\kappa}_{\mathrm{FO}} \langle \mathfrak{A}, \bar{b} \rangle .$$

By Lemma C4.4.10, it follows that $\langle \mathfrak{A}, \bar{a} \rangle \cong \langle \mathfrak{A}, \bar{b} \rangle$. □

Lemma 1.4. *Let T be a first-order theory that admits quantifier elimination for* $\mathrm{FO}_{\infty\aleph_0}$. *Every model of T is $\aleph_0$-homogeneous.*

Proof. If $\mathfrak{A}$ is a model of T then we have $\mathfrak{A} \cong^{\aleph_0}_{\mathrm{o}} \mathfrak{A}$, by Theorem D1.2.9. This implies that $\mathfrak{A} \cong^{\aleph_0}_{\mathrm{FO}} \mathfrak{A}$. □

We have shown in Section C4.4 that $\cong^{\kappa}_{\mathrm{FO}}$ is an equivalence relation on the class of all κ-homogeneous structures. In the following lemmas we will study the corresponding equivalence classes. We will show that we have $\mathfrak{A} \cong^{\kappa}_{\mathrm{FO}} \mathfrak{B}$ if and only if both structures realise the same types.

Lemma 1.5. *Let $\mathfrak{B}$ be κ-homogeneous and suppose that $\mathfrak{A}$ is a structure such that, for all $n < \omega$, every n-type realised in $\mathfrak{A}$ is also realised in $\mathfrak{B}$. For each $\bar{a} \in A^{<\kappa}$, there exists a sequence $\bar{b} \in B^{<\kappa}$ such that*

$$\langle \mathfrak{A}, \bar{a} \rangle \equiv \langle \mathfrak{B}, \bar{b} \rangle .$$

Proof. Let $\bar{a} \in A^{\alpha}$, for $\alpha < \kappa$. We prove the statement by induction on α. If $\alpha < \omega$ then, since $\mathfrak{A}$ and $\mathfrak{B}$ realise the same α-types, we can find some tuple $\bar{b}$ with $\mathrm{tp}(\bar{b}/\mathfrak{B}) = \mathrm{tp}(\bar{a}/\mathfrak{A})$. If $\lambda := |\alpha| < \alpha$ then we can fix a bijection $g : \lambda \to \alpha$ and the claim follows if we apply the inductive hypothesis to the reordered sequence $(a_{g(i)})_{i<\lambda}$.

It therefore remains to consider the case that α is an infinite cardinal. We construct $(b_i)_{i<\alpha}$ by induction on i such that, at every step $\beta \leq \alpha$ we have

$$\langle \mathfrak{A}, (a_i)_{i<\beta} \rangle \equiv \langle \mathfrak{B}, (b_i)_{i<\beta} \rangle .$$

For $\beta = 0$, we have $\mathfrak{A} \equiv \mathfrak{B}$ since the unique complete 0-type $\mathrm{Th}(\mathfrak{A})$ realised in $\mathfrak{A}$ is also realised in $\mathfrak{B}$. If β is a limit ordinal then there is nothing to do. Suppose that $\beta = \gamma + 1$ is a successor and we have already defined $(b_i)_{i<\gamma}$. Since α is a limit we have $\beta < \alpha$. Therefore, we can apply the inductive hypothesis for α and it follows that there is some sequence $(c_i)_{i<\beta}$ such that

$$\langle \mathfrak{A}, (a_i)_{i<\beta} \rangle \equiv \langle \mathfrak{B}, (c_i)_{i<\beta} \rangle .$$

In particular, we have

$$\langle \mathfrak{B}, (b_i)_{i<\gamma} \rangle \equiv \langle \mathfrak{A}, (a_i)_{i<\gamma} \rangle \equiv \langle \mathfrak{B}, (c_i)_{i<\gamma} \rangle ,$$

and, since $\mathfrak{B}$ is κ-homogeneous, we can find some element $b_\gamma \in B$ such that

$$\langle \mathfrak{B}, (b_i)_{i<\gamma}, b_\gamma \rangle \equiv \langle \mathfrak{B}, (c_i)_{i<\gamma}, c_\gamma \rangle \equiv \langle \mathfrak{A}, (a_i)_{i<\gamma}, a_\gamma \rangle . \qquad \square$$

Proposition 1.6. *Let $\mathfrak{B}$ be κ-homogeneous and suppose that $\mathfrak{A}$ is a structure such that, for all $n < \omega$, every n-type realised in $\mathfrak{A}$ is also realised in $\mathfrak{B}$. Then $\mathfrak{A} \sqsubseteq^{\kappa}_{\mathrm{FO}} \mathfrak{B}$.*

Proof. Since $I^{\kappa}_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B})$ is always κ-complete we only need to prove the forth property. Let $\bar{a} \mapsto \bar{b} \in I^{\kappa}_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B})$ and $c \in A$. By the preceding lemma, we can find a sequence $\bar{b}'d' \subseteq B$ such that

$$\langle \mathfrak{A}, \bar{a}c \rangle \equiv \langle \mathfrak{B}, \bar{b}'d' \rangle .$$

In particular, we have $\langle \mathfrak{B}, \bar{b} \rangle \equiv \langle \mathfrak{B}, \bar{b}' \rangle$. Since $\mathfrak{B}$ is κ-homogeneous we can therefore find some element $d \in B$ such that

$$\langle \mathfrak{B}, \bar{b}d \rangle \equiv \langle \mathfrak{B}, \bar{b}'d' \rangle \equiv \langle \mathfrak{B}, \bar{b} \rangle .$$

Hence, $\bar{a}c \mapsto \bar{b}d \in I^{\kappa}_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B})$. □

Corollary 1.7. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be κ-homogeneous structures. We have*

$$\mathfrak{A} \cong^{\kappa}_{\mathrm{FO}} \mathfrak{B} \quad \text{iff} \quad \mathfrak{A} \text{ and } \mathfrak{B} \text{ realise the same } n\text{-types, for all } n < \omega.$$

Corollary 1.8. *If $\mathfrak{A}$ and $\mathfrak{B}$ are $\aleph_0$-homogeneous structures that realise the same n-types, for all $n < \omega$, and $\bar{a} \in A^{<\omega}$, $\bar{b} \in B^{<\omega}$ are finite tuples then*

$$\langle \mathfrak{A}, \bar{a} \rangle \equiv_{\mathrm{FO}} \langle \mathfrak{B}, \bar{b} \rangle \quad \text{implies} \quad \langle \mathfrak{A}, \bar{a} \rangle \equiv_{\infty} \langle \mathfrak{B}, \bar{b} \rangle .$$

Proof. This follows by Proposition 1.6 and Theorem D1.2.13. □

Theorem 1.9. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be homogeneous structures of the same size $|A| = |B|$. If, for every $n < \omega$, $\mathfrak{A}$ and $\mathfrak{B}$ realise the same n-types then $\mathfrak{A} \cong \mathfrak{B}$.*

Proof. Let $\kappa := |A| = |B|$. By Proposition 1.6, we have $\mathfrak{A} \sqsubseteq^{\kappa}_{\mathrm{FO}} \mathfrak{B}$ and $\mathfrak{A} \sqsupseteq^{\kappa}_{\mathrm{FO}} \mathfrak{B}$. Hence, the claim follows from Lemma C4.4.10 (a). □

Corollary 1.10. *A complete first-order theory T has, up to isomorphism, for every cardinal κ at most $2^{2^{|T|}}$ homogeneous models of size κ.*

Proof. For every set $X \subseteq S^{<\omega}(T)$, there is, according to the preceding theorem, at most one homogeneous model of size κ that realises exactly the types in X. Since $|S^{<\omega}(T)| \leq 2^{|T|}$ the claim follows. □

To build κ-homogeneous structures we can use the following lemma. We will defer the proof of the fact that every structure has a κ-homogeneous elementary extension to Section 3 where it will follow from a much stronger result.

Lemma 1.11. *Let $\mathfrak{A}$ be a Σ-structure and $\bar{a}, \bar{b} \in A^{\alpha}$ tuples with $\langle \mathfrak{A}, \bar{a} \rangle \equiv \langle \mathfrak{A}, \bar{b} \rangle$.*

(a) *There exists an elementary extension* $\mathfrak{B} \geq \mathfrak{A}$ *such that*

$$\langle \mathfrak{B}, \bar{a} \rangle \equiv_\infty \langle \mathfrak{B}, \bar{b} \rangle \quad \textit{and} \quad |B| \leq |A| \oplus |\Sigma| \oplus |\alpha| \oplus \aleph_0 .$$

(b) *There exists an elementary extension* $\mathfrak{B} \geq \mathfrak{A}$ *and an automorphism* $\pi \in \operatorname{Aut} \mathfrak{B}$ *with* $\pi(\bar{a}) = \bar{b}$.

Proof. (a) For $0 \leq k < \omega$, let I_k be a new $2k$-ary relation symbol and set

$$\beta_k := (\forall \bar{x} \bar{y}. I_k \bar{x} \bar{y})[\forall u \exists v I_{k+1} \bar{x} u \bar{y} v \wedge \forall v \exists u I_{k+1} \bar{x} u \bar{y} v] ,$$

and $\quad \psi_k^\varphi := (\forall \bar{x} \bar{y}. I_k \bar{x} \bar{y})[\varphi(\bar{a}, \bar{x}) \leftrightarrow \varphi(\bar{b}, \bar{y})] .$

The formula β_k says that I_k has the back-and-forth property with respect to I_{k+1}, and the ψ_k^φ hold if every tuple $\langle \bar{c}, \bar{d} \rangle \in I_k$ corresponds to a partial isomorphism $\bar{c} \mapsto \bar{d}$ from $\langle \mathfrak{A}, \bar{a} \rangle$ to $\langle \mathfrak{A}, \bar{b} \rangle$. Setting

$$\Phi := \operatorname{Th}(\mathfrak{A}_A) \cup \{I_0\} \cup \{\, \beta_k \wedge \psi_k^\varphi \mid k < \omega,\ \varphi \text{ an atomic formula} \,\} ,$$

we have

$$\mathfrak{B} \vDash \Phi \quad \text{iff} \quad \mathfrak{B} \geq \mathfrak{A} \quad \text{and} \quad \langle\rangle \mapsto \langle\rangle \in I_\infty(\langle \mathfrak{B}, \bar{a} \rangle, \langle \mathfrak{B}, \bar{b} \rangle) .$$

If Φ is satisfiable then we can, therefore, use the Theorem of Löwenheim and Skolem to find the desired structure $\mathfrak{B}$. To prove that Φ is satisfiable let $\Phi_0 \subseteq \Phi$ be finite. There is some $m < \omega$ and a finite set Δ of atomic formulae such that

$$\Phi_0 \subseteq \operatorname{Th}(\mathfrak{A}_A) \cup \{I_0\} \cup \{\, \beta_k \wedge \psi_k^\varphi \mid k < m, \varphi \in \Delta \,\} .$$

Let $\bar{a}'$ and $\bar{b}'$ be the subsequences of, respectively, $\bar{a}$ and $\bar{b}$ that appear in Δ. Since $\operatorname{tp}(\bar{a}') = \operatorname{tp}(\bar{b}')$ we can obtain a model $\langle \mathfrak{A}_A, (I_k)_{k<m} \rangle \vDash \Phi_0$ by setting

$$I_k := \left\{ \bar{c}\bar{d} \in A^{2k} \;\middle|\; \langle \mathfrak{A}, \bar{a}'\bar{c} \rangle \equiv_{m-k} \langle \mathfrak{A}, \bar{b}'\bar{d} \rangle \right\} .$$

(b) Let f be a new unary function symbol and set

$$\begin{aligned}\Phi := \operatorname{Th}(\mathfrak{A}_A) &\cup \{ f a_i = b_i \mid i < \alpha \} \\ &\cup \{\forall x \exists y f y = x\} \\ &\cup \{ \forall \bar{x}(\varphi(\bar{x}) \leftrightarrow \varphi(f\bar{x})) \mid \varphi \text{ an atomic formula} \}.\end{aligned}$$

If $\mathfrak{B} \models \Phi$ then $f^{\mathfrak{B}}$ is the desired automorphism. Therefore, it is sufficient to prove that Φ is satisfiable.

Let $\Phi_0 \subseteq \Phi$ be finite. There are finitely many indices $k_0, \dots, k_{n-1} < \alpha$, a finite set $C \subseteq A$, a finite signature $\Sigma_0 \subseteq \Sigma$, and a finite set Δ of atomic formulae over Σ_0 such that

$$\begin{aligned}\Phi_0 \subseteq \operatorname{Th}(\mathfrak{A}_C) &\cup \{ f a_{k_i} = b_{k_i} \mid i < n \} \\ &\cup \{\forall x \exists y f y = x\} \\ &\cup \{ \forall \bar{x}(\varphi(\bar{x}) \leftrightarrow \varphi(f\bar{x})) \mid \varphi \in \Delta \}.\end{aligned}$$

To simplify notation, set $\bar{a}' = a_{k_0} \dots a_{k_{n-1}}$ and $\bar{b}' = b_{k_0} \dots b_{k_{n-1}}$. By the Theorem of Löwenheim and Skolem, we can find a countable elementary substructure $\mathfrak{A}_0 \preceq \mathfrak{A}|_{\Sigma_0}$ with $C \cup \bar{a}'\bar{b}' \subseteq A_0$.

By (a), there exists a countable elementary extension $\mathfrak{B}_0 \succeq \mathfrak{A}_0$ such that

$$\langle \mathfrak{B}_0, \bar{a}' \rangle \equiv_\infty \langle \mathfrak{B}_0, \bar{b}' \rangle .$$

Hence, by Lemma C4.4.10, it follows that

$$\langle \mathfrak{B}_0, \bar{a}' \rangle \cong \langle \mathfrak{B}_0, \bar{b}' \rangle ,$$

and there is some automorphism $\pi \in \operatorname{Aut} \mathfrak{B}_0$ with $\pi(\bar{a}') = \bar{b}'$. Consequently, $\langle \mathfrak{B}_0, \pi \rangle$ is the desired model of Φ_0. □

Exercise 1.2. Let κ be an infinite cardinal. Prove that every structure has a κ-homogeneous elementary extension.

2. Saturated structures

We have shown in the previous section that κ-homogeneous structures can be ordered with respect to the set of types they realise. In this section we consider structures that are maximal in this ordering, i.e., homogeneous structures realising every type.

Definition 2.1. Let $\mathfrak{A}$ be a Σ-structure and κ a cardinal.

(a) $\mathfrak{A}$ is *κ-saturated* if, for all sets $C \subseteq A$ of size $|C| < \kappa$, every type $\mathfrak{p} \in S^{<\omega}(C)$ is realised in $\mathfrak{A}$. A structure $\mathfrak{A}$ is called *saturated* if it is $|A|$-saturated.

(b) $\mathfrak{A}$ is *κ-universal* if there exist elementary embeddings $\mathfrak{B} \to \mathfrak{A}$, for all Σ-structures $\mathfrak{B}$ of size $|B| < \kappa$ such that $\mathfrak{B} \equiv \mathfrak{A}$.

Similarly to homogeneous structures we can characterise κ-saturated structures in terms of the relation $\sqsubseteq^{\kappa}_{\mathrm{FO}}$.

Lemma 2.2. *A structure $\mathfrak{B}$ is κ-saturated if and only if*

$$\langle \mathfrak{A}, \bar{a} \rangle \equiv \langle \mathfrak{B}, \bar{b} \rangle \quad \textit{implies} \quad \langle \mathfrak{A}, \bar{a} \rangle \sqsubseteq^{\kappa}_{\mathrm{FO}} \langle \mathfrak{B}, \bar{b} \rangle ,$$

for all sequences $\bar{a} \in A^{<\kappa}$ and $\bar{b} \in B^{<\kappa}$.

Proof. ($\Rightarrow$) Suppose that $\langle \mathfrak{A}, \bar{a} \rangle \equiv \langle \mathfrak{B}, \bar{b} \rangle$. We have $\bar{a} \mapsto \bar{b} \in I^{\kappa}_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B})$ and $I^{\kappa}_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B})$ is κ-complete. Therefore, we only need to prove the forth property. Suppose that $\bar{c} \mapsto \bar{d} \in I^{\kappa}_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B})$ and $e \in A$. Set $\mathfrak{p} := \mathrm{tp}(e/\mathfrak{A}_{\bar{c}})$ and let $\mathfrak{q}$ be the type obtained from $\mathfrak{p}$ by replacing the constants $\bar{c}$ by $\bar{d}$. Note that $\mathfrak{q}$ really is a type since $\langle \mathfrak{A}, \bar{c} \rangle \equiv \langle \mathfrak{B}, \bar{d} \rangle$. As $|\bar{d}| < \kappa$ and $\mathfrak{B}$ is κ-saturated we can find some element $f \in B$ realising $\mathfrak{q}$. Therefore,

$$\langle \mathfrak{A}, \bar{c}e \rangle \equiv \langle \mathfrak{B}, \bar{d}f \rangle , \quad \text{that is,} \quad \bar{c}e \mapsto \bar{d}f \in I^{\kappa}_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B}) .$$

($\Leftarrow$) Let $C \subseteq B$ be a set of size $|C| < \kappa$ and $\mathfrak{p} \in S^n(C)$. There exists an elementary extension $\mathfrak{A} \succeq \mathfrak{B}$ in which $\mathfrak{p}$ is realised by some tuple $\bar{a}$. Let $\bar{c}$ be an enumeration of C. Since $\langle \mathfrak{A}, \bar{c} \rangle \equiv \langle \mathfrak{B}, \bar{c} \rangle$ we have

$$\langle \mathfrak{A}, \bar{c} \rangle \sqsubseteq^{\kappa}_{\mathrm{FO}} \langle \mathfrak{B}, \bar{c} \rangle .$$

Hence, by Lemma C4.4.9 we can find a tuple $\bar{b} \in B^n$ such that

$$\langle \mathfrak{A}, \bar{c}\bar{a} \rangle \sqsubseteq^{\kappa}_{\mathrm{FO}} \langle \mathfrak{B}, \bar{c}\bar{b} \rangle .$$

Consequently, $\bar{b}$ is a realisation of $\mathfrak{p}$ in B. ☐

Corollary 2.3. *For κ-saturated structures $\mathfrak{A}$ and $\mathfrak{B}$, we have*

$$\langle \mathfrak{A}, \bar{a} \rangle \cong^{\kappa}_{\mathrm{FO}} \langle \mathfrak{B}, \bar{b} \rangle \quad \text{iff} \quad \langle \mathfrak{A}, \bar{a} \rangle \equiv \langle \mathfrak{B}, \bar{b} \rangle ,$$

for all $\bar{a} \in A^{<\kappa}$ and $\bar{b} \in B^{<\kappa}$.

We will prove below that every κ-saturated structure is κ-homogeneous. Hence, the next corollary is a special case of Corollary 1.8.

Corollary 2.4. *If $\mathfrak{A}$ and $\mathfrak{B}$ are $\aleph_0$-saturated then*

$$\mathfrak{A} \equiv \mathfrak{B} \quad \textit{implies} \quad \mathfrak{A} \cong_{\infty} \mathfrak{B} .$$

For an example let us take a look at saturated linear orders.

Lemma 2.5. *Every $\aleph_1$-saturated dense linear order is incomplete.*

Proof. Let $a_0 < a_1 < \dots$ be a strictly increasing sequence of length ω and set $A := \{ a_n \mid n < \omega \}$. We claim that $\sup A$ does not exist. For a contradiction, suppose that the supremum c exists. Choose a type $\mathfrak{p}$ over $A \cup \{c\}$ containing the formulae

$$x < c \quad \text{and} \quad a_n < x \quad \text{for } n < \omega .$$

Any realisation b of $\mathfrak{p}$ is an upper bound of A. Hence, $b < c = \sup A$ yields the desired contradiction. ☐

Lemma 2.6. *A linear order is κ-saturated if, and only if, it is κ-dense.*

Proof. We have already shown in Lemma C4.4.6 that every κ-dense linear order is κ-saturated. For the converse, suppose that $\mathfrak{A} = \langle A, \leq \rangle$

is κ-saturated and let $C, D \subseteq A$ sets of size $|C|, |D| < \kappa$ with $C < D$. Let $\mathfrak{p} \in S^1(C \cup D)$ be any type with

$$\mathfrak{p} \supseteq \{ c < x \mid c \in C \} \cup \{ x < d \mid d \in D \} .$$

Since $\mathfrak{A}$ is κ-saturated there is some element $a \in A$ realising $\mathfrak{p}$. Hence, $C < a < D$ and $\mathfrak{A}$ is κ-dense. □

Lemma 2.7. *Let $(\mathfrak{A}^i)_{i<\lambda}$ be an elementary chain of κ-saturated structures. If $\kappa \leq \operatorname{cf} \lambda$ then the union $\bigcup_i \mathfrak{A}^i$ is also κ-saturated.*

Proof. Let $C \subseteq \bigcup_i A^i$ be a set of size $|C| < \kappa$ and suppose that $\mathfrak{p} \in S^{<\omega}(C)$ is a type over C. Since $|C| < \kappa \leq \operatorname{cf} \lambda$ there is some $\alpha < \lambda$ such that $C \subseteq A^\alpha$. Hence, there is a tuple $\bar{a} \subseteq A^\alpha \subseteq \bigcup_i A^i$ realising $\mathfrak{p}$. □

By definition a structure is κ-saturated if it realises every n-type, for $n < \omega$, with less than κ-parameters. In fact, it is sufficient to realise all 1-types.

Lemma 2.8. *Let $\kappa \geq \aleph_0$. A structure $\mathfrak{A}$ is κ-saturated if, and only if, whenever $C \subseteq A$ is of size $|C| < \kappa$ then every 1-type in $S^1(C)$ is realised in $\mathfrak{A}$.*

Exercise 2.1. Prove the preceding lemma.

Theorem 2.9. *Let $\mathfrak{A}$ be a Σ-structure. The following statements are equivalent:*

(1) *$\mathfrak{A}$ is κ-saturated.*

(2) *$\mathfrak{A}$ is κ-homogeneous and it realises every type in $S^\kappa(\varnothing)$.*

(3) *$\mathfrak{A}$ is κ-homogeneous and it realises every type in $S^{<\kappa}(\varnothing)$.*

If $\kappa \geq |\Sigma| \oplus \aleph_0$ then the following statement is also equivalent to the ones above.

(4) *$\mathfrak{A}$ is κ-homogeneous and κ^+-universal.*

Proof. (1) ⇒ (2) Let $\mathfrak{A}$ be κ-saturated. By Lemma 2.2, $\mathfrak{A} \equiv \mathfrak{A}$ implies $\mathfrak{A} \sqsubseteq^{\kappa}_{\mathrm{FO}} \mathfrak{A}$. Therefore, we have $\mathfrak{A} \cong^{\kappa}_{\mathrm{FO}} \mathfrak{A}$, that is, $\mathfrak{A}$ is κ-homogeneous.

It remains to prove that $\mathfrak{A}$ realises every type $\mathfrak{p} \in S^{\kappa}(\emptyset)$. For $\alpha < \kappa$, let $\mathfrak{p}_\alpha := \mathfrak{p} \cap \mathrm{FO}^{\alpha}[\Sigma]$ be the restriction of $\mathfrak{p}$ to the first α variables. By induction on α, we construct a sequence $(a_\alpha)_{\alpha<\kappa}$ such that the subsequence $(a_i)_{i<\alpha}$ realises $\mathfrak{p}_\alpha$. Suppose we have already defined a_i, for $i < \alpha$. Let

$$\mathfrak{q}_\alpha := \{ \varphi(a_{i_0}, \dots, a_{i_{k-1}}, x_\alpha) \mid \varphi(x_{i_0}, \dots, x_{i_{k-1}}, x_\alpha) \in \mathfrak{p} \text{ for } i_0, \dots, i_{k-1} < \alpha \}.$$

Since $\mathfrak{A}$ is κ-saturated we can find some element a_α such that

$$\mathrm{tp}(a_\alpha / \{ a_i \mid i < \alpha \}) = \mathfrak{q}_\alpha .$$

Hence, $(a_i)_{i\leq\alpha}$ realises $\mathfrak{p}_{\alpha+1}$.

(2) ⇒ (3) is trivial.

(3) ⇒ (1) Let $\mathfrak{p} \in S^n(U)$ where $|U| < \kappa$. Let $(c_i)_{i<\lambda}$ be an enumeration of U and let $\mathfrak{q} \in S^{\lambda+n}(\emptyset)$ be the type

$$\mathfrak{q} := \{ \varphi(x_{i_0}, \dots, x_{i_{k-1}}, x_\lambda, \dots, x_{\lambda+n-1}) \mid \varphi(c_{i_0}, \dots, c_{i_{k-1}}, x_0, \dots, x_{n-1}) \in \mathfrak{p} \}.$$

By assumption we can find sequences $\bar{a} \in A^\lambda$ and $\bar{b} \in A^n$ such that $\mathrm{tp}(\bar{a}\bar{b}) = \mathfrak{q}$. Since

$$\langle \mathfrak{A}, \bar{c} \rangle \equiv \langle \mathfrak{A}, \bar{a} \rangle$$

and $\mathfrak{A}$ is κ-homogeneous it follows that there is some tuple $\bar{d} \in A^n$ such that

$$\langle \mathfrak{A}, \bar{c}\bar{d} \rangle \equiv \langle \mathfrak{A}, \bar{a}\bar{b} \rangle .$$

Consequently $\mathrm{tp}(\bar{d}/\bar{c}) = \mathfrak{p}$.

(2) ⇒ (4) Suppose that $\mathfrak{A}$ realises every type in $S^{\kappa}(\emptyset)$. We claim that $\mathfrak{A}$ is κ^+-universal. Let $\mathfrak{B}$ be a structure of size $|B| \leq \kappa$ with $\mathfrak{B} \equiv \mathfrak{A}$. Choose

an enumeration $\bar{b}$ of B and let $\mathfrak{p} := \mathrm{tp}(\bar{b}/\mathfrak{B})$. Then $\mathfrak{p} \in S^{\leq\kappa}(\varnothing)$. Hence, there exists a sequence $\bar{a} \subseteq A$ realising $\mathfrak{p}$. The function $\bar{b} \mapsto \bar{a}$ is the desired elementary embedding.

(4) ⇒ (1) Suppose that $\mathfrak{A}$ is κ^+-universal. We show that $\mathfrak{A}$ realises every type $\mathfrak{p} \in S^\kappa(\varnothing)$. For each such $\mathfrak{p}$ we can find a structure $\mathfrak{B} \equiv \mathfrak{A}$ and a tuple $\bar{b} \subseteq B$ with $\mathrm{tp}(\bar{b}/\mathfrak{B}) = \mathfrak{p}$. By the Theorem of Löwenheim and Skolem we may assume that $|B| \leq \kappa$. Hence, there exists an elementary embedding $h : \mathfrak{B} \to \mathfrak{A}$. The sequence $h(\bar{b})$ is a realisation of $\mathfrak{p}$ in $\mathfrak{A}$. □

Theorem 2.10. *If $\mathfrak{A} \equiv \mathfrak{B}$ are saturated structures of the same size $|A| = |B|$ then $\mathfrak{A} \cong \mathfrak{B}$.*

Proof. Let $\kappa := |A| = |B|$. By Lemma 2.2, we have $\mathfrak{A} \cong^\kappa_{\mathrm{FO}} \mathfrak{B}$. Therefore, the claim follows from Lemma C4.4.10 (a). □

Every structure has a κ-saturated elementary extension. There are two ways to construct such extensions: (i) we can form an ultrapower, or (ii) we can take the union of an infinite elementary chain where each structure realises every type over the universe of the preceding structure. In the following proofs we will employ the first method. Below, where we construct saturated structures and projectively κ-saturated ones, we will choose the second method.

Proposition 2.11. *Let $\mathfrak{u}$ be a regular ultrafilter over an infinite set I and let $(\mathfrak{A}^i)_{i \in I}$ be a family of structures. Every countable partial type $\mathfrak{p}$ over $\prod_i A_i/\mathfrak{u}$ is realised in $\prod_i \mathfrak{A}_i/\mathfrak{u}$.*

Proof. Let $(\varphi_n)_{n<\omega}$ be an enumeration of $\mathfrak{p}$. Since $\mathfrak{u}$ is regular, we can find sets $(s_n)_{n<\omega}$ in $\mathfrak{u}$ such that, for every $i \in I$, the set

$$\{\, n < \omega \mid i \in s_n \,\}$$

is finite. Setting $w_n := s_0 \cap \dots \cap s_n \in \mathfrak{u}$ we obtain a strictly decreasing sequence $w_0 \supset w_1 \supset w_2 \supset \cdots$ of sets $w_n \in \mathfrak{u}$. By choice of $(s_n)_n$ we have

$$\bigcap_{n<\omega} w_n = \bigcap_{n<\omega} s_n = \varnothing \,.$$

Set $\psi_n := \varphi_0 \wedge \dots \wedge \varphi_n$ and let $[\bar{c}^n]_{\mathfrak{u}}$ be the parameters appearing in ψ_n. According to the Theorem of Łoś,

$$\prod_i \mathfrak{A}_i/\mathfrak{u} \vDash \exists \bar{x} \psi_n(\bar{x}; [\bar{c}^n]_{\mathfrak{u}}) \quad \text{implies} \quad [\![\exists \bar{x} \psi_n(\bar{x}; \bar{c}^n)]\!] \in \mathfrak{u} .$$

Hence, the sets

$$w_n^\circ := \{ i \in w_n \mid \mathfrak{A}_i \vDash \exists \bar{x} \psi_n(\bar{x}; \bar{c}_i^n) \} = w_n \cap [\![\exists \bar{x} \psi_n]\!]$$

are in $\mathfrak{u}$. We define a sequence $(\bar{a}_i)_{i \in I}$ as follows. If $i \notin w_0^\circ$, we choose an arbitrary tuple $\bar{a}_i \subseteq A_i$. Otherwise, let n be the maximal number such that $i \in w_n^\circ$ and let $\bar{a}_i \subseteq A_i$ be a tuple such that $\mathfrak{A}_i \vDash \psi_n(\bar{a}_i; \bar{c}_i^n)$.

We claim that $[\bar{a}]_{\mathfrak{u}}$ realises $\mathfrak{p}$. Consider $\varphi_n \in \mathfrak{p}$. Then

$$[\![\varphi_n(\bar{a}_i)]\!] \supseteq [\![\psi_n(\bar{a}_i)]\!] \supseteq w_n^\circ \in \mathfrak{u} \quad \text{implies} \quad [\![\varphi_n(\bar{a}_i)]\!] \in \mathfrak{u} .$$

By the Theorem of Łoś it follows that $\prod_i \mathfrak{A}_i/\mathfrak{u} \vDash \varphi_n([\bar{a}]_{\mathfrak{u}})$. □

Corollary 2.12. *Let $\mathfrak{u}$ be a regular ultrafilter of an infinite set I and let Σ be a countable signature. For every sequence $(\mathfrak{A}_i)_{i \in I}$ of Σ-structures, the ultraproduct $\prod_{i \in I} \mathfrak{A}_i/\mathfrak{u}$ is $\aleph_1$-saturated.*

Proposition 2.13. *Let $\mathfrak{u}$ be an ultrafilter over a set I of size $\kappa := |I|$. The following statements are equivalent:*

(1) *$\mathfrak{u}$ is regular.*

(2) *For each theory T and every family $(\mathfrak{A}_i)_{i \in I}$ of models of T, the ultraproduct $\prod_i \mathfrak{A}_i/\mathfrak{u}$ realises every partial type $\mathfrak{p}$ over $\varnothing$ with $|\mathfrak{p}| \leq \kappa$.*

(3) *For every structure $\mathfrak{M}$, the ultrapower $\mathfrak{M}^{\mathfrak{u}}$ realises every partial type $\mathfrak{p}$ over M with $|\mathfrak{p}| \leq \kappa$.*

Proof. (1) ⇒ (2) Since $|\mathfrak{p}| \leq |I|$ and $\mathfrak{u}$ is regular we can find sets $(s_\varphi)_{\varphi \in \mathfrak{p}}$ in $\mathfrak{u}$ such that the sets

$$\Phi_i := \{ \varphi \in \mathfrak{p} \mid i \in s_\varphi \}$$

are finite. For every $i \in I$, there exists a tuple $\bar{a}^i \subseteq A_i$ realising the finite type Φ_i. We claim that $\bar{a} := (\bar{a}^i)_i$ realises $\mathfrak{p}$. Let $\varphi \in \mathfrak{p}$. For every $k \in s_\varphi$, we have $k \in [\![\varphi(\bar{a}^i)]\!]_i$. Hence, $s_\varphi \subseteq [\![\varphi(\bar{a}^i)]\!]_i \in \mathfrak{u}$ which implies, by the Theorem of Łoś, that $\prod_i \mathfrak{A}_i/\mathfrak{u} \models \varphi([\bar{a}]_\mathfrak{u})$.

(2) ⇒ (3) follows by setting $\mathfrak{A}_i := \mathfrak{M}_M$, for each $i \in I$.

(3) ⇒ (1) We consider the structure $\mathfrak{M} := \langle M, \subseteq \rangle$ where

$$M := \{ X \subseteq I \mid |X| < \aleph_0 \},$$

and the type

$$\mathfrak{p} := \{ \{k\} \subseteq x \mid k \in I \},$$

which is finitely satisfiable in $\mathfrak{M}$. By (3), there is an element $[a]_\mathfrak{u}$ of $\mathfrak{M}^\mathfrak{u}$ realising $\mathfrak{p}$. For $k \in I$, we set

$$s_k := \{ i \in I \mid \{k\} \subseteq a_i \} = [\![\{k\} \subseteq a_i]\!].$$

Since $\mathfrak{M}^\mathfrak{u} \models \{k\} \subseteq [a]_\mathfrak{u}$ it follows by the Theorem of Łoś that $s_k \in \mathfrak{u}$. Furthermore, each a_i being finite there are only finitely many s_k with $i \in s_k$. Hence, the family $(s_k)_{k \in I}$ witnesses that $\mathfrak{u}$ is regular. □

Proposition 2.14. *Let I be an infinite set, $\mathfrak{u}$ a regular ultrafilter on I, $\kappa := |I|$, and Σ a signature of size $|\Sigma| \leq \kappa$. If $\mathfrak{A}_i$ and $\mathfrak{B}_i$, for $i \in I$, are Σ-structures such that $\mathfrak{A}_i \equiv \mathfrak{B}_i$, for all $i \in I$, then*

$$\prod_{i \in I} \mathfrak{A}_i/\mathfrak{u} \cong^\kappa_{\mathrm{iso}} \prod_{i \in I} \mathfrak{B}_i/\mathfrak{u}.$$

Proof. Below we need our structures to be relational. Therefore, we replace $\mathfrak{A}_i$ and $\mathfrak{B}_i$ by their *relational variants* $\mathfrak{A}_i^*$ and $\mathfrak{B}_i^*$ as follows. Let $\Sigma_{\mathrm{rel}} \subseteq \Sigma$ be the set of relation symbols and $\Sigma_{\mathrm{fun}} \subseteq \Sigma$ the set of function symbols. We replace every function symbol $f \in \Sigma_{\mathrm{fun}}$ of type $\bar{s} \to t$ by a new relation symbol R_f of type $\bar{s}t$. The resulting signature is

$$\Sigma^* := \Sigma_{\mathrm{rel}} \cup \{ R_f \mid f \in \Sigma_{\mathrm{fun}} \}.$$

To every Σ-structure $\mathfrak{M}$, we associate a Σ^*-structure $\mathfrak{M}^*$ by expanding $\mathfrak{M}|_{\Sigma_{\mathrm{rel}}}$ by the graphs

$$R_f^{\mathfrak{M}^*} := \{ \bar{a}b \mid f^{\mathfrak{M}}(\bar{a}) = b \}$$

of the functions $f \in \Sigma_{\mathrm{fun}}$.

Since $\mathfrak{u}$ is regular there exists a sequence $(s_\alpha)_{\alpha<\kappa}$ of sets $s_\alpha \in \mathfrak{u}$ such that, for every $i \in I$, the set $\{ \alpha < \kappa \mid i \in s_\alpha \}$ is finite. Fix an enumeration $\langle \Sigma^*_\alpha, k_\alpha \rangle_{\alpha<\kappa}$ of all pairs $\langle \Sigma^*_\alpha, k_\alpha \rangle$ consisting of finite subsets $\Sigma^*_\alpha \subseteq \Sigma^*$ and $k_\alpha \subseteq \kappa$. For $i \in I$ and $\gamma < \kappa$, set

$$\begin{aligned} \Gamma_i &:= \bigcup \{ \Sigma^*_\alpha \mid i \in s_\alpha \} , \\ K_i &:= \bigcup \{ k_\alpha \mid i \in s_\alpha \} , \\ m_i^\gamma &:= |\{ \alpha \in K_i \mid \alpha \geq \gamma \}| . \end{aligned}$$

We claim that

$$J : \prod_{i \in I} \mathfrak{A}_i/\mathfrak{u} \cong^\kappa_{\mathrm{iso}} \prod_{i \in I} \mathfrak{B}_i/\mathfrak{u} ,$$

where $J \subseteq \mathrm{pIso}_\kappa(\prod_i \mathfrak{A}_i/\mathfrak{u}, \prod_i \mathfrak{B}_i/\mathfrak{u})$ is the following set of partial isomorphisms $\bar{a} \mapsto \bar{b}$. Let $\bar{a} = (a_\nu)_{\nu<\gamma}$ and $\bar{b} = (b_\nu)_{\nu<\gamma}$ where $\gamma < \kappa$ and $a_\nu = [(a^i_\nu)_{i\in I}]_\mathfrak{u}$ and $b_\nu = [(b^i_\nu)_{i\in I}]_\mathfrak{u}$. Then $\bar{a} \mapsto \bar{b} \in J$ if, and only if,

$$\langle \mathfrak{A}^*_i|_{\Gamma_i}, (a^i_\nu)_{\nu\in K_i} \rangle \cong_{m_i^\gamma} \langle \mathfrak{B}^*_i|_{\Gamma_i}, (b^i_\nu)_{\nu\in K_i} \rangle , \quad \text{for all } i \in I .$$

It is straightforward to check that J is κ-complete and κ-bounded. To show that $\langle\rangle \mapsto \langle\rangle \in J$, note that each Γ_i is finite and relational. Hence, we can use Corollary C4.3.6 to show that

$$\mathfrak{A}^*_i|_{\Gamma_i} \equiv \mathfrak{B}^*_i|_{\Gamma_i} \quad \text{implies} \quad \mathfrak{A}^*_i|_{\Gamma_i} \cong_\omega \mathfrak{B}^*_i|_{\Gamma_i} .$$

It remains to prove that J has the back-and-forth property with respect to itself. By symmetry, it is sufficient to prove the forth property. Let $\bar{a} \mapsto \bar{b} \in J$ and $c = [(c^i)_{i\in I}]_\mathfrak{u} \in \prod_i A_i/\mathfrak{u}$. To find a matching element

$d = [(d^i)_{i\in I}]_{\mathfrak{u}} \in \prod_i B_i/\mathfrak{u}$ we consider each component d_i separately. Let $\bar{a} = (a_\nu)_{\nu<\gamma}$ and $\bar{b} = (b_\nu)_{\nu<\gamma}$ as above. By definition, $\bar{a} \mapsto \bar{b} \in J$ implies that

$$\big\langle \mathfrak{A}_i^*|_{\Gamma_i}, (a_\nu^i)_{\nu\in K_i} \big\rangle \cong_{m_i^\gamma} \big\langle \mathfrak{B}_i^*|_{\Gamma_i}, (b_\nu^i)_{\nu\in K_i} \big\rangle .$$

If $\gamma \notin K_i$, we take an arbitrary element $d_i \in B_i$. Otherwise, there exists some $d_i \in B_i$ such that

$$\big\langle \mathfrak{A}_i^*|_{\Gamma_i}, (a_\nu^i)_{\nu\in K_i}, c^i \big\rangle \cong_{m_i^\gamma - 1} \big\langle \mathfrak{B}_i^*|_{\Gamma_i}, (b_\nu^i)_{\nu\in K_i}, d^i \big\rangle .$$

Since $\gamma \in K_i$ implies $m_i^{\gamma+1} = m_i^\gamma - 1$, it follows in both cases that

$$\big\langle \mathfrak{A}_i^*|_{\Gamma_i}, (a_\nu^i)_{\nu\in K_i}, c^i \big\rangle \cong_{m_i^{\gamma+1}} \big\langle \mathfrak{B}_i^*|_{\Gamma_i}, (b_\nu^i)_{\nu\in K_i}, d^i \big\rangle . \qquad \square$$

We have seen that we can find κ-saturated elementary extensions, for all cardinals κ. For saturated elementary extensions the situation is different. The next results give conditions on when such extensions exist.

Proposition 2.15. *Let T be a countable complete first-order theory with infinite models. The following statements are equivalent:*

(1) *T has a countable saturated model.*

(2) *T has a countable $\aleph_1$-universal model.*

(3) $|S^{\bar{s}}(T)| \leq \aleph_0$, *for all finite tuples $\bar{s}$.*

Proof. (1) ⇒ (2) follows from Theorem 2.9.

(2) ⇒ (3) Let $\mathfrak{M}$ be a countable $\aleph_1$-universal model of T. Each type $\mathfrak{p} \in S^{\bar{s}}(T)$ is realised in some countable model. Hence, it is also realised in $\mathfrak{M}$. Since $\mathfrak{M}$ is countable it follows that $|S^{\bar{s}}(T)| \leq \aleph_0$.

(3) ⇒ (1) First, let us show that $|S^{<\omega}(A)| \leq \aleph_0$, for every finite set A. Let $\bar{a}$ be an enumeration of A and $\bar{t}$ the sorts of $\bar{a}$. For every finite tuple of sorts $\bar{s}$ there exists an injective function $f : S^{\bar{s}}(A) \to S^{\bar{s}\bar{t}}(T)$ sending a type $\mathfrak{p} \in S^{\bar{s}}(A)$ to the type

$$f(\mathfrak{p}) := \{\, \varphi(\bar{x}, \bar{y}) \mid \varphi(\bar{x}, \bar{a}) \in \mathfrak{p} \,\} .$$

Consequently, $|S^{\bar{s}}(A)| \leq |S^{\bar{s}\bar{t}}(T)| \leq \aleph_0$. Since T is countable there are only countably many sorts. Therefore it follows that $S^{<\omega}(A)$ is countable as well.

To find the desired saturated model of T we construct an elementary chain $(\mathfrak{M}_n)_{n<\omega}$ of countable models of T such that each $\mathfrak{M}_{n+1}$ realises every type over a finite subset $A \subseteq \mathfrak{M}_n$. Then the union $\mathfrak{M}_\omega := \bigcup_{n<\omega} \mathfrak{M}_n$ will be the desired countable $\aleph_0$-saturated model of T.

We start with an arbitrary countable model $\mathfrak{M}_0$ of T. Given $\mathfrak{M}_n$ we construct $\mathfrak{M}_{n+1}$ as follows. Let F be the class of all finite subsets of M_n and set $P := \bigcup_{A \in F} S^{<\omega}(A)$. By the above remarks it follows that P is countable. Fix an enumeration $(\mathfrak{p}_k)_{k<\omega}$ of P. Using Lemma C3.5.2 we construct an elementary chain $(\mathfrak{A}_n^k)_{k<\omega}$ of countable structures with $\mathfrak{A}_n^0 := \mathfrak{M}_n$ such that $\mathfrak{p}_k$ is realised in $\mathfrak{A}_n^{k+1}$. Their union $\bigcup_k \mathfrak{A}_n^k$ is the desired structure $\mathfrak{M}_{n+1}$. □

For the existence of uncountable saturated structures we can only give a sufficient condition at the moment. A more precise characterisation will be presented in Theorem **??** below.

Theorem 2.16. *Let T be a complete theory with infinite models. If T is κ-stable, for a regular cardinal $\kappa \geq |T|$, then T has a saturated model of size κ.*

Proof. We construct an elementary chain $(\mathfrak{A}_i)_{i\leq\kappa}$ of models $\mathfrak{A}_i \models T$ with $|A_i| = \kappa$. We start with an arbitrary model $\mathfrak{A}_0$ of size κ. For limit ordinals δ, we set $\mathfrak{A}_\delta := \bigcup_{i<\delta} \mathfrak{A}_i$. For the successor step, suppose that we have already defined $\mathfrak{A}_i$. Since T is κ-stable we have $|S^s(A_i)| \leq \kappa$, for all sorts s. Furthermore, there are at most $|T| \leq \kappa$ sorts. Hence, we can use Corollary C3.5.3 to find an elementary extension $\mathfrak{A}_{i+1} \succeq \mathfrak{A}_i$ of size κ that realises every type in $\bigcup_s S^s(A_i)$.

We claim that the limit $\mathfrak{A}_\kappa$ is saturated. It is sufficient to prove that every 1-type over a set $U \subseteq A_\kappa$ of size $|U| < \kappa$ is realised in $\mathfrak{A}_\kappa$. Since κ is regular there exists an index $\alpha < \kappa$ with $U \subseteq A_\alpha$. Consequently, every 1-type over U is realised in $\mathfrak{A}_{\alpha+1} \preceq \mathfrak{A}_\kappa$. □

We conclude this section with a closer look at definable relations in κ-saturated structures. We have already proved in Lemma C5.6.17 that the closure ordinal of a least fixed point on an $\aleph_0$-saturated structure is at most ω.

Lemma 2.17. *Suppose that* $\mathfrak{A}$ *is* κ*-saturated and let* $\varphi(\bar{x})$ *be a first-order formula with* $|\bar{x}| < \omega$. *Either* $|\varphi^{\mathfrak{A}}| < \aleph_0$ *or* $|\varphi^{\mathfrak{A}}| \geq \kappa$.

Proof. Suppose that $\varphi^{\mathfrak{A}}$ is infinite. We construct a sequence $(\bar{a}^i)_{i<\kappa}$ of distinct tuples satisfying φ. Suppose that we have already defined $\bar{a}^i$, for $i < \alpha$. The set

$$\Gamma_\alpha(\bar{x}) := \{\varphi(\bar{x})\} \cup \{\, \bar{x} \neq \bar{a}^i \mid i < \alpha \,\}$$

is a partial type since $\varphi^{\mathfrak{A}}$ is infinite. Since $\mathfrak{A}$ is κ-saturated we can therefore find a tuple $\bar{a}^\alpha$ realising $\Gamma_\alpha(\bar{x})$. □

Proposition 2.18. *A first-order theory* T *admits quantifier elimination if and only if we have*

$$\mathfrak{A} \equiv_0 \mathfrak{B} \quad \textit{implies} \quad \mathfrak{A} \cong_0^{\aleph_0} \mathfrak{B}\,,$$

for all $\aleph_0$*-saturated models* $\mathfrak{A}, \mathfrak{B}$ *of* T.

Proof. ($\Leftarrow$) follows from Corollary D1.2.12. For ($\Rightarrow$), note that, according to Theorem D1.2.6, if $\mathfrak{A}$ and $\mathfrak{B}$ are models of T then we have

$$I_0^{\aleph_0}(\mathfrak{A}, \mathfrak{B}) = I_{\mathrm{FO}}^{\aleph_0}(\mathfrak{A}, \mathfrak{B})\,.$$

Furthermore, if $\mathfrak{A}$ and $\mathfrak{B}$ are $\aleph_0$-saturated then we have

$$I_{\mathrm{FO}}^{\aleph_0}(\mathfrak{A}, \mathfrak{B}) = I_\infty^{\aleph_0}(\mathfrak{A}, \mathfrak{B})\,,$$

by Corollary 2.3. Since $\mathfrak{A} \equiv_0 \mathfrak{B}$ implies $\langle\rangle \mapsto \langle\rangle \in I_\infty^{\aleph_0}(\mathfrak{A}, \mathfrak{B})$, it follows that $\mathfrak{A} \cong_0^{\aleph_0} \mathfrak{B}$. □

Proposition 2.19. *If* $\mathfrak{A}$ *is* κ*-saturated then so is* $\mathcal{I}(\mathfrak{A})$, *for every first-order interpretation* $\mathcal{I}$.

Proof. Recall that interpretations are comorphisms, that is, for every formula $\varphi(\bar{x})$, there is a formula $\varphi^{\mathcal{I}}(\bar{x})$ such that

$$\mathcal{I}(\mathfrak{A}) \vDash \varphi(\mathcal{I}(\bar{a})) \quad \text{iff} \quad \mathfrak{A} \vDash \varphi^{\mathcal{I}}(\bar{a}) .$$

Suppose that $\mathfrak{p} \in S^n(U)$ where $U \subseteq \mathcal{I}[A]$ is of size $|U| < \kappa$. Then there is some set $V \subseteq A$ of size $|V| = |U|$ with $U = \mathcal{I}[V]$. Since $\mathfrak{A}$ is κ-saturated we can find a tuple $\bar{a} \in A^n$ realising the partial type

$$\mathfrak{p}^{\mathcal{I}} := \{ \varphi^{\mathcal{I}}(\bar{x}, \bar{c}) \mid \varphi(\bar{x}, \mathcal{I}(\bar{c})) \in \mathfrak{p}, \ \bar{c} \subseteq V \}$$

over V. It follows that $\mathcal{I}(\bar{a})$ realises $\mathfrak{p}$. □

3. *Projectively saturated structures*

In a saturated structure every type over sets of a certain size is realised. We can extend this requirement by also including types with *second-order* variables. Structures that realise also all types of this form are called *projectively saturated.*

Definition 3.1. Let Σ and Ξ be disjoint signatures and $T \subseteq \mathrm{FO}^{\circ}[\Sigma]$ a first-order theory.

(a) A Ξ*-type* is a subset $\mathfrak{p} \subseteq \mathrm{FO}^{\circ}[\Sigma \cup \Xi]$ such that $T \cup \mathfrak{p}$ is consistent. $\mathfrak{p}$ is *complete* if $\mathfrak{p} = \mathrm{Th}(\mathfrak{A})$ for some $(\Sigma \cup \Xi)$-structure $\mathfrak{A}$ satisfying T. The set of all complete Ξ-types is denoted by $S^{\Xi}(T)$.

(b) A Σ-structure $\mathfrak{A}$ *realises* a Ξ-type $\mathfrak{p}$ if it has a $(\Sigma \cup \Xi)$-expansion $\mathfrak{A}_+$ with $\mathfrak{A}_+ \vDash \mathfrak{p}$.

(c) We call a structure $\mathfrak{A}$ *projectively κ-saturated* if it realises every $\{\xi\}$-type over a set of less than κ parameters, for all relation symbols and function symbols ξ.

Lemma 3.2. *Every projectively κ-saturated structure is κ-saturated and strongly κ-homogeneous.*

Proof. Let $\mathfrak{M}$ be a projectively κ-saturated Σ-structure.

First, we show that $\mathfrak{M}$ is κ-saturated. Let $A \subseteq M$ be a subset of size $|A| < \kappa$ and let $\mathfrak{p} \in S^n(A)$. We have to find some $\bar{c} \in M^n$ with $\operatorname{tp}(\bar{c}/A) = \mathfrak{p}$. Let $\mathfrak{N}$ be some elementary extension of $\mathfrak{M}$ that realises $\mathfrak{p}$ and fix a tuple $\bar{d} \in N^n$ of type $\mathfrak{p}$. Let $R \notin \Sigma$ be a new n-ary relation symbol and set $R^{\mathfrak{N}} = \{\bar{d}\}$. Since $\mathfrak{M}$ is projectively κ-saturated there exists a relation $R^{\mathfrak{M}}$ such that

$$\langle \mathfrak{M}, R^{\mathfrak{M}}, \bar{a} \rangle \equiv \langle \mathfrak{N}, R^{\mathfrak{N}}, \bar{a} \rangle ,$$

where $\bar{a}$ is some enumeration of A. It follows that $R^{\mathfrak{M}}$ contains exactly one tuple $\bar{c}$ and we have $\operatorname{tp}(\bar{c}/A) = \operatorname{tp}(\bar{d}/A) = \mathfrak{p}$.

It remains to show that $\mathfrak{M}$ is strongly κ-homogeneous. Let $\bar{a}, \bar{b} \in M^\alpha$, for $\alpha < \kappa$, be sequences such that $\langle \mathfrak{M}, \bar{a} \rangle \equiv \langle \mathfrak{M}, \bar{b} \rangle$. Set

$$\begin{aligned} \Phi(f) := {} & \operatorname{Th}(\mathfrak{M}, \bar{a}, \bar{b}) \\ & \cup \{ f a_i = b_i \mid i < \alpha \} \\ & \cup \{ \forall x \exists y f y = x \} \\ & \cup \{ \forall \bar{x} (\varphi(\bar{x}) \leftrightarrow \varphi(f\bar{x})) \mid \varphi \in \mathrm{FO} \} , \end{aligned}$$

where $f \notin \Sigma$ is a new unary function symbol. By Lemma 1.11, we know that $\Phi(f)$ is satisfiable. Hence, $\Phi(f)$ is an $\{f\}$-type over $\bar{a}\bar{b}$ and there exist a function $\pi : M \to M$ such that $\langle \mathfrak{M}, \bar{a}\bar{b} \rangle \vDash \Phi(\pi)$. In particular, π is an automorphism of $\mathfrak{M}$ with $\pi(\bar{a}) = \bar{b}$. ☐

Theorem 3.3. *Let $\mathfrak{A}$ be a Σ-structure and $\kappa > |\Sigma| \oplus \aleph_0$ a regular cardinal. There exists a projectively κ-saturated elementary extension $\mathfrak{B} \succeq \mathfrak{A}$ of size $|B| \leq |A|^{<\kappa}$.*

Proof. If $\mathfrak{A}$ is finite then it is already projectively κ-saturated, for all κ. Therefore, we may assume that $\mathfrak{A}$ is infinite. Let us write $\mathfrak{C} \sqsubseteq \mathfrak{D}$ if $\mathfrak{D}$ is an expansion of some elementary extension of $\mathfrak{C}$. If $(\mathfrak{C}_i)_{i<\alpha}$ is a $\sqsubseteq$-chain then we can form its union $\bigcup_{i<\alpha} \mathfrak{C}_i$ and, by the same proof as for elementary chains, it follows that $\mathfrak{C}_k \sqsubseteq \bigcup_{i<\alpha} \mathfrak{C}_i$.

Set $\mu := |\Sigma| \oplus \aleph_0$ and $\lambda := (|A| \oplus \mu^+)^{<\kappa}$. Then $\lambda^{<\kappa} = \lambda \geq \kappa$. We will construct a $\sqsubseteq$-chain $(\mathfrak{C}_\alpha)_{\alpha<\lambda\kappa}$ of length $\lambda\kappa$ where the structure $\mathfrak{C}_\alpha$ is of size $|C_\alpha| = \lambda \otimes (\alpha \oplus 1)$. For simplicity, we assume that C_α is the set of ordinals less than $\lambda(\alpha + 1)$. The Σ-reduct of the union $\bigcup_{\alpha<\lambda\kappa} \mathfrak{C}_\alpha$ will be the desired structure $\mathfrak{B} \succeq \mathfrak{A}$. Note that $B = \lambda\kappa$ has size $|B| = \lambda \otimes \kappa = \lambda$.

For every finite tuple $\bar{s}$ of sorts and each sort t fix a new relation symbol $R_{\bar{s}}$ of type $\bar{s}$ and a new function symbol $f_{\bar{s}t}$ of type $\bar{s} \to t$. Let Ξ be the set of these symbols. For $U \subseteq B$ we can consider $T := \mathrm{Th}(\mathfrak{A})$ as an incomplete theory over the signature Σ_U. Hence, we have the type space $S^\Xi(U) := S(\mathrm{FO}[\Sigma_U \cup \Xi]/T)$. Fix an enumeration $(\mathfrak{p}_i)_{i<\lambda\kappa}$ of all $\{\xi\}$-types $\mathfrak{p}_i \in S^{\{\xi\}}(U_i)$, for all possible $\xi \in \Xi$ and all subsets $U_i \subseteq B$ of size $|U_i| < \kappa$. For every $\nu < \kappa$, there are $|B|^\nu = \lambda^\nu \leq \lambda^{<\kappa} = \lambda$ subsets of size ν and $2^{\mu\oplus\nu} \leq \lambda^{<\kappa} = \lambda$ different $\{\xi\}$-types with ν parameters. Therefore, the above enumeration contains $\lambda \otimes \lambda = \lambda$ different types. Consequently, we can choose the sequence $\langle \mathfrak{p}_i \rangle_{i<\lambda\kappa}$ such that, for every $\alpha < \kappa$, each $\{\xi\}$-type $\mathfrak{p}$ appears at least once with some index $\lambda\alpha \leq i < \lambda(\alpha + 1)$. In particular, we assume that every type appears cofinally often in our enumeration.

We start the construction of $(\mathfrak{C}_i)_i$ with an arbitrary elementary extension $\mathfrak{C}_0 \succeq \mathfrak{A}$ of size $|C_0| = \lambda$. For limit ordinals δ, we set $\mathfrak{C}_\delta := \bigcup_{\alpha<\delta} \mathfrak{C}_\alpha$. For the successor step, suppose that $\mathfrak{C}_\alpha$ has already been defined.

If $U_\alpha \nsubseteq C_\alpha = \lambda(\alpha + 1)$ or if $\mathfrak{p}_\alpha$ is inconsistent with $\mathrm{Th}((\mathfrak{C}_\alpha)_{C_\alpha})$ then we choose an arbitrary elementary extension $\mathfrak{C}_{\alpha+1} \succeq \mathfrak{C}_\alpha$ with universe $\lambda(\alpha+2)$. Otherwise, let $\mathfrak{D}$ be a model of $\mathfrak{p}_\alpha \cup \mathrm{Th}((\mathfrak{C}_\alpha)_{C_\alpha})$. By the Theorem of Löwenheim and Skolem we can choose $\mathfrak{D}$ of size $|D| = \lambda$. Hence, we may assume that $D = \lambda(\alpha + 2)$. By construction, we have $\mathfrak{C}_\alpha \sqsubseteq \mathfrak{D}$ and we can set $\mathfrak{C}_{\alpha+1} := \mathfrak{D}$.

This concludes the construction of $(\mathfrak{C}_\alpha)_\alpha$. Let $\mathfrak{D} := \bigcup_{\alpha<\lambda\kappa} \mathfrak{C}_\alpha$. We claim that $\mathfrak{B} := \mathfrak{D}|_\Sigma$ is a projectively κ-saturated elementary extension of $\mathfrak{A}$. Since $\mathfrak{A} \preceq \mathfrak{C}_0 \sqsubseteq \mathfrak{D}$ we have $\mathfrak{A} \preceq \mathfrak{B}$. Let $V \subseteq B$ be a set of size $|V| < \kappa$ and let $\mathfrak{p}$ be a $\{\xi\}$-type over V. We have to find a relation or function $\xi^{\mathfrak{B}}$ such that $\langle \mathfrak{B}_V, \xi^{\mathfrak{B}} \rangle \vDash \mathfrak{p}$. Since $V \subseteq \lambda\kappa$, $|V| < \kappa$, and κ is regular there is some ordinal α such that $V \subseteq \lambda\alpha$. By construction, there is some index i in the range $\lambda\alpha \leq i < \lambda(\alpha+1)$ such that $\mathfrak{p} = \mathfrak{p}_i$ and $V = U_i$. Consequently, $(\mathfrak{C}_{i+1})_{U_i} \vDash \mathfrak{p}_i$ implies $\langle \mathfrak{B}_V, \xi^{\mathfrak{C}_{i+1}} \rangle \vDash \mathfrak{p}$. □

Corollary 3.4. *Let $\kappa \geq |\Sigma| \oplus \aleph_0$. Every Σ-structure $\mathfrak{A}$ has a projectively κ^+-saturated elementary extension of size at most $|A|^\kappa$.*

In the definition of a projectively saturated structure we only require that every type with one free second-order variable is realised. In fact, we can add several relations at the same time.

Proposition 3.5. *Let $\mathfrak{A}$ be a projectively κ-saturated Σ-structure. Then $\mathfrak{A}$ realises every Ξ-type over less than κ parameters with $|\Xi| < \kappa$.*

Proof. Let $\mathfrak{p}$ be a Ξ-type and $\mathfrak{B} \vDash \mathfrak{p}$ a structure of size κ realising $\mathfrak{p}$. Fix an arbitrary bijection $f : B \times B \to B$ and let $(\xi_i)_{i<\alpha}$ be an enumeration of Ξ. We choose α different elements $c_i \in B$, $i < \alpha$. Using the pairing function f we can replace each relation or function ξ_i by a unary relation P_i. Finally, we define a 4-ary relation R by

$$\begin{aligned} R := {} & \{ \langle a, a, b, f(a,b) \rangle \mid a, b \in B \} \\ & \cup \{ \langle c_i, a, a, b \rangle \mid b \in P_i, a \in B, a \neq c_i \} . \end{aligned}$$

Note that $\mathfrak{B}$ is definable in the structure $\mathfrak{B}' := \langle \mathfrak{B}|_\Sigma, R, (P_i)_i, (c_i)_i \rangle$. Since $\mathfrak{A}$ is projectively κ-saturated it has an expansion $\mathfrak{A}' \equiv \mathfrak{B}'$. We can apply the definition of $\mathfrak{B}$ in $\mathfrak{B}'$ to the structure $\mathfrak{A}'$ to obtain the desired $(\Sigma \cup \Xi)$-expansion $\mathfrak{A}_+$ of $\mathfrak{A}$ with $\mathfrak{A}_+ \equiv \mathfrak{B}$. □

4. *Pseudo-saturated structures*

Depending on the model of set theory there can be first-order theories without saturated models. But if we slightly weaken the definition of saturation then we can prove that such models always exist.

Definition 4.1. A structure $\mathfrak{A}$ is *pseudo-saturated,* or *special,* if there exists an elementary chain $(\mathfrak{A}_\kappa)_{\kappa<|A|}$, indexed by cardinals κ, such that $\mathfrak{A} = \bigcup_\kappa \mathfrak{A}_\kappa$ and every $\mathfrak{A}_\kappa$ is κ^+-saturated.

Lemma 4.2. *Every saturated structure is pseudo-saturated.*

Proof. If $\mathfrak{A}$ is saturated then it is κ^+-saturated, for all $\kappa < |A|$. Therefore, we can obtain the desired chain $(\mathfrak{A}_\kappa)_\kappa$ be setting $\mathfrak{A}_\kappa := \mathfrak{A}$ for all κ. □

By a *strong limit cardinal* we mean a cardinal of the form $\beth_\delta$ where δ is either 0 or a limit ordinal.

Theorem 4.3. *Let $\mathfrak{A}$ be an infinite Σ-structure and κ a strong limit cardinal with $\kappa > |A| \oplus |\Sigma|$. Then $\mathfrak{A}$ has a pseudo-saturated elementary extensions of size κ.*

Proof. Suppose that $\kappa = \beth_\delta$. Fix a strictly increasing sequence $(\lambda_i)_{i<\operatorname{cf}\delta}$ of cardinals $\lambda_i < \beth_\delta$ such that

$$\beth_\delta = \sup\{\, \lambda_i \mid i < \operatorname{cf}\delta \,\} = \sup\{\, 2^{\lambda_i} \mid i < \operatorname{cf}\delta \,\}.$$

By removing some elements of this sequence, we may assume that $\lambda_0 > |A| \oplus |\Sigma|$. We construct an elementary chain $(\mathfrak{B}_i)_{i<\operatorname{cf}\delta}$ such that

- $\mathfrak{B}_0 = \mathfrak{A}$,
- each $\mathfrak{B}_{i+1}$ is a λ_i^+-saturated structure of size $|B_{i+1}| = 2^{\lambda_i}$, and
- $|B_\gamma| \leq 2^{\lambda_\gamma}$, for limit ordinals γ.

The first structure $\mathfrak{B}_0$ is already defined. If $i = j+1$ is a successor then $|B_j| \leq 2^{\lambda_j}$ implies that we can apply Corollary 3.4 to find a λ_i^+-saturated elementary extension $\mathfrak{B}_{j+1} \succeq \mathfrak{B}_j$ of size $|B_i| = |B_j|^{\lambda_i} = 2^{\lambda_i}$. Finally, for limit ordinals γ, we can set $\mathfrak{B}_\gamma := \bigcup_{i<\gamma} \mathfrak{B}_i$ since

$$|B_\gamma| = \sup\{\, 2^{\lambda_i} \mid i < \gamma \,\} \leq 2^{\lambda_\gamma}.$$

The structure $\mathfrak{B} := \bigcup_i \mathfrak{B}_i$ is an elementary extension of $\mathfrak{B}_0 = \mathfrak{A}$ of size $|B| = \sup\{\, 2^{\lambda_i} \mid i < \operatorname{cf}\delta \,\} = \kappa$. We claim that $\mathfrak{B} := \bigcup_i \mathfrak{B}_i$ is pseudo-saturated. Let g be an increasing function from the set of all cardinals less than κ to the ordinal $\operatorname{cf}\delta$ such that $\lambda_{g(\mu)} \geq \mu$, for all $\mu < \kappa$. Then $\mathfrak{B}_{g(\mu)+1}$ is $\lambda^+_{g(\mu)}$-saturated and the chain $(\mathfrak{B}_{g(\mu)+1})_{\mu<\kappa}$ witnesses that $\mathfrak{B}$ is pseudo-saturated. □

Corollary 4.4. *Let $T \subseteq \mathrm{FO}[\Sigma]$ be a consistent first-order theory.*

(a) *T has a pseudo-saturated model.*

(b) *If T has infinite models and $\kappa > |\mathrm{FO}[\Sigma]|$ is a strong limit cardinal then T has a pseudo-saturated model of size κ.*

Proof. (b) By the Theorem of Löwenheim and Skolem T has a model $\mathfrak{A}$ of size $|A| = |\mathrm{FO}[\Sigma]|$. Therefore, we can apply the preceding theorem to obtain a pseudo-saturated elementary extension $\mathfrak{B} \succeq \mathfrak{A}$ of size κ.

(a) If T has infinite models then the claim follows from (b). Otherwise, T has a finite model and every finite structure is saturated. □

Theorem 4.5. *If $\mathfrak{A} \equiv \mathfrak{B}$ are pseudo-saturated structures of the same size $|A| = |B|$ then $\mathfrak{A} \cong \mathfrak{B}$.*

Proof. Suppose that $\mathfrak{A} = \bigcup_\kappa \mathfrak{A}_\kappa$ and $\mathfrak{B} = \bigcup_\kappa \mathfrak{B}_\kappa$. Choose subsets $C_\kappa \subseteq A_\kappa$ and $D_\kappa \subseteq B_\kappa$ of size $|C_\kappa| = |D_\kappa| = \kappa$ such that

$$\bigcup_\kappa C_\kappa = A \quad \text{and} \quad \bigcup_\kappa D_\kappa = B\,.$$

By induction on κ, we construct an increasing chain of partial isomorphisms $(p_\kappa)_\kappa$ with $p_\kappa \in I^\kappa_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B})$ such that

$$C_\kappa \subseteq \operatorname{dom} p_\kappa \subseteq A_\kappa \quad \text{and} \quad D_\kappa \subseteq \operatorname{rng} p_\kappa \subseteq B_\kappa\,.$$

The union $p := \bigcup_\kappa p_\kappa$ is the desired isomorphism.

Let $p_0 := \langle\rangle \mapsto \langle\rangle$. If κ is a limit cardinal then we set $p_\kappa := \bigcup_{\lambda<\kappa} p_\lambda$. Since $I^\kappa_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B})$ is κ-complete, we have $p_\kappa \in I^\kappa_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B})$. Finally, suppose that $\kappa = \lambda^+$ and $p_\lambda = \bar{a} \mapsto \bar{b} \in I^\lambda_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B})$ has already been defined. Let $\bar{c}$ be an enumeration of C_κ and $\bar{d}$ one of D_κ. Since $\mathfrak{A}_\kappa$ and $\mathfrak{B}_\kappa$ are κ^+-saturated, we have

$$\langle \mathfrak{A}, \bar{a} \rangle \cong^{\kappa^+}_{\mathrm{FO}} \langle \mathfrak{B}, \bar{b} \rangle\,.$$

As $|\bar{c}| = |\bar{d}| = \kappa < \kappa^+$ we can apply Lemma C4.4.9 to find sequences $\bar{e} \in (A_\kappa)^\kappa$ and $\bar{f} \in (B_\kappa)^\kappa$ such that

$$\langle \mathfrak{A}, \bar{a}\bar{c}\bar{e} \rangle \cong^{\kappa^+}_{\mathrm{FO}} \langle \mathfrak{B}, \bar{b}\bar{f}\bar{d} \rangle\,.$$

In particular, $p_\kappa := \bar{a}\bar{c}\bar{e} \mapsto \bar{b}\bar{f}\bar{d} \in I^\kappa_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B})$. □

Lemma 4.6. *Let $\mathfrak{A}$ be a pseudo-saturated Σ-structure of size $|A| = \kappa$.*

(a) *The expansion $\langle \mathfrak{A}, \bar{a} \rangle$ is pseudo-saturated, for every sequence $\bar{a} \in A^\alpha$ of length $\alpha < \operatorname{cf} \kappa$.*

(b) *The reduct $\mathfrak{A}|_\Gamma$ is pseudo-saturated, for every $\Gamma \subseteq \Sigma$.*

Proof. (b) follows immediately from the definition.

(a) Let $\mathfrak{A} = \bigcup_{\lambda<\kappa} \mathfrak{A}_\lambda$ where $\mathfrak{A}_\lambda$ is λ^+-saturated. Since $\alpha < \operatorname{cf} \kappa$ there is some index $\mu < \kappa$ with $\bar{a} \subseteq A_\mu$. It follows that $\langle \mathfrak{A}_\lambda, \bar{a} \rangle$ is λ^+-saturated, for every $\lambda \geq \mu$. Consequently, $\langle \mathfrak{A}, \bar{a} \rangle = \bigcup_{\lambda<\kappa} \langle \mathfrak{A}_{\lambda \oplus \mu}, \bar{a} \rangle$ is pseudo-saturated. □

As an easy corollary of Theorem 4.5 we see that every pseudo-saturated structure $\mathfrak{A}$ is $\operatorname{cf}(|A|)$-homogeneous. In fact, we will show below that it is even projectively $\operatorname{cf}(|A|)$-saturated.

Proposition 4.7. *Every pseudo-saturated structure $\mathfrak{A}$ of size $|A| = \kappa$ is strongly $\operatorname{cf}(\kappa)$-homogeneous.*

Proof. Suppose that $\langle \mathfrak{A}, \bar{a} \rangle \equiv \langle \mathfrak{A}, \bar{b} \rangle$, for $\bar{a}, \bar{b} \in A^\alpha$ with $\alpha < \operatorname{cf} \kappa$. The expansions $\langle \mathfrak{A}, \bar{a} \rangle$ and $\langle \mathfrak{A}, \bar{b} \rangle$ are pseudo-saturated, by Lemma 4.6 (a). Consequently, it follows by Theorem 4.5 that they are isomorphic. □

Every pseudo-saturated structure of size κ is projectively $\operatorname{cf}(\kappa)$-saturated and κ^+-universal. To prove this fact we need some technical lemmas.

Lemma 4.8. *Let $\mathfrak{A}$ be a Σ-structure and $\mathfrak{B}$ a Σ_+-structure with $\Sigma \subseteq \Sigma_+$. If $\mathfrak{A}$ and $\mathfrak{B}$ are pseudo-saturated, $\mathfrak{A} \equiv \mathfrak{B}|_\Sigma$, and $|\Sigma_+| \leq |A| \leq |B|$ then there exists an elementary embedding $h : \mathfrak{A} \to \mathfrak{B}|_\Sigma$ such that the set* rng h *induces a substructure of $\mathfrak{B}$.*

Proof. Suppose that $\mathfrak{A} = \bigcup_\lambda \mathfrak{A}_\lambda$ and $\mathfrak{B} = \bigcup_\lambda \mathfrak{B}_\lambda$. Let $(a_\alpha)_{\alpha<\kappa}$ be an enumeration of A such that $a_\alpha \in A_{|\alpha|}$, for all α. We choose a bijection $\tau : \kappa \to T[\Sigma_+, A]$ such that

$$\tau(\alpha) = t(a_{i_0}, \dots, a_{i_{n-1}}) \quad \text{implies} \quad i_0, \dots, i_{n-1} < \alpha\,.$$

To define h we construct an increasing sequence $(p_\alpha)_{\alpha<\kappa}$ of partial elementary maps $p_\alpha \in I_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B})$ such that, for all $\alpha < \kappa$,

- $\operatorname{dom} p_\alpha \subseteq A_{|\alpha|}$ and $\operatorname{rng} p_\alpha \subseteq B_{|\alpha|}$,
- $|p_\alpha| \leq |2\alpha|$,
- $a_\alpha \in \operatorname{dom} p_{\alpha+1}$,
- if $\tau(\alpha) = t(\bar{a})$ then $t^{\mathfrak{B}}[p_\alpha(\bar{a})] \in \operatorname{rng} p_{\alpha+1}$.

The limit $h := \bigcup_\alpha p_\alpha$ will be the desired elementary embedding.

We start the construction with $p_0 := \varnothing$. For limit ordinals δ, we set $p_\delta := \bigcup_{\alpha<\delta} p_\alpha$. For the successor step, suppose that $p_\alpha = \bar{c} \mapsto \bar{d}$ has already been defined. Suppose that $\tau(\alpha) = t(\bar{a})$ and let $y := t^{\mathfrak{B}}[p_\alpha(\bar{a})]$. As $\mathfrak{A}_{|\alpha|}$ is $|\alpha|^+$-saturated there is some element $x \in A_{|\alpha|}$ such that

$$\langle \mathfrak{A}, \bar{c}x \rangle \equiv \langle \mathfrak{B}, \bar{d}y \rangle .$$

Similarly, since $\mathfrak{B}_{|\alpha|}$ is $|\alpha|^+$-saturated we can find an element $z \in B_{|\alpha|}$ with

$$\langle \mathfrak{A}, \bar{c}xa_\alpha \rangle \equiv \langle \mathfrak{B}, \bar{d}yz \rangle .$$

We set $p_{\alpha+1} := \bar{c}xa_\alpha \mapsto \bar{d}yz$. □

Theorem 4.9. *Let $\mathfrak{A}$ be a pseudo-saturated Σ-structure and Ξ a signature disjoint from Σ. If $|A| \geq |\Sigma| \oplus |\Xi|$ then $\mathfrak{A}$ realises every Ξ-type $\mathfrak{p} \in S^{\Xi}(\varnothing)$.*

Proof. Let $\mathfrak{p}^* \subseteq \mathrm{FO}^0[\Gamma]$ be a Skolemisation of $\mathfrak{p}$ and fix a pseudo-saturated model $\mathfrak{B}$ realising $\mathfrak{p}^*$ such that $\mathfrak{B}|_\Sigma \equiv \mathfrak{A}$ and $|B| \geq |A|$. We can use Lemma 4.8 to find exists an elementary embedding $h : \mathfrak{A} \to \mathfrak{B}|_\Sigma$ whose range $B_0 := \operatorname{rng} h$ induces a substructure $\mathfrak{B}_0$ of $\mathfrak{B}$. We define a Γ-expansion $\mathfrak{A}_*$ of $\mathfrak{A}$ by setting

$$\xi^{\mathfrak{A}_*} := h^{-1}[\xi^{\mathfrak{B}_0}] , \quad \text{for } \xi \in \Gamma \setminus \Sigma .$$

It follows that $h : \mathfrak{A}_* \cong \mathfrak{B}_0$. Since $\mathfrak{p}^*$ is a Skolem theory we have $\mathfrak{B}_0 \preceq \mathfrak{B}$. This implies that $\mathfrak{A}_* \cong \mathfrak{B}_0 \vDash \mathfrak{p}^*$. Consequently, $\mathfrak{A}_+ := \mathfrak{A}_*|_{\Sigma\cup\Xi}$ is the desired model of $\mathfrak{p}$. □

Corollary 4.10. *Let $\mathfrak{A}$ be a pseudo-saturated structure of size $|A| = \kappa$ and let Δ be a set of first-order formulae that is closed under conjunctions. If $\mathfrak{B}$ is any structure of size $|B| \leq \kappa$ with $\mathfrak{B} \leq_{\exists\Delta} \mathfrak{A}$ then there exists a Δ-embedding $\mathfrak{B} \to \mathfrak{A}$.*

Proof. Let $\Phi := \mathrm{Th}_\Delta(\mathfrak{B}_B)$. If we can show that $\Phi \cup \mathrm{Th}(\mathfrak{A})$ is consistent then we can use Theorem 4.9 to find an expansion $\mathfrak{A}_C$ of $\mathfrak{A}$ satisfying Φ. Hence, the Diagram Lemma implies that there exists a Δ-embedding $\mathfrak{B} \to \mathfrak{A}$.

It remains to prove that $\Phi \cup \mathrm{Th}(\mathfrak{A})$ is consistent. Suppose otherwise. Then there are finitely many formulae $\varphi_0(\bar{b}_0), \dots, \varphi_{n-1}(\bar{b}_{n-1}) \in \Phi$ with parameters $\bar{b}_i \subseteq B$ such that

$$\mathrm{Th}(\mathfrak{A}) \vDash \neg\varphi_0(\bar{b}_0) \vee \dots \vee \neg\varphi_{n-1}(\bar{b}_{n-1}) .$$

Since Φ is closed under conjunction we may assume w.l.o.g. that $n = 1$. Consequently,

$$\mathfrak{A} \vDash \neg\exists\bar{x}\varphi_0(\bar{x}) .$$

But $\mathfrak{B} \vDash \exists\bar{x}\varphi_0(\bar{x})$ and $\mathfrak{B} \leq_{\exists\Delta} \mathfrak{A}$ implies that $\mathfrak{A} \vDash \exists\bar{x}\varphi_0(\bar{x})$. Contradiction. □

Theorem 4.11. *A pseudo-saturated structure of size κ is κ^+-universal and projectively $\mathrm{cf}(\kappa)$-saturated.*

Proof. Let $\mathfrak{A}$ be pseudo-saturated. If $\mathfrak{B} \equiv \mathfrak{A}$ is a structure of size $|B| \leq \kappa$ then we can use Corollary 4.10 to find an elementary embedding $\mathfrak{B} \to \mathfrak{A}$. Consequently, $\mathfrak{A}$ is κ^+-universal.

For the second claim suppose that $\bar{a} \in A^\alpha$ is a sequence of $\alpha < \mathrm{cf}\,\kappa$ elements. Then $\langle \mathfrak{A}, \bar{a} \rangle$ is pseudo-saturated by Lemma 4.6 (a). It follows by Theorem 4.9 that $\langle \mathfrak{A}, \bar{a} \rangle$ is projectively 1-saturated. Consequently, $\mathfrak{A}$ is projectively $\mathrm{cf}(\kappa)$-saturated. □

Corollary 4.12. *If $\mathfrak{A}$ is pseudo-saturated and $|A|$ is regular then $\mathfrak{A}$ is saturated.*

Corollary 4.13. *Every saturated structure of size κ is projectively κ-saturated.*

Proof. Suppose that $\mathfrak{A}$ is saturated. Then so is $\langle \mathfrak{A}, \bar{a} \rangle$, for every $\bar{a} \in A^{<\kappa}$. Since saturated structures are pseudo-saturated it follows that every expansion $\langle \mathfrak{A}, \bar{a} \rangle$ by less than κ constants is projectively 1-saturated. Consequently, $\mathfrak{A}$ is projectively κ-saturated. □

We conclude this section with a few results about definable relations in pseudo-saturated and projectively saturated structures. We start with an analogue of Lemma 2.17.

Lemma 4.14. *Suppose that $\mathfrak{A}$ is pseudo-saturated and let $\varphi(\bar{x}, \bar{c})$ be a first-order formula with parameters $\bar{c} \subseteq A$ where $|\bar{x}| < \omega$. Then $\varphi(\bar{x}, \bar{c})^{\mathfrak{A}}$ is either finite or $|\varphi(\bar{x}, \bar{c})^{\mathfrak{A}}| = |A|$.*

Proof. Suppose that $\mathfrak{A} = \bigcup_\lambda \mathfrak{A}_\lambda$. If $\varphi^{\mathfrak{A}}$ is infinite then, by Lemma 2.17, we have $|\varphi^{\mathfrak{A}_\lambda}| \geq \lambda^+$. Consequently,

$$|\varphi^{\mathfrak{A}}| \geq |\varphi^{\mathfrak{A}_\lambda}| \geq \lambda^+ , \quad \text{for all } \lambda < |A| ,$$

implies that $|\varphi^{\mathfrak{A}}| = |A|$. □

Lemma 4.15. *If $\mathfrak{A}$ is pseudo-saturated then so is $\mathcal{I}(\mathfrak{A})$, for every first-order interpretation $\mathcal{I}$.*

Proof. Suppose that $\mathfrak{A} = \bigcup_\kappa \mathfrak{A}_\kappa$ where each $\mathfrak{A}_\kappa$ is κ^+-saturated. Note that

$$\mathfrak{A}_\kappa \preceq \mathfrak{A}_\lambda \quad \text{implies} \quad \mathcal{I}(\mathfrak{A}_\kappa) \preceq \mathcal{I}(\mathfrak{A}_\lambda) , \quad \text{for } \kappa \leq \lambda .$$

Hence, the structures $\mathcal{I}(\mathfrak{A}_\kappa)$ form an elementary chain with limit

$$\bigcup_{\kappa < |A|} \mathcal{I}(\mathfrak{A}_\kappa) = \mathcal{I}(\mathfrak{A}) .$$

Furthermore, according to Proposition 2.19, each structure $\mathcal{I}(\mathfrak{A}_\kappa)$ is κ^+-saturated. Hence, $\mathcal{I}(\mathfrak{A})$ is pseudo-saturated. □

Lemma 4.16. *Let $\mathcal{I}$ be a first-order interpretation from Σ to Γ and let $\kappa > |\Sigma| \oplus |\Gamma|$ be a cardinal. If $\mathfrak{A}$ is projectively κ-saturated then so is $\mathcal{I}(\mathfrak{A})$.*

Proof. Let $\bar{a} \subseteq \mathcal{I}(A)$ be a sequence of less than κ-parameters and suppose that $\mathfrak{p}$ is a $\{\xi\}$-type over $\bar{a}$. We can find parameters $\bar{c} \subseteq A$ and an interpretation $\mathcal{J}$ with $\mathcal{J}(\mathfrak{A}, \bar{c}) = \langle \mathcal{I}(\mathfrak{A}), \bar{a} \rangle$. Replacing $\mathfrak{A}$ by $\langle \mathfrak{A}, \bar{c} \rangle$ and $\mathcal{I}$ by $\mathcal{J}$ we can therefore simplify notation by omitting the parameters.

To show that $\mathfrak{p}$ is realised in $\mathcal{I}(\mathfrak{A})$ fix a $(\Gamma \cup \{\xi\})$-structure $\mathfrak{B} \vDash \mathfrak{p}$ realising $\mathfrak{p}$. Let λ be a strong limit cardinal with $\lambda > |\Sigma| \oplus |\Gamma|$ and choose pseudo-saturated structures $\mathfrak{A}_+$ and $\mathfrak{B}_+$ of size λ such that $\mathfrak{A}_+ \equiv \mathfrak{A}$ and $\mathfrak{B}_+ \equiv \mathfrak{B}$. Then $\mathcal{I}(\mathfrak{A}_+) \equiv \mathfrak{B}_+|_\Gamma$ implies, by Theorem 4.5, that $\mathcal{I}(\mathfrak{A}_+) \cong \mathfrak{B}_+|_\Gamma$. Let $\xi^{\mathcal{I}(\mathfrak{A}_+)}$ be the relation on $\mathcal{I}(\mathfrak{A}_+)$ induced by this isomorphism and let $\xi^{\mathfrak{A}_+}$ be its preimage under $\mathcal{I}$. Similarly, for every $\zeta \in \Gamma$, let $\zeta^{\mathfrak{A}_+}$ be the preimage of $\zeta^{\mathcal{I}(\mathfrak{A}_+)}$ under $\mathcal{I}$. W.l.o.g. assume that Σ and Γ are disjoint. Let $\mathfrak{A}_*$ be the $(\Sigma \cup \Gamma \cup \{\xi\})$-expansion of $\langle \mathfrak{A}_+, \xi^{\mathfrak{A}_+} \rangle$ by all these relations and functions $\zeta^{\mathfrak{A}_+}$. We can extend $\mathcal{I}$ to an interpretation $\mathcal{J}$ with

$$\mathcal{J}(\mathfrak{A}_*) = \left\langle \mathcal{I}(\mathfrak{A}_+), \xi^{\mathcal{I}(\mathfrak{A}_+)} \right\rangle .$$

Since $\kappa > |\Sigma| \oplus |\Gamma|$ we can use Proposition 3.5 to find a $(\Sigma \cup \Gamma \cup \{\xi\})$-expansion $\mathfrak{A}'$ of $\mathfrak{A}$ with $\mathfrak{A}' \equiv \mathfrak{A}_*$. It follows that $\mathcal{J}(\mathfrak{A}')$ is an $(\Gamma \cup \{\xi\})$-expansion of $\mathcal{I}(\mathfrak{A})$ with $\mathcal{J}(\mathfrak{A}') \equiv \mathcal{J}(\mathfrak{A}_*) \equiv \mathfrak{B}_+ \equiv \mathfrak{B}$. □

E2. *Definability and automorphisms*

1. *Definability in projectively saturated models*

As an application of the notions introduced in the previous chapter we study the relationship between definable relations and automorphisms.

Definition 1.1. Let L be an algebraic logic, $\mathfrak{M}$ a structure, and $U \subseteq M$ a set of parameters.

(a) A tuple $\bar{a} \subseteq M$ is *L-definable over U* if there is an L-formula $\varphi(\bar{x}; \bar{c})$ with parameters $\bar{c} \subseteq U$ such that $\varphi(\bar{x}; \bar{c})^{\mathfrak{M}} = \{\bar{a}\}$.

(b) The *L-definitional closure* of U is the set

$$\mathrm{dcl}_L(U) := \{ a \in M \mid a \text{ is } L\text{-definable over } U \}.$$

The set U is *L-definitional closed* if it is a fixed point of dcl_L.

(c) We say that an L-formula $\varphi(\bar{x}; \bar{c})$ with parameters $\bar{c} \subseteq M$ is *algebraic* if $\varphi(\bar{x}; \bar{c})^{\mathfrak{M}}$ is finite. An L-type $\mathfrak{p}$ is *algebraic* if it implies an algebraic formula.

We call a tuple $\bar{a} \subseteq M$ *L-algebraic over U* if there is an algebraic L-formula $\varphi(\bar{x}; \bar{c})$ with parameters $\bar{c} \subseteq U$ such that $\mathfrak{M} \vDash \varphi(\bar{a}; \bar{c})$.

(d) The *L-algebraic closure* of U is the set

$$\mathrm{acl}_L(U) := \{ a \in M \mid a \text{ is } L\text{-algebraic over } U \}.$$

The set U is *L-algebraically closed* if it is a fixed point of acl_L.

(e) For $L = \mathrm{FO}$ we simply say that $\bar{a}$ is *definable* or *algebraic* over U and we write $\mathrm{dcl}(U)$ and $\mathrm{acl}(U)$ without the index L.

Lemma 1.2. *Let $\mathfrak{M}$ be a structure. The operators $\mathrm{dcl}_{\mathrm{FO}}$ and $\mathrm{acl}_{\mathrm{FO}}$ are closure operators on M with finite character.*

Proof. Every element $a \in U$ is definable over U by the formula $x = a$. Consequently, $U \subseteq \mathrm{dcl}_{\mathrm{FO}}(U) \subseteq \mathrm{acl}_{\mathrm{FO}}(U)$.

If a is definable or algebraic over U by the formula $\varphi(x;\bar{c})$, the same formula can be used to show that a is definable or algebraic over any set $V \supseteq \bar{c}$. Consequently, $U \subseteq V$ implies $\mathrm{dcl}(U) \subseteq \mathrm{dcl}(V)$ and $\mathrm{acl}(U) \subseteq \mathrm{acl}(V)$. Furthermore, it follows that $a \in \mathrm{dcl}(\bar{c})$ or $a \in \mathrm{acl}(\bar{c})$, respectively. Hence, these operators have finite character.

Finally, suppose that a is definable over $\mathrm{dcl}(U)$. Let $\varphi(x;\bar{c},\bar{d})$ be the corresponding formula where $\bar{d} \subseteq U$ and $\bar{c} \subseteq \mathrm{dcl}(U) \setminus U$. For every element c_i, there is a formula ψ_i over U with $\psi_i^{\mathbb{M}} = \{c_i\}$. We can define a over U by the formula

$$\varphi'(x;\bar{d}) := \exists\bar{y}\Big[\bigwedge_i \psi_i(y_i) \wedge \varphi(x;\bar{y},\bar{d})\Big].$$

The proof for acl is analogous. Suppose that a is algebraic over $\mathrm{acl}(U)$ and let $\varphi(x;\bar{c},\bar{d})$ be the formula witnessing this fact where $\bar{d} \subseteq U$ and $\bar{c} \subseteq \mathrm{acl}(U) \setminus U$. For every element c_i, fix a formula ψ_i over U such that $\psi_i^{\mathfrak{M}}$ is a finite set containing c_i. Set $m := |\varphi(x,\bar{c},\bar{d})^{\mathfrak{M}}|$. The following formula shows that a is algebraic over U.

$$\varphi'(x;\bar{d}) := \exists\bar{y}\Big[\bigwedge_i \psi_i(y_i) \wedge \vartheta(\bar{y}) \wedge \varphi(x;\bar{y},\bar{d})\Big],$$

where

$$\vartheta(\bar{y}) := \forall z_0 \cdots \forall z_m\Big[\bigwedge_i \varphi(z_i;\bar{y},\bar{d}) \rightarrow \bigvee_{i<k} z_i = z_k\Big]$$

states that there are at most m elements z satisfying $\varphi(z;\bar{y},\bar{d})$. □

For strongly κ-homogeneous structures there is a tight relationship between types and automorphisms.

Lemma 1.3. *Let $\mathfrak{M}$ be strongly κ-homogeneous and $U \subseteq M$ a set of size $|U| < \kappa$. For $\bar{a}, \bar{b} \in M^{<\kappa}$, the following statements are equivalent:*

(1) $\operatorname{tp}(\bar{a}/U) = \operatorname{tp}(\bar{b}/U)$

(2) *There is some automorphism* $\pi \in \operatorname{Aut}\mathfrak{M}$ *with*

$$\pi \upharpoonright U = \mathrm{id}_U \quad \textit{and} \quad \pi(\bar{a}) = \bar{b}\,.$$

Proof. (1) ⇒ (2) follows from the definition of a strongly κ-homogeneous structure, while (2) ⇒ (1) follows from the fact that isomorphisms preserve first-order formulae. □

As a consequence we can express the definitional closure and the algebraic closure in terms of automorphisms.

Definition 1.4. Let $\mathfrak{M}$ be a structure and $U \subseteq M$.

(a) Let ξ and ζ be two tuples or two relations in M. We say that ζ is a *conjugate* of ξ *over* U if ξ is mapped to ζ by an automorphism of $\mathfrak{M}$ that fixes U pointwise.

For a sets of formulae Φ and Ψ we similarly say that Ψ is a *conjugate* of Φ over U if there exists an automorphism π fixing U pointwise such that

$$\Psi = \big\{\, \varphi(\bar{x}; \pi(\bar{c})) \;\big|\; \varphi(\bar{x}; \bar{c}) \in \Phi \,\big\}\,.$$

(b) We define the following two closure operators on M:

$$\mathrm{dcl}_{\mathrm{Aut}}(U) := \{\, a \in M \mid a \text{ has exactly one conjugate over } U \,\}\,,$$
$$\mathrm{acl}_{\mathrm{Aut}}(U) := \{\, a \in M \mid a \text{ has only finitely many conjugates over } U \,\}\,.$$

Exercise 1.1. Let $\mathfrak{M}$ be a structure. Prove that $\mathrm{dcl}_{\mathrm{Aut}}$ and $\mathrm{acl}_{\mathrm{Aut}}$ are closure operators on M.

Example. Let $\mathfrak{V}$ be a vector space and let $U \subseteq V$. Then

$$\mathrm{dcl}_{\mathrm{Aut}}(U) = \langle\!\langle U \rangle\!\rangle_{\mathfrak{V}}\,.$$

Remark. Let $\mathfrak{M}$ be a structure and $U \subseteq M$. We can write the pointwise stabiliser of U in $\operatorname{Aut}\mathfrak{M}$ and its setwise stabiliser as

$$(\operatorname{Aut}\mathfrak{M})_{(U)} = \operatorname{Aut}\mathfrak{M}_U \quad \text{and} \quad (\operatorname{Aut}\mathfrak{M})_{\{U\}} = \operatorname{Aut}\langle\mathfrak{M}, U\rangle .$$

In arbitrary structures the relationship between dcl_L and $\mathrm{dcl}_{\mathrm{Aut}}$ and between acl_L and $\mathrm{acl}_{\mathrm{Aut}}$ is as follows.

Lemma 1.5. *Let L be an algebraic logic, $\mathfrak{M}$ a structure, and $U \subseteq M$.*

(a) $\mathrm{dcl}_L(U) \subseteq \mathrm{dcl}_{\mathrm{Aut}}(U)$

(b) $\mathrm{acl}_L(U) \subseteq \mathrm{acl}_{\mathrm{Aut}}(U)$

Proof. (a) If there is an automorphism π with $\pi \restriction U = \mathrm{id}_U$ and $\pi(a) = b$, for $a \neq b$, then

$$\mathfrak{M} \vDash \varphi(a;\bar{c}) \leftrightarrow \varphi(b;\bar{c}),$$

for all L-formulae φ and all parameters $\bar{c} \subseteq U$. Consequently, a is not L-definable over U.

(b) Similarly, if the orbit of a under $\operatorname{Aut}\mathfrak{M}_U$ is infinite then every formula satisfied by a is also satisfied by infinitely many other elements. Hence, a is not L-algebraic over U. □

For sufficiently saturated structures the two closure operators coincide.

Theorem 1.6. *Let $\mathfrak{M}$ be κ-saturated and strongly κ-homogeneous, $a \in M$ an element, and let $U \subseteq M$ be a set of size $|U| < \kappa$.*

(a) *The following statements are equivalent:*

  (1) $a \in \mathrm{dcl}_{\mathrm{FO}}(U)$

  (2) $a \in \mathrm{dcl}_{\mathrm{Aut}}(U)$

  (3) $\mathrm{tp}(a/U)$ *has exactly one realisation in* $\mathfrak{M}$.

(b) *The following statements are equivalent:*

  (1) $a \in \mathrm{acl}_{\mathrm{FO}}(U)$

  (2) $a \in \mathrm{acl}_{\mathrm{Aut}}(U)$

(3) $\operatorname{tp}(a/U)$ *has only finitely many realisations in* $\mathfrak{M}$.

Proof. (a) (2) ⇔ (3) follows by Lemma 1.3.

(1) ⇒ (3) Fix a formula $\varphi(x)$ over U that defines a. Since $\varphi \in \operatorname{tp}(a/U)$, it follows that a is the only realisation of $\operatorname{tp}(a/U)$.

(3) ⇒ (1) Suppose that $a \notin \operatorname{dcl}_{\mathrm{FO}}(U)$. It follows that, for every finite set Φ of first-order formulae over U, there is some element $b \neq a$ such that

$$\mathfrak{M} \vDash \bigwedge \Phi(a) \leftrightarrow \bigwedge \Phi(b) .$$

By the Compactness Theorem and the fact that $\mathfrak{M}$ is κ-saturated, it follows that we can find some element $b \neq a$ with

$$\operatorname{tp}(a/U) = \operatorname{tp}(b/U) .$$

(b) (2) ⇔ (3) follows by Lemma 1.3.

(1) ⇒ (3) Fix a formula $\varphi(x)$ over U such that $\varphi^{\mathfrak{M}}$ is a finite set containing a. Since $\varphi \in \operatorname{tp}(a/U)$ it follows that there are at most $|\varphi^{\mathfrak{M}}|$ realisations of $\operatorname{tp}(a/U)$.

(3) ⇒ (1) We can use an analogous argument as in (a) to show that $a \notin \operatorname{acl}_{\mathrm{FO}}(U)$ implies that there are infinitely many realisations of $\operatorname{tp}(a/U)$. □

Corollary 1.7. *Let* $\mathfrak{M}$ *be a structure and* $U \subseteq M$. *Then*

$$\pi[\operatorname{acl}(U)] = \operatorname{acl}(U) , \quad \textit{for all } \pi \in \operatorname{Aut} \mathfrak{M}_U .$$

Proof. Let $a \in \operatorname{acl}(U)$. To show that $\pi(a) \in \operatorname{acl}(U)$ we consider the set $A \subseteq M$ of all realisations of $\operatorname{tp}(a/U)$. By Theorem 1.6, A is a finite set with $A \subseteq \operatorname{acl}(U)$. Consequently, $\pi(a) \in A \subseteq \operatorname{acl}(U)$. □

Corollary 1.8. *Let* $\mathfrak{M}$ *be* κ*-saturated and strongly* κ*-homogeneous, and let* $A, B \subseteq M$ *be sets of size* $|A|, |B| < \kappa$.

(a) *The following statements are equivalent:*

(1) $A \subseteq \operatorname{dcl}(B)$

(2) $\operatorname{dcl}(A) \subseteq \operatorname{dcl}(B)$
(3) $\operatorname{Aut} \mathfrak{M}_A \supseteq \operatorname{Aut} \mathfrak{M}_B$.

(b) *The following statements are equivalent:*

(1) $A \subseteq \operatorname{dcl}(B)$ *and* $B \subseteq \operatorname{dcl}(A)$
(2) $\operatorname{dcl}(A) = \operatorname{dcl}(B)$
(3) $\operatorname{Aut} \mathfrak{M}_A = \operatorname{Aut} \mathfrak{M}_B$.

Proof. (b) follows from (a).

(a) (1) ⇔ (2) Clearly, $\operatorname{dcl}(A) \subseteq \operatorname{dcl}(B)$ implies $A \subseteq \operatorname{dcl}(A) \subseteq \operatorname{dcl}(B)$. Conversely, $A \subseteq \operatorname{dcl}(B)$ implies $\operatorname{dcl}(A) \subseteq \operatorname{dcl}(\operatorname{dcl}(B)) = \operatorname{dcl}(B)$.

(1) ⇒ (3) Suppose that $A \subseteq \operatorname{dcl}(B)$ and let $\pi \in \operatorname{Aut} \mathfrak{M}_B$. Then it follows by Theorem 1.6 and definition of $\operatorname{dcl}_{\mathrm{Aut}}(B)$ that

$$\pi(a) = a\,, \quad \text{for all } a \in \operatorname{dcl}_{\mathrm{Aut}}(B) = \operatorname{dcl}(B) \supseteq A\,.$$

Hence, $\pi \in \operatorname{Aut} \mathfrak{M}_A$.

(3) ⇒ (1) Suppose that $\operatorname{Aut} \mathfrak{M}_A \supseteq \operatorname{Aut} \mathfrak{M}_B$ and let $a \in A$. Then $a \in \operatorname{dcl}_{\mathrm{Aut}}(A)$ implies that

$$\pi(a) = a\,, \quad \text{for all } \pi \in \operatorname{Aut} \mathfrak{M}_A\,.$$

In particular, we have

$$\pi(a) = a\,, \quad \text{for all } \pi \in \operatorname{Aut} \mathfrak{M}_B\,.$$

By Theorem 1.6 and definition of $\operatorname{dcl}_{\mathrm{Aut}}(B)$, it follows that

$$a \in \operatorname{dcl}_{\mathrm{Aut}}(B) = \operatorname{dcl}(B)\,. \qquad \square$$

As an application of Theorem 1.6, we present the following characterisation of the algebraic closure.

Lemma 1.9. *Let $\mathfrak{M}$ be a Σ-structure that is κ-saturated and strongly κ-homogeneous, for some cardinal $\kappa > |\Sigma|$, and let $U \subseteq M$ be a set of size $|U| < \kappa$. Then*

$$\operatorname{acl}(U) = \bigcap \{\, A \mid \mathfrak{A} \preceq \mathfrak{M} \textit{ with } U \subseteq A \,\}\,.$$

Proof. (⊆) Let $\mathfrak{A} \preceq \mathfrak{M}$ be an elementary substructure containing U. To show that $\operatorname{acl}(U) \subseteq A$, consider an element $a \in \operatorname{acl}(U)$. There exists an algebraic formula $\varphi(x)$ over U with $a \in \varphi^{\mathfrak{M}}$. Let $m := |\varphi^{\mathfrak{M}}|$. Then

$$\mathfrak{M} \vDash \exists^m x \varphi(x) \quad \text{implies} \quad \mathfrak{A} \vDash \exists^m x \varphi(x) .$$

Since $\varphi^{\mathfrak{A}} \subseteq \varphi^{\mathfrak{M}}$ it follows that $\varphi^{\mathfrak{A}} = \varphi^{\mathfrak{M}}$. Hence, $a \in \varphi^{\mathfrak{A}} \subseteq A$.

(⊇) Suppose that $a \notin \operatorname{acl}(U)$. We have to find an elementary substructure $\mathfrak{A} \preceq \mathfrak{M}$ containing U such that $a \notin A$. By Theorem 1.6 and the fact that $\mathfrak{M}$ is κ-saturated, there exists a sequence $(b_\alpha)_{\alpha<\kappa}$ of distinct elements such that

$$\operatorname{tp}(b_\alpha/U) = \operatorname{tp}(a/U) , \quad \text{for all } \alpha < \kappa .$$

Using the Theorem of Löwenheim and Skolem, we can find an elementary substructure $\mathfrak{A}_0 \preceq \mathfrak{M}$ containing U with

$$|A_0| \leq |U| \oplus |\Sigma| < \kappa .$$

There exists an index $\alpha < \kappa$ with $b_\alpha \notin A_0$. Since $\mathfrak{M}$ is strongly κ-homogeneous, we can find an automorphism π with $\pi \upharpoonright U = \mathrm{id}_U$ and $\pi(b_\alpha) = a$. Set $\mathfrak{A} := \pi[\mathfrak{A}_0]$. Then $\mathfrak{A} \preceq \mathfrak{M}$ contains U but not a. □

After considering the definability of single elements we now study the relationship between automorphisms and definable relations. Our first result gives a characterisation of those relations that are definable over a set U of parameters.

Lemma 1.10. *Suppose that $\mathfrak{M}$ is κ-saturated and strongly κ-homogeneous and let $U \subseteq M$ be a set of size $|U| < \kappa$. An M-definable relation $R \subseteq M^n$ is U-definable if, and only if, $\pi[R] = R$, for all $\pi \in \operatorname{Aut} \mathfrak{M}_U$.*

Proof. Clearly, a U-definable relation is invariant under all automorphisms of $\mathfrak{M}$ that fix U pointwise. For the converse, suppose that R is defined by the formula $\varphi(\bar{x}; \bar{c})$ with $\bar{c} \subseteq M$. Consider the set

$$\begin{aligned} \Phi := \{ &\varphi(\bar{x}; \bar{c}) \wedge \neg\varphi(\bar{x}'; \bar{c}) \} \\ \cup \{ &\psi(\bar{x}) \leftrightarrow \psi(\bar{x}') \mid \psi \text{ a formula over } U \} . \end{aligned}$$

If $\Phi(\bar{x}, \bar{x}') \cup \mathrm{Th}(\mathfrak{M}_M)$ is satisfiable then Φ is a partial type and, since $\mathfrak{M}$ is κ-saturated, there are elements $\bar{a}, \bar{b} \in M^n$ satisfying Φ. Let $\pi_0 : U \cup \bar{a} \to U \cup \bar{b}$ be the function with $\pi_0 \upharpoonright U = \mathrm{id}_U$ and $\pi_0(\bar{a}) = \bar{b}$. By choice of $\bar{a}$ and $\bar{b}$ this is an elementary partial function. Since $\mathfrak{M}$ is strongly κ-homogeneous, we can extend it to an automorphism $\pi : M \to M$. But we have $\bar{a} \in \varphi^{\mathfrak{M}} = R$ and $\pi(\bar{a}) = \bar{b} \notin \varphi^{\mathfrak{M}} = R$. Hence, R is not invariant under automorphisms of $\mathrm{Aut}\,\mathfrak{M}_U$. A contradiction.

Consequently, $\Phi \cup \mathrm{Th}(\mathfrak{M}_M)$ is not satisfiable. Hence, there are finitely many formulae $\psi_0, \dots, \psi_{m-1}$ over U such that

$$\mathfrak{M} \vDash \forall \bar{x} \forall \bar{x}' \Big[\bigwedge_i [\psi_i(\bar{x}) \leftrightarrow \psi_i(\bar{x}')] \to [\varphi(\bar{x}; \bar{c}) \leftrightarrow \varphi(\bar{x}'; \bar{c})] \Big] .$$

For $I \subseteq [m]$, define

$$\chi_I(\bar{x}) := \bigwedge_{i \in I} \psi_i(\bar{x}) \wedge \bigwedge_{i \notin I} \neg \psi_i(\bar{x}) ,$$

and let

$$S := \big\{ I \subseteq [m] \bigm| \mathfrak{M} \vDash \chi_I(\bar{a}) \text{ for some } \bar{a} \in R \big\} .$$

It follows that

$$\bar{a} \in R \quad \text{iff} \quad \mathfrak{M} \vDash \bigvee_{I \in S} \chi_I(\bar{a}) .$$

Consequently, the formula $\bigvee_{I \in S} \chi_I(\bar{x})$ defines R over U. □

An analogous result for relations with finitely many conjugates will be given in Lemma 3.11 below.

If the structure $\mathfrak{M}$ is even projectively saturated, we can drop the assumption that the relation R is M-definable. In particular, the following result implies that FO has the Beth property.

Theorem 1.11. *Let Σ, Ξ be disjoint signatures, $\kappa > |\Xi|$, and $T \subseteq \mathrm{FO}^0[\Sigma]$ a first-order theory. For a complete Ξ-type $\mathfrak{p} \in S^{\Xi}(T)$ and a relation symbol $R \in \Xi$, the following statements are equivalent:*

(1) *There is an* $\mathrm{FO}^{<\omega}[\Sigma]$*-formula* $\varphi(\bar{x})$ *such that*

$$\mathfrak{p} \vDash \forall \bar{x}[R\bar{x} \leftrightarrow \varphi(\bar{x})] .$$

(2) *If* $\mathfrak{M}$ *is a model of* T *and* $\mathfrak{N}_0, \mathfrak{N}_1$ *are realisations of* $\mathfrak{p}$ *in* $\mathfrak{M}$ *then* $R^{\mathfrak{N}_0} = R^{\mathfrak{N}_1}$.

(3) *There is a model* $\mathfrak{M}$ *of* T *which is either projectively* κ*-saturated, or saturated and of cardinality at least* $|\Sigma \cup \Xi|$*, such that*

$$R^{\mathfrak{N}_0} = R^{\mathfrak{N}_1} , \quad \text{for every pair } \mathfrak{N}_0, \mathfrak{N}_1 \text{ of realisations of } \mathfrak{p} \text{ in } \mathfrak{M}.$$

(4) *There is a model* $\mathfrak{M}$ *of* T *which is either projectively* κ*-saturated, or saturated and of cardinality at least* $|\Sigma \cup \Xi|$*, such that*

$$\pi[R^{\mathfrak{M}_+}] = R^{\mathfrak{M}_+} , \quad \text{for every realisation } \mathfrak{M}_+ \text{ of } \mathfrak{p} \text{ in } \mathfrak{M} \text{ and each automorphism } \pi \in \operatorname{Aut}\mathfrak{M}.$$

Proof. The implications (1) ⇒ (2) and (3) ⇒ (4) are trivial. (2) ⇒ (3) is also trivial, except for the existence of $\mathfrak{M}$ which follows by Corollary E1.3.4.

(4) ⇒ (1) The proof is similar to that of the preceding lemma. Let $\bar{s}$ be the type of R. We choose new constant symbols $\bar{c}$ and $\bar{d}$ and we set

$$\Phi := \mathfrak{p} \cup \{R\bar{c}, \neg R\bar{d}\} \cup \{ \psi(\bar{c}) \leftrightarrow \psi(\bar{d}) \mid \psi \in \mathrm{FO}^{\bar{s}}[\Sigma] \} .$$

If Φ is inconsistent, there are finitely many formulae $\psi_0, \dots, \psi_{m-1} \in \mathrm{FO}^{\bar{s}}[\Sigma]$ such that

$$\mathfrak{p} \vDash \forall \bar{x}\bar{y}\Big[\bigwedge_{i<m} [\psi_i(\bar{x}) \leftrightarrow \psi_i(\bar{y})] \to (R\bar{x} \leftrightarrow R\bar{y})\Big] .$$

As above we define

$$\chi_I(\bar{x}) := \bigwedge_{i \in I} \psi_i(\bar{x}) \wedge \bigwedge_{i \notin I} \neg\psi_i(\bar{x}) , \quad \text{for } I \subseteq [m] .$$

For every $I \subseteq [m]$, it follows that we either have

$$\mathfrak{p} \vDash \chi_I(\bar{x}) \to R\bar{x} \quad \text{or} \quad \mathfrak{p} \vDash \chi_I(\bar{x}) \to \neg R\bar{x}\,.$$

Consequently, we can define R by the formula

$$\varphi(\bar{x}) := \bigvee_{I \in S} \chi_I(\bar{x}) \quad \text{where} \quad S := \{\, I \subseteq [m] \mid \mathfrak{p} \vDash \chi_I(\bar{x}) \to R\bar{x} \,\}\,.$$

It remains to consider the case where Φ has a model $\mathfrak{A}$. We claim that this is impossible. Since $\mathfrak{p}$ is complete it follows that $\mathfrak{A}|_\Sigma \equiv \mathfrak{M}|_\Sigma$. Consequently, we can use Proposition E1.3.5 to expand $\mathfrak{M}|_\Sigma$ to a model $\mathfrak{M}^+$ of Φ. Let $\bar{a}$ and $\bar{b}$ be the values of the constants $\bar{c}$ and $\bar{d}$ in $\mathfrak{M}^+$, respectively. Then

$$\langle \mathfrak{M}|_\Sigma, \bar{a} \rangle \equiv \langle \mathfrak{M}|_\Sigma, \bar{b} \rangle\,.$$

Since $\mathfrak{M}|_\Sigma$ is strongly $\aleph_0$-homogeneous it follows that there is some automorphism $\pi \in \operatorname{Aut} \mathfrak{M}|_\Sigma$ with $\pi(\bar{a}) = \bar{b}$. But $\bar{a} \in R^{\mathfrak{M}^+}$ and $\pi(\bar{a}) = \bar{b} \notin R^{\mathfrak{M}^+}$ contradicts our choice of $\mathfrak{M}$. □

Corollary 1.12. *Let Σ, Ξ be disjoint signatures, $R \in \Xi$ a relation symbol, and $T \subseteq \mathrm{FO}^0[\Sigma]$ a complete first-order theory. If $\mathfrak{p} \in S^\Xi(T)$ is a complete Ξ-type such that, for every realisation $\mathfrak{M}$ of $\mathfrak{p}$ and all automorphisms $\pi \in \operatorname{Aut} \mathfrak{M}|_\Sigma$, we have*

$$\pi[R^{\mathfrak{M}}] = R^{\mathfrak{M}},$$

then there is an $\mathrm{FO}^{<\omega}[\Sigma]$-formula $\varphi(\bar{x})$ such that

$$\mathfrak{p} \vDash \forall \bar{x}[R\bar{x} \leftrightarrow \varphi(\bar{x})]\,.$$

Proof. Since T has a projectively $|\Xi|^+$-saturated model, the claim follows from Theorem 1.11. □

Corollary 1.13. *Let Σ, Ξ be disjoint signatures, $R \in \Xi$ a relation symbol, and $T \subseteq \mathrm{FO}^0[\Sigma]$ a first-order theory. If $\mathfrak{p}$ is a Ξ-type such that, for every realisation $\mathfrak{M}$ of $\mathfrak{p}$ and all automorphisms $\pi \in \mathrm{Aut}\,\mathfrak{M}|_\Sigma$, we have*

$$\pi[R^{\mathfrak{M}}] = R^{\mathfrak{M}},$$

then there are finitely many formulae $\varphi_0(\bar{x}), \dots, \varphi_{n-1}(\bar{x}) \in \mathrm{FO}^{<\omega}[\Sigma]$ such that

$$\mathfrak{p} \vDash \bigvee_{i<n} \forall \bar{x}[R\bar{x} \leftrightarrow \varphi_i(\bar{x})] .$$

Proof. If $\mathfrak{q} \supseteq \mathfrak{p}$ is a complete Ξ-type, we can use the preceding corollary to find a formula $\varphi_{\mathfrak{q}}(\bar{x})$ defining R modulo $\mathfrak{q}$. Consequently,

$$\mathfrak{p} \vDash \bigvee \{ R\bar{x} \leftrightarrow \varphi_{\mathfrak{q}}(\bar{x}) \mid \mathfrak{q} \supseteq \mathfrak{p} \text{ complete} \} .$$

By compactness, it follows that there are finitely many complete types $\mathfrak{q}_0, \dots, \mathfrak{q}_{n-1} \supseteq \mathfrak{p}$ with

$$\mathfrak{p} \vDash \bigvee_{i<n} [R\bar{x} \leftrightarrow \varphi_{\mathfrak{q}_i}(\bar{x})] .$$

□

Below we will frequently work in projectively saturated elementary extensions of a given model. In order to simplify the presentation and to avoid having to include phrases like 'there exists an elementary extension such that', it turned out to be a good idea to fix such an extension once and for all. If this structure is sufficiently saturated, we can use the Amalgamation Theorem and Theorem E1.2.9 to embed all other models we consider into it.

Thus, let us fix a projectively κ-saturated model $\mathbb{M}$ of T where κ is some very large cardinal. We call $\mathbb{M}$ the *monster model* of T. All models $\mathfrak{M}$ of T we will consider are tacitly assumed to be elementary substructures of $\mathbb{M}$ of size $|M| < \kappa$.

We call a relation $R \subseteq \mathbb{M}^n$ *small* if $|R| < \kappa$. Otherwise, it is *large*. To distinguish small and large relations we denote the latter by blackboard bold symbols $\mathbb{A}, \mathbb{B}, \mathbb{C}, \dots$. Note that, by Lemma E1.2.17, definable relations are either finite or large. Mostly, we will only consider types $\mathfrak{p} \in S^{\bar{s}}(U)$ over small sets U of parameters. Note that every such type is realised in $\mathbb{M}$. Similarly, we will tacitly assume that all parameter-definable relations are defined over a small set of parameters.

To simplify notation, we will drop the model $\mathbb{M}$ and write just $\bar{a} \equiv_U \bar{b}$ instead of $\langle \mathbb{M}_U, \bar{a} \rangle \equiv \langle \mathbb{M}_U, \bar{b} \rangle$. By Lemma 1.3, it follows that $\bar{a} \equiv_U \bar{b}$ if, and only if, there exists a U-automorphism π of $\mathbb{M}$ mapping $\bar{a}$ to $\bar{b}$. We extend this notation to sequences of sets $A_0, \dots, A_n, B_0, \dots, B_n \subseteq \mathbb{M}$ by defining

$$A_0 \dots A_n \equiv_U B_0 \dots B_n$$

if there are enumerations $\bar{a}_i$ of A_i and $\bar{b}_i$ of B_i such that

$$\mathrm{tp}(\bar{a}_0 \dots \bar{a}_n / U) = \mathrm{tp}(\bar{b}_0 \dots \bar{b}_n / U)\,.$$

2. *Imaginary elements and canonical parameters*

In this section we present a construction adding to a given structure new elements representing all definable relations. More generally, we add elements for every class of a definable equivalence relation.

Definition 2.1. Let $\mathfrak{M}$ be an S-sorted structure. An *equivalence formula* is a formula $\chi(\bar{x}, \bar{y})$ without parameters defining an equivalence relation on $M^{\bar{s}}$, for some $\bar{s} \in S^{<\omega}$. The tuple $\bar{s}$ is called the *type* of χ. We denote the equivalence class of a tuple $\bar{a} \in M^{\bar{s}}$ by $[\bar{a}]_\chi$. The elements of the quotient $M^{\bar{s}}/\chi^{\mathfrak{M}}$ are called *imaginary elements.*

Given $\mathfrak{M}$ we construct a new structure $\mathfrak{M}^{\mathrm{eq}}$ by adding all imaginary elements.

Definition 2.2. Let $\mathfrak{M}$ be an S-sorted Σ-structure.

(a) Set

$$S^{\mathrm{eq}} := \{\, \chi \mid \chi \text{ an equivalence formula} \,\}\,,$$
$$\Sigma^{\mathrm{eq}} := \Sigma \cup \{\, p_\chi \mid \chi \in S^{\mathrm{eq}} \,\}\,.$$

We regard S as a subset of S^{eq} via the identification of $s \in S$ with the formula $(x = y) \in S^{\mathrm{eq}}$, where x and y are variables of sort s.

We construct an S^{eq}-sorted Σ^{eq}-structure $\mathfrak{M}^{\mathrm{eq}}$ as follows. For every equivalence formula χ of type $\bar{s}$, the domain of sort χ is

$$M^{\mathrm{eq}}_{\chi} := M^{\bar{s}} / \chi^{\mathfrak{M}}.$$

By the identification of $s \in S$ with $(x = y) \in S^{\mathrm{eq}}$, we obtain an embedding of M into M^{eq}. We interpret the symbols of $\Sigma \subseteq \Sigma^{\mathrm{eq}}$ in $\mathfrak{M}^{\mathrm{eq}}$ according to this embedding. The new function symbols p_{χ} are interpreted as the canonical projections $M^{\bar{s}} \to M^{\bar{s}} / \chi^{\mathfrak{M}}$.

(b) To avoid ambiguities we denote the definable closure and the algebraic closure of a subset $U \subseteq M^{\mathrm{eq}}$ by $\mathrm{dcl}^{\mathrm{eq}}(U)$ and $\mathrm{acl}^{\mathrm{eq}}(U)$, respectively, while $\mathrm{dcl}(U)$ and $\mathrm{acl}(U)$ are the closures of U in the original structure $\mathfrak{M}$.

Remark. (a) Every finite tuple $\bar{a} \in M^{\bar{s}}$ is encoded in $\mathfrak{M}^{\mathrm{eq}}$ as a single element $[\bar{a}]_{\chi} \in M^{\mathrm{eq}}$ of sort

$$\chi(\bar{x}, \bar{y}) := x_0 = y_0 \wedge \cdots \wedge x_{n-1} = y_{n-1}\,,$$

where the variables x_i and y_i have sort s_i.

(b) For each formula $\varphi(\bar{x})$, we can define the equivalence formula

$$\chi(\bar{x}, \bar{y}) := \varphi(\bar{x}) \leftrightarrow \varphi(\bar{y})\,.$$

There are two imaginary elements of sort χ: one representing $\varphi^{\mathfrak{M}}$, the other one representing $\neg\varphi^{\mathfrak{M}}$. Consequently, $\mathfrak{M}^{\mathrm{eq}}$ contains imaginary elements for all relations definable without parameters.

The next proposition shows that, when considering the logical properties of a structure, the transition from $\mathfrak{M}$ to $\mathfrak{M}^{\mathrm{eq}}$ does not change much. But we will see below that, when studying automorphisms, this construction allows us in certain cases to replace setwise stabilisers by pointwise ones.

Proposition 2.3. *Let $\mathfrak{M}$ be a structure.*

(a) *$\mathfrak{M}$ is a relativised reduct of $\mathfrak{M}^{\mathrm{eq}}$.*

(b) *There exists a first-order interpretation mapping $\mathfrak{M}$ to $\mathfrak{M}^{\mathrm{eq}}$.*

(c) *For every formula $\varphi(\bar{x}) \in \mathrm{FO}^{\bar{s}}[\Sigma^{\mathrm{eq}}]$, we can construct a formula $\varphi'(\bar{x}) \in \mathrm{FO}^{\bar{s}}[\Sigma]$ such that*

$$\mathfrak{M}^{\mathrm{eq}} \vDash \varphi(\bar{a}) \quad \text{iff} \quad \mathfrak{M} \vDash \varphi'(\bar{a}), \quad \textit{for all } \bar{a} \in M^{\bar{s}}.$$

(d) *$\mathfrak{A} \equiv \mathfrak{B}$ implies $\mathfrak{A}^{\mathrm{eq}} \equiv \mathfrak{B}^{\mathrm{eq}}$.*

(e) *$M^{\mathrm{eq}} = \langle\!\langle M \rangle\!\rangle_{\mathfrak{M}^{\mathrm{eq}}}$.*

(f) *Every element of M^{eq} is definable over M.*

(g) *Every elementary embedding $g : \mathfrak{A} \to \mathfrak{B}$ can be extended to an elementary embedding $\mathfrak{A}^{\mathrm{eq}} \to \mathfrak{B}^{\mathrm{eq}}$ in a unique way.*

(h) *The restriction map*

$$\rho : \mathfrak{Aut}\,\mathfrak{M}^{\mathrm{eq}} \to \mathfrak{Aut}\,\mathfrak{M} : \pi \mapsto \pi \upharpoonright M$$

is a group isomorphism.

(i) *For every $U \subseteq M$, we have*

$$\mathrm{dcl}(U) = \mathrm{dcl}^{\mathrm{eq}}(U) \cap M \quad \textit{and} \quad \mathrm{acl}(U) = \mathrm{acl}^{\mathrm{eq}}(U) \cap M.$$

Proof. (a) and (b) follow immediately from the definition of $\mathfrak{M}^{\mathrm{eq}}$.

(c) and (d) follow from (b) via Lemma C1.5.9 and Corollary C1.5.13, respectively.

(e) Every imaginary element $[\bar{a}]_\chi \in M^{\mathrm{eq}}$ is denoted by a term $p_\chi \bar{a}$ with parameters $\bar{a} \subseteq M$.

(f) follows immediately from (e).

(g) Let $g : \mathfrak{A} \to \mathfrak{B}$ be an elementary embedding. It follows by (b) and Lemma C2.2.10 that the map $[\bar{a}]_\chi \mapsto [g(\bar{a})]_\chi$ is an elementary embedding $\mathfrak{A}^{\mathrm{eq}} \to \mathfrak{B}^{\mathrm{eq}}$ extending g. For uniqueness, suppose that there are elementary embeddings $h_0, h_1 : \mathfrak{A}^{\mathrm{eq}} \to \mathfrak{B}^{\mathrm{eq}}$ with $h_0 \upharpoonright A = h_1 \upharpoonright A$. By Theorem B3.1.9, it follows that $h_0 \upharpoonright \langle\!\langle A \rangle\!\rangle_{\mathfrak{A}^{\mathrm{eq}}} = h_1 \upharpoonright \langle\!\langle A \rangle\!\rangle_{\mathfrak{A}^{\mathrm{eq}}}$. Hence, (e) implies that $h_0 = h_1$.

(h) First, note that ρ is well-defined since it follows by Lemma C2.2.10 and (a) that, for all $\pi \in \operatorname{Aut} \mathfrak{M}^{\mathrm{eq}}$, the restriction $\pi \restriction M$ is indeed an automorphism of $\mathfrak{M}$. Furthermore, ρ is obviously a group homomorphism. Hence, it remains to show that it is bijective. For surjectivity, note that, by (b), every automorphism of $\mathfrak{M}$ can be extended to one of $\mathfrak{M}^{\mathrm{eq}}$. For injectivity, note that, by (g), every automorphism of $\mathfrak{M}$ can be extended to at most one of $\mathfrak{M}^{\mathrm{eq}}$.

(i) To see that $\operatorname{acl}(U) \subseteq \operatorname{acl}^{\mathrm{eq}}(U)$ note that, if there is a formula φ over U defining a finite set X in $\mathfrak{M}$ then the same formula can be used to define X in $\mathfrak{M}^{\mathrm{eq}}$. For the converse, suppose that φ is a formula over U defining a finite set $X \subseteq M$ in $\mathfrak{M}^{\mathrm{eq}}$. By (c), we can find a formula φ' over U defining the same set in $\mathfrak{M}$. The claim for the definable closure is proved analogously. □

According to the preceding proposition, the first-order theory of $\mathfrak{M}^{\mathrm{eq}}$ only depends on the theory of $\mathfrak{M}$. Consequently, we can extend the operation ${}^{\mathrm{eq}}$ to theories.

Definition 2.4. For a complete first-order theory T, we denote the theory $\operatorname{Th}(\mathbb{M}^{\mathrm{eq}})$ by T^{eq}.

It also follows that adding imaginary elements does not change the structure of the type spaces.

Corollary 2.5. *Let $U \subseteq \mathbb{M}^{\mathrm{eq}}$ and $U_0 \subseteq \mathbb{M}$ be sets.*

$$\operatorname{dcl}^{\mathrm{eq}}(U) = \operatorname{dcl}^{\mathrm{eq}}(U_0) \quad \textit{implies} \quad \mathfrak{S}^{\bar{s}}(T^{\mathrm{eq}}(U)) \cong \mathfrak{S}^{\bar{s}}(T(U_0)) .$$

Proof. Since $\operatorname{dcl}^{\mathrm{eq}}(U) = \operatorname{dcl}^{\mathrm{eq}}(U_0)$, it follows by Proposition 2.3 and Lemma C3.3.4 that $\mathrm{FO}^{\bar{s}}[\Sigma_{U_0}]/T(U_0)$ is a retract of $\mathrm{FO}^{\bar{s}}[\Sigma^{\mathrm{eq}}_U]/T^{\mathrm{eq}}(U)$. Consequently, the claim follows by Corollary C3.3.3. □

As a consequence, many logical properties of $\mathfrak{M}$ and T transfer to $\mathfrak{M}^{\mathrm{eq}}$ and T^{eq}. We give two examples.

Lemma 2.6. *Let T be a complete first-order theory, $\mathfrak{M}$ a structure, and κ an infinite cardinal.*

(a) *$\mathfrak{M}$ is κ-saturated if, and only if, $\mathfrak{M}^{\mathrm{eq}}$ is κ-saturated.*

(b) *T is κ-stable if, and only if, T^{eq} is κ-stable.*

Proof. (a) We have seen in Proposition E1.2.19 that κ-saturation is preserved under interpretations.

(b) ($\Leftarrow$) Suppose that T^{eq} is κ-stable. To show that T is κ-stable, consider a set $U \subseteq \mathbb{M}$ of size $|U| \leq \kappa$. By Corollary 2.5, we have

$$\mathfrak{S}^{\bar{s}}(T(U)) \cong \mathfrak{S}^{\bar{s}}(T^{\mathrm{eq}}(U))\,.$$

Consequently, $|S^{\bar{s}}(T(U))| = |S^{\bar{s}}(T^{\mathrm{eq}}(U))| \leq \kappa$.

($\Rightarrow$) Suppose that T is κ-stable and let $U \subseteq \mathbb{M}^{\mathrm{eq}}$ be a set of size $|U| \leq \kappa$. There exists a set $C \subseteq \mathbb{M}$ of size $|C| \leq |U| \oplus \aleph_0 \leq \kappa$ with $U \subseteq \mathrm{dcl}^{\mathrm{eq}}(C)$. By Corollary 2.5, we have

$$\mathfrak{S}^{\bar{s}}(T(C)) \cong \mathfrak{S}^{\bar{s}}(T^{\mathrm{eq}}(U \cup C))\,.$$

Consequently, $|S^{\bar{s}}(T^{\mathrm{eq}}(U))| \leq |S^{\bar{s}}(T^{\mathrm{eq}}(U \cup C))| = |S^{\bar{s}}(T(C))| \leq \kappa$. □

We have seen that the operation of adding imaginary elements is well-behaved. But what do we gain by it? As an example, consider the following problem. Suppose that a relation $\mathbb{R}$ is defined by a formula $\varphi(\bar{x};\bar{c})$ with parameters $\bar{c}$. There might be many other parameters $\bar{d}$ such that $\varphi(\bar{x};\bar{d})$ defines the same relation $\mathbb{R}$. Sometimes, we would like the parameter $\bar{c}$ to be unique. Using imaginary elements, this can be done. We start by defining the equivalence formula

$$\chi(\bar{y},\bar{y}') := \forall \bar{x}\big[\varphi(\bar{x};\bar{y}) \leftrightarrow \varphi(\bar{x};\bar{y}')\big]\,.$$

Then two tuples $\bar{a}$ and $\bar{b}$ are equivalent if $\varphi(\bar{x};\bar{a})$ and $\varphi(\bar{x};\bar{b})$ define the same relation. Consequently, the tuples in $[\bar{c}]_\chi$ are precisely those defining $\mathbb{R}$. The imaginary element $e := [\bar{c}]_\chi$ is a unique representative of this set. We obtain a formula

$$\psi(\bar{x};z) := \exists y\big[\varphi(\bar{x};\bar{y}) \wedge p_\chi \bar{y} = z\big]$$

such that e is the unique element such that $\psi(\bar{x};e)$ defines $\mathbb{R}$. Let us formalise this construction.

Definition 2.7. Let $\varphi(\bar{x};\bar{y})$ be a formula.

(a) The *parameter equivalence* for φ is the formula

$$\chi(\bar{y},\bar{y}') := \forall\bar{x}\big[\varphi(\bar{x};\bar{y}) \leftrightarrow \varphi(\bar{x};\bar{y}')\big].$$

(b) A tuple $\bar{c}$ is a *canonical parameter* of a relation $\mathbb{R}$ if there exists a formula $\psi(\bar{x};\bar{y})$ such that $\bar{c}$ is the unique tuple satisfying

$$\psi(\bar{x};\bar{c})^{\mathbb{M}} = \mathbb{R}.$$

In this case, we call the formula $\psi(\bar{x};\bar{c})$ a *canonical definition* of $\mathbb{R}$.

In this terminology we can state the above remark as follows.

Lemma 2.8. *Let χ be the parameter equivalence of a formula $\varphi(\bar{x};\bar{y})$. For every tuple $\bar{c}$, the imaginary element $[\bar{c}]_\chi \in \mathbb{M}^{\mathrm{eq}}_\chi$ is a canonical parameter of $\varphi(\bar{x};\bar{c})^{\mathbb{M}}$.*

Proof. The formula

$$\psi(\bar{x};[\bar{c}]_\chi) := \exists\bar{y}\big[\varphi(\bar{x};\bar{y}) \wedge p_\chi\bar{y} = [\bar{c}]_\chi\big]$$

is a canonical definition of $\varphi(\bar{x};\bar{c})^{\mathbb{M}}$. □

Corollary 2.9. *Every relation $\mathbb{R} \subseteq \mathbb{M}^{\bar{s}}$ that is definable over a set $U \subseteq \mathbb{M}$ has a canonical parameter $e \in \mathrm{dcl}^{\mathrm{eq}}(U)$.*

Thus, all parameter-definable relations $\mathbb{R} \subseteq \mathbb{M}^{\bar{s}}$ have canonical parameters in $\mathbb{M}^{\mathrm{eq}}$. We will see in Corollary 2.12 below that the same is true for parameter-definable relations in $\mathbb{M}^{\mathrm{eq}}$. The reason for this is that performing the operation $^{\mathrm{eq}}$ twice does not offer any additional benefit: according to the following proposition there exist, for every sort $\chi \in (S^{\mathrm{eq}})^{\mathrm{eq}}$, a sort $\eta \in S^{\mathrm{eq}}$ and a definable bijection $(M^{\mathrm{eq}})^{\mathrm{eq}}_\chi \to M^{\mathrm{eq}}_\eta$. Hence, every doubly imaginary element is already present as a singly imaginary one.

Proposition 2.10. *For every equivalence formula $\chi(\bar{x}, \bar{y})$ with type $\bar{\zeta} \in (S^{\mathrm{eq}})^n$, there exist a sort $\eta \in S^{\mathrm{eq}}$ and a definable, surjective function*

$$f : (M^{\mathrm{eq}})^{\bar{\zeta}} \to (M^{\mathrm{eq}})^{\eta}$$

such that $\ker f = \chi^{\mathfrak{M}^{\mathrm{eq}}}$.

Proof. Each sort $\zeta_i \in S^{\mathrm{eq}}$ is itself an equivalence formula of some type $\bar{s}_i \in S^{<\omega}$. We set

$$\eta(\bar{x}_0 \dots \bar{x}_{n-1}, \bar{y}_0 \dots \bar{y}_{n-1}) := \chi\big(p_{\zeta_0}\bar{x}_0, \dots, p_{\zeta_{n-1}}\bar{x}_{n-1},\ p_{\zeta_0}\bar{y}_0, \dots, p_{\zeta_{n-1}}\bar{y}_{n-1}\big)\,.$$

Then $\eta \in S^{\mathrm{eq}}$ is an equivalence formula of type $\bar{s}_0 \dots \bar{s}_{n-1}$. We claim that the desired function $f : (M^{\mathrm{eq}})^{\bar{\zeta}} \to (M^{\mathrm{eq}})^{\eta}$ is defined by the formula

$$\varphi(\bar{x}, y) := \exists \bar{z}_0 \cdots \exists \bar{z}_{n-1}\Big[\bigwedge_{i<n} x_i = p_{\zeta_i}\bar{z}_i \wedge p_\eta \bar{z}_0 \dots \bar{z}_{n-1} = y\Big]\,.$$

Note that

$$\mathfrak{M}^{\mathrm{eq}} \vDash \varphi(\bar{\alpha}, b)$$

if, and only if, there are tuples $\bar{a}_0, \dots, \bar{a}_{n-1}$ such that

$$\bar{\alpha} = \langle [\bar{a}_0]_{\zeta_0}, \dots, [\bar{a}_{n-1}]_{\zeta_{n-1}} \rangle \quad \text{and} \quad b = [\bar{a}_0 \dots \bar{a}_{n-1}]_\eta\,.$$

Since the equivalence class $[\bar{a}_0 \dots \bar{a}_{n-1}]_\eta$ does not depend on the particular choice of representatives $\bar{a}_i \in [\bar{a}_i]_{\zeta_i}$, the element b is uniquely determined by $\bar{\alpha}$. Thus, φ defines a function $f : (M^{\mathrm{eq}})^{\bar{\zeta}} \to (M^{\mathrm{eq}})^{\eta}$.

To see that f is surjective, note that, for every element $[\bar{a}_0 \dots \bar{a}_{n-1}]_\eta \in (M^{\mathrm{eq}})^{\eta}$, we have

$$[\bar{a}_0 \dots \bar{a}_{n-1}]_\eta = f([\bar{a}_0]_{\zeta_0}, \dots, [\bar{a}_{n-1}]_{\zeta_{n-1}}) \in \operatorname{rng} f\,.$$

Hence, it remains to compute the kernel. Let $\bar{\alpha}, \bar{\alpha}' \in (M^{\mathrm{eq}})^{\bar{\zeta}}$ and suppose that $\bar{\alpha} = \langle [\bar{a}_0]_{\zeta_0}, \dots, [\bar{a}_{n-1}]_{\zeta_{n-1}} \rangle$ and $\bar{\alpha}' = \langle [\bar{a}'_0]_{\zeta_0}, \dots, [\bar{a}'_{n-1}]_{\zeta_{n-1}} \rangle$. Then

$$\begin{aligned} f(\bar{\alpha}) = f(\bar{\alpha}') \quad &\text{iff} \quad \mathfrak{M}^{\mathrm{eq}} \vDash \exists y[\varphi(\bar{\alpha}, y) \wedge \varphi(\bar{\alpha}', y)] \\ &\text{iff} \quad [\bar{a}_0 \dots \bar{a}_{n-1}]_\eta = [\bar{a}'_0 \dots \bar{a}'_{n-1}]_\eta \\ &\text{iff} \quad \mathfrak{M}^{\mathrm{eq}} \vDash \eta(\bar{a}_0 \dots \bar{a}_{n-1}, \bar{a}'_0 \dots \bar{a}'_{n-1}) \\ &\text{iff} \quad \mathfrak{M}^{\mathrm{eq}} \vDash \chi(\bar{\alpha}, \bar{\alpha}') . \end{aligned}$$ □

We obtain the following generalisation of Lemma 2.8.

Corollary 2.11. *Let $\mathfrak{M}$ be a structure. For every formula $\varphi(\bar{x}; \bar{y})$, there exists a formula $\psi(\bar{x}; \bar{z})$ such that, for every tuple $\bar{b} \subseteq M^{\mathrm{eq}}$, there is a unique tuple $\bar{c} \subseteq M^{\mathrm{eq}}$ with*

$$\varphi(\bar{x}; \bar{b})^{\mathfrak{M}^{\mathrm{eq}}} = \psi(\bar{x}; \bar{c})^{\mathfrak{M}^{\mathrm{eq}}} .$$

Proof. Let $\varphi(\bar{x}; \bar{y})$ be a formula with parameter equivalence $\chi(\bar{y}, \bar{y}')$. According to Proposition 2.10 there exists a definable and surjective function $f : (M^{\mathrm{eq}})^{\bar{\zeta}} \to (M^{\mathrm{eq}})^\eta$ such that $\ker f = \chi^{\mathfrak{M}}$. We claim that the formula

$$\psi(\bar{x}; \bar{z}) := \exists \bar{y}[\varphi(\bar{x}; \bar{y}) \wedge f(\bar{y}) = \bar{z}]$$

has the desired properties.

We start by proving that $\varphi(\bar{x}; \bar{b})^{\mathfrak{M}^{\mathrm{eq}}} = \psi(\bar{x}; \bar{c})^{\mathfrak{M}^{\mathrm{eq}}}$ where $\bar{c} := f(\bar{b})$. Clearly, every tuple satisfying $\varphi(\bar{x}; \bar{b})$ also satisfies $\psi(\bar{x}; \bar{c})$. Conversely, suppose that $\bar{a}$ satisfies $\psi(\bar{x}; \bar{c})$. Then there is some tuple $\bar{b}' \in f^{-1}(\bar{c})$ such that $\bar{a} \in \varphi(\bar{x}; \bar{b}')^{\mathfrak{M}^{\mathrm{eq}}}$. By definition of f, it follows that $\bar{b}' \in [\bar{b}]_\chi$. Hence, $\varphi(\bar{x}; \bar{b}')^{\mathfrak{M}^{\mathrm{eq}}} = \varphi(\bar{x}; \bar{b})^{\mathfrak{M}^{\mathrm{eq}}}$. Consequently, $\bar{a}$ satisfies $\varphi(\bar{x}; \bar{b})$.

It remains to show that $\bar{c}$ is unique. Hence, suppose that $\bar{c}'$ is some tuple with $\varphi(\bar{x}; \bar{b})^{\mathfrak{M}^{\mathrm{eq}}} = \psi(\bar{x}; \bar{c}')^{\mathfrak{M}^{\mathrm{eq}}}$. As f is surjective, there exists an element $\bar{b}' \in f^{-1}(\bar{c}')$. Since

$$\varphi(\bar{x}; \bar{b}')^{\mathfrak{M}^{\mathrm{eq}}} = \psi(\bar{x}; \bar{c}')^{\mathfrak{M}^{\mathrm{eq}}} = \varphi(\bar{x}; \bar{b})^{\mathfrak{M}^{\mathrm{eq}}} ,$$

it follows that $\mathfrak{M} \vDash \chi(\bar{b}, \bar{b}')$. Consequently, $\bar{c}' = f(\bar{b}') = f(\bar{b}) = \bar{c}$. □

Corollary 2.12. *Every parameter-definable relation in* $\mathbb{M}^{\mathrm{eq}}$ *has a canonical parameter.*

3. *Galois bases*

We can characterise canonical parameters also in a more algebraic way via automorphisms.

Definition 3.1. A *Galois base,* or *canonical base,* of a relation $\mathbb{R} \subseteq \mathbb{M}^{\bar{s}}$ is a set $B \subseteq \mathbb{M}$ such that

$$\pi[\mathbb{R}] = \mathbb{R} \quad \text{iff} \quad \pi \restriction B = \mathrm{id}_B\,, \quad \text{for all } \pi \in \mathrm{Aut}\,\mathbb{M}\,.$$

Remark. According to the definition, B is a Galois base or $\mathbb{R}$ if, and only if, in $\mathrm{Aut}\,\mathbb{M}$ the setwise stabiliser of $\mathbb{R}$ coincides with the pointwise stabiliser of B, i.e., if $\mathrm{Aut}\langle\mathbb{M},\mathbb{R}\rangle = \mathrm{Aut}\,\mathbb{M}_B$.

From the results of Section 1 it follows that, for parameter-definable relations, Galois bases are the same as canonical parameters. But note that the notion of a Galois base also applies to relations that are not definable. Before giving the proof, let us present some technical lemmas. The first one is an immediate consequence of Lemma 1.10.

Lemma 3.2. *If B is a Galois base of a parameter-definable relation $\mathbb{R}$, then $\mathbb{R}$ is definable over B.*

Lemma 3.3. *Let $\mathbb{R} \subseteq \mathbb{M}^{\bar{s}}$ be a relation and $B \subseteq \mathbb{M}$ a set. The following statements are equivalent:*

(1) *B is a Galois base of $\mathbb{R}$ in the structure $\mathbb{M}$.*

(2) *B is a Galois base of $\mathbb{R}$ in the structure $\mathbb{M}^{\mathrm{eq}}$.*

Proof. As the restriction map $\pi \mapsto \pi \restriction M$ is an isomorphism between $\mathfrak{Aut}\,\mathbb{M}^{\mathrm{eq}}$ and $\mathfrak{Aut}\,\mathbb{M}$, the following two statements are equivalent:

- $\pi[\mathbb{R}] = \mathbb{R} \quad \Leftrightarrow \quad \pi \restriction B = \mathrm{id}_B\,, \quad$ for all $\pi \in \mathrm{Aut}\,\mathbb{M}$.
- $\pi[\mathbb{R}] = \mathbb{R} \quad \Leftrightarrow \quad \pi \restriction B = \mathrm{id}_B\,, \quad$ for all $\pi \in \mathrm{Aut}\,\mathbb{M}^{\mathrm{eq}}$.

□

Lemma 3.4. *Let* $\mathbb{R}$ *be a relation and A, B sets.*

(a) *If* $\operatorname{dcl}(A) = \operatorname{dcl}(B)$, *then A is a Galois base of* $\mathbb{R}$ *if, and only if, B is a Galois base of* $\mathbb{R}$.

(b) *If A and B are both Galois bases of* $\mathbb{R}$, *then* $\operatorname{dcl}(A) = \operatorname{dcl}(B)$.

Proof. (a) Suppose that A is a Galois base of $\mathbb{R}$. By Corollary 1.8, it follows that

$$\operatorname{Aut}\mathbb{M}_B = \operatorname{Aut}\mathbb{M}_A = \operatorname{Aut}\langle\mathbb{M},\mathbb{R}\rangle\,.$$

Hence, B is a Galois base of $\mathbb{R}$.

(b) Since both A and B are Galois bases, we have

$$\operatorname{Aut}\mathbb{M}_B = \operatorname{Aut}\langle\mathbb{M},\mathbb{R}\rangle = \operatorname{Aut}\mathbb{M}_A\,.$$

Therefore it follows by Corollary 1.8 that $\operatorname{dcl}(A) = \operatorname{dcl}(B)$. □

With these preparations we can prove that, for parameter-definable relations, Galois bases and canonical parameters are the same.

Proposition 3.5. *Let* $\mathbb{R}$ *be a parameter-definable relation and* $\bar{b}$ *a tuple. The following statements are equivalent:*

(1) $\bar{b}$ *is a Galois base of* $\mathbb{R}$.

(2) $\bar{b}$ *is a canonical parameter of* $\mathbb{R}$.

(3) $\operatorname{dcl}^{\mathrm{eq}}(\bar{b})$ *is the least* $\operatorname{dcl}^{\mathrm{eq}}$*-closed set over which* $\mathbb{R}$ *is definable.*

Proof. (2) ⇒ (1) Suppose that $\psi(\bar{x};\bar{b})$ is a canonical definition of $\mathbb{R}$. To show that $\bar{b}$ is a Galois base of $\mathbb{R}$, consider an automorphism π of $\mathbb{M}$. Then

$$\pi(\bar{b}) = \bar{b} \quad \text{implies} \quad \pi[\mathbb{R}] = \psi(\bar{x};\pi(\bar{b}))^{\mathbb{M}} = \psi(\bar{x};\bar{b})^{\mathbb{M}} = \mathbb{R}\,.$$

Conversely,

$$\pi[\mathbb{R}] = \mathbb{R} \quad \text{implies} \quad \psi(\bar{x}; \pi(\bar{b}))^{\mathbb{M}} = \psi(\bar{x}; \bar{b})^{\mathbb{M}}.$$

By uniqueness of $\bar{b}$, it follows that $\pi(\bar{b}) = \bar{b}$.

(1) ⇒ (2) Suppose that $\bar{b}$ is a Galois base of $\mathbb{R}$. By Lemma 3.2, there exists a formula $\varphi(\bar{x}; \bar{z})$ such that

$$\mathbb{R} = \varphi(\bar{x}; \bar{b})^{\mathbb{M}}.$$

First, let us show that there is no tuple $\bar{b}' \neq \bar{b}$ with

$$\bar{b}' \equiv_{\varnothing} \bar{b} \quad \text{and} \quad \varphi(\bar{x}; \bar{b}')^{\mathbb{M}} = \varphi(\bar{x}; \bar{b})^{\mathbb{M}}.$$

For a contradiction, suppose otherwise. Since $\bar{b}$ and $\bar{b}'$ have the same type, there exists an automorphism π with $\pi(\bar{b}) = \bar{b}'$. It follows that

$$\pi[\mathbb{R}] = \pi[\varphi(\bar{x}; \bar{b})^{\mathbb{M}}] = \varphi(\bar{x}; \pi(\bar{b}))^{\mathbb{M}} = \varphi(\bar{x}; \bar{b}')^{\mathbb{M}} = \mathbb{R}\,.$$

Since $\bar{b}$ is a Galois base of $\mathbb{R}$, this implies that $\pi(\bar{b}) = \bar{b}$. Hence, $\bar{b}' = \bar{b}$. Contradiction.

Set $\Phi(\bar{x}) := \operatorname{tp}(\bar{b})$. We have shown that

$$\Phi(\bar{y}) \cup \Phi(\bar{y}') \cup \{\forall \bar{x}[\varphi(\bar{x}; \bar{y}) \leftrightarrow \varphi(\bar{x}; \bar{y}')]\} \vDash \bar{y} = \bar{y}'\,.$$

By compactness, there exists a finite subset $\Phi_0 \subseteq \Phi$ such that

$$\Phi_0(\bar{y}) \cup \Phi_0(\bar{y}') \cup \{\forall \bar{x}[\varphi(\bar{x}; \bar{y}) \leftrightarrow \varphi(\bar{x}; \bar{y}')]\} \vDash \bar{y} = \bar{y}'\,.$$

Consequently, we obtain a canonical definition of $\mathbb{R}$ by setting

$$\psi(\bar{x}; \bar{b}) := \varphi(\bar{x}; \bar{b}) \wedge \bigwedge \Phi_0(\bar{b})\,.$$

(2) ⇒ (3) Let $\bar{b}$ be a Galois base of $\mathbb{R}$. We have seen in Lemma 3.2 that $\mathbb{R}$ is definable over $\bar{b}$. Suppose that $\mathbb{R}$ is definable over a $\mathrm{dcl}^{\mathrm{eq}}$-closed set $A \subseteq \mathbb{M}^{\mathrm{eq}}$. For $\pi \in \operatorname{Aut} \mathbb{M}^{\mathrm{eq}}$, it follows that

$$\pi \restriction A = \mathrm{id}_A \quad \text{implies} \quad \pi[\mathbb{R}] = \mathbb{R} \quad \text{implies} \quad \pi(\bar{b}) = \bar{b}\,.$$

Consequently, $\operatorname{Aut}\mathbb{M}^{\mathrm{eq}}_A \subseteq \operatorname{Aut}\mathbb{M}^{\mathrm{eq}}_{\bar b}$ and it follows by Corollary 1.8 that $\bar b \subseteq \mathrm{dcl}^{\mathrm{eq}}(A)$.

(3) ⇒ (1) We have seen in Corollary 2.9 that $\mathbb{R}$ has a canonical parameter $e \in \mathbb{M}^{\mathrm{eq}}$. By (3), this implies that $\mathrm{dcl}^{\mathrm{eq}}(\bar b) \subseteq \mathrm{dcl}^{\mathrm{eq}}(e)$. Conversely, since $\mathbb{R}$ is definable over $\bar b$, it follows by the already proved implication (2) ⇒ (3) that $\mathrm{dcl}^{\mathrm{eq}}(e) \subseteq \mathrm{dcl}^{\mathrm{eq}}(\bar b)$. Consequently, $\mathrm{dcl}^{\mathrm{eq}}(e) = \mathrm{dcl}^{\mathrm{eq}}(\bar b)$. Note that, by the already established implication (1) ⇒ (2), e is a Galois base of $\mathbb{R}$. Therefore, we can use Lemma 3.4 (a) to show that $\bar b$ is also a Galois base of $\mathbb{R}$. □

Relations that are not definable still might have a Galois base. Of particular interest are relations that are definable by types.

Definition 3.6. A *Galois base* of a type $\mathfrak{p} \in S^{\bar s}(\mathbb{M})$ is a Galois base of the relation $\mathfrak{p}^{\mathbb{M}}$ defined by it.

For types, Galois bases do not necessarily exists. But if they do, they are unique up to definable equivalence.

Definition 3.7. For a type $\mathfrak{p}$ with Galois base B, we set

$$\mathrm{Gb}(\mathfrak{p}) := \mathrm{dcl}^{\mathrm{eq}}(B)\,.$$

Remark. By the Lemma 3.4, it follows that $\mathrm{Gb}(\mathfrak{p})$ is the maximal Galois base of $\mathfrak{p}$ and that it does not depend on the choice of B.

Lemma 3.8. *Let T be a complete first-order theory and $\mathfrak{p} \in S^{\bar s}(\mathbb{M})$ a type. If $\mathfrak{p}$ is definable over $U \subseteq \mathbb{M}$, it has a Galois base $B \subseteq \mathrm{dcl}^{\mathrm{eq}}(U)$ of size $|B| \leq |T|$.*

Proof. Let $\varphi(\bar x;\bar y)$ be a formula without parameters and let $\delta_\varphi(\bar y)$ be a φ-definition of $\mathfrak{p}$ over U. By Corollary 2.9, the relation $\mathbb{R}_\varphi := (\delta_\varphi)^{\mathbb{M}}$ has a Galois base $b_\varphi \in \mathrm{dcl}^{\mathrm{eq}}(U)$. Set $B := \{\, b_\varphi \mid \varphi \text{ a formula}\,\}$. Then $|B| \leq |T|$ and $B \subseteq \mathrm{dcl}^{\mathrm{eq}}(U)$. To show that B is a Galois base of $\mathfrak{p}$, consider an automorphism $\pi \in \operatorname{Aut}\mathbb{M}^{\mathrm{eq}}$. Then

$$\begin{aligned} \pi(\mathfrak{p}) = \mathfrak{p} \quad &\text{iff} \quad \pi[\mathbb{R}_\varphi] = \mathbb{R}_\varphi\,, \quad \text{for all } \varphi \\ &\text{iff} \quad \pi(b_\varphi) = b_\varphi\,, \quad \text{for all } \varphi \\ &\text{iff} \quad \pi \upharpoonright B = \mathrm{id}_B\,, \end{aligned}$$

as desired. □

Corollary 3.9. *In a stable first-order theory T, every complete type over a set U has a Galois base in* $\mathrm{dcl}^{\mathrm{eq}}(U)$.

Proof. Let $\mathfrak{p}$ be a complete type over U. According to Theorem C3.5.17, $\mathfrak{p}$ is definable over U. Hence, the claim follows by Lemma 3.8. □

Lemma 3.10. *Let $\mathfrak{p} \in S^{\bar{s}}(\mathbb{M})$ be a definable type and $U \subseteq \mathbb{M}$ a set of parameters. Then $\mathfrak{p}$ is definable over U if, and only if,* $\mathrm{Gb}(\mathfrak{p}) \subseteq \mathrm{dcl}^{\mathrm{eq}}(U)$.

Proof. ($\Rightarrow$) follows by Lemma 3.8.

($\Leftarrow$) According to Lemma 3.8, $\mathfrak{p}$ has a Galois base B. Since $\mathfrak{p}$ is definable we can find, for every formula $\varphi(\bar{x};\bar{y})$, a definable relation $\mathbb{R}_\varphi$ such that

$$\varphi(\bar{x};\bar{c}) \in \mathfrak{p} \quad \text{iff} \quad \bar{c} \in \mathbb{R}_\varphi\,.$$

Since $B \subseteq \mathrm{Gb}(\mathfrak{p}) \subseteq \mathrm{dcl}^{\mathrm{eq}}(U)$, it is sufficient to show that $\mathbb{R}_\varphi$ is definable over B. For each automorphism $\pi \in \operatorname{Aut}\mathbb{M}^{\mathrm{eq}}_B$, we have $\pi[\mathfrak{p}] = \mathfrak{p}$. Consequently, $\pi[\mathbb{R}_\varphi] = \mathbb{R}_\varphi$. Therefore, Lemma 3.2 implies that $\mathbb{R}_\varphi$ is definable over B. □

We conclude this section with a characterisation of the algebraic closure in $\mathbb{M}^{\mathrm{eq}}$. We start with an analogue of Lemma 1.10 for the algebraic closure.

Lemma 3.11. *A parameter-definable relation $\mathbb{R}$ has finitely many conjugates over a set $U \subseteq \mathbb{M}$ if, and only if, $\mathbb{R}$ is definable over* $\mathrm{acl}^{\mathrm{eq}}(U)$.

Proof. ($\Leftarrow$) Suppose that $\mathbb{R}$ is definable over $\bar{c} \subseteq \mathrm{acl}^{\mathrm{eq}}(U)$. Then

$$|\{\, \pi[\mathbb{R}] \mid \pi \in \operatorname{Aut}\mathbb{M}^{\mathrm{eq}}_U \,\}| \leq |\{\, \pi(\bar{c}) \mid \pi \in \operatorname{Aut}\mathbb{M}^{\mathrm{eq}}_U \,\}| < \aleph_0 \,.$$

Hence, $\mathbb{R}$ has only finitely many conjugates over U.

($\Rightarrow$) Suppose that $\mathbb{R}$ has only finitely many conjugates over U and let $\bar{b}$ be a Galois base of $\mathbb{R}$. Then

$$|\{\, \pi(\bar{b}) \mid \pi \in \operatorname{Aut}\mathbb{M}^{\mathrm{eq}}_U \,\}| \leq |\{\, \pi[\mathbb{R}] \mid \pi \in \operatorname{Aut}\mathbb{M}^{\mathrm{eq}}_U \,\}| < \aleph_0 \,.$$

By Theorem 1.6, it follows that $\bar{b} \subseteq \mathrm{acl}^{\mathrm{eq}}(U)$. Furthermore, we have seen in Lemma 3.2 that $\mathbb{R}$ is definable over $\bar{b}$. □

The algebraic closure of a set U in $\mathbb{M}^{\mathrm{eq}}$ can be characterised as follows.

Definition 3.12. Let $U \subseteq \mathbb{M}$ be a set of parameters and $\bar{s}$ a finite tuple of sorts. We denote by $\mathrm{FE}^{\bar{s}}(U)$ the set of all formulae $\chi(\bar{x}, \bar{y})$ over U where $\bar{x}$ and $\bar{y}$ have sort $\bar{s}$ such that $\chi^{\mathbb{M}}$ is an equivalence relation on $\mathbb{M}^{\bar{s}}$ with finitely many classes.

Lemma 3.13. *Let $\bar{a}, \bar{b} \in \mathbb{M}^{\bar{s}}$ be finite tuples and $U \subseteq \mathbb{M}$ a set of parameters. Then*

$$\bar{a} \equiv_{\mathrm{acl}^{\mathrm{eq}}(U)} \bar{b} \quad \text{iff} \quad \mathbb{M} \vDash \chi(\bar{a}, \bar{b}) \quad \text{for all } \chi \in \mathrm{FE}^{\bar{s}}(U)\,.$$

Proof. ($\Rightarrow$) Let $\chi \in \mathrm{FE}^{\bar{s}}(U)$ and let $\mathbb{B} := [\bar{b}]_{\chi^{\mathbb{M}}} \subseteq \mathbb{M}^{\bar{s}}$ be the $\chi^{\mathbb{M}}$-class of $\bar{b}$. The conjugates of $\mathbb{B}$ over U are $\chi^{\mathbb{M}}$-classes. Since there are only finitely many such classes, it follows by Lemma 3.11 (b) that $\mathbb{B}$ is definable over $\mathrm{acl}^{\mathrm{eq}}(U)$. Therefore, we can use Proposition 3.5 and Corollary 2.9 to find a canonical definition $\psi(\bar{x}; e)$ of $\mathbb{B}$ where $e \in \mathrm{dcl}^{\mathrm{eq}}(\mathrm{acl}^{\mathrm{eq}}(U)) = \mathrm{acl}^{\mathrm{eq}}(U)$. Since

$$\bar{a} \equiv_{\mathrm{acl}^{\mathrm{eq}}(U)} \bar{b}\,,$$

it follows that

$$\mathbb{M} \vDash \psi(\bar{b}; e) \quad \text{implies} \quad \mathbb{M} \vDash \psi(\bar{a}; e)\,.$$

Hence, $\bar{a} \in \mathbb{B}$ implies $\mathbb{M} \vDash \chi(\bar{a}, \bar{b})$.

($\Leftarrow$) Suppose that $\mathbb{M} \vDash \varphi(\bar{a}; \bar{c})$, for $\bar{c} \subseteq \mathrm{acl}^{\mathrm{eq}}(U)$. We have to show that $\mathbb{M} \vDash \varphi(\bar{b}; \bar{c})$. There exists a formula $\psi(\bar{x})$ over U such that $\psi^{\mathbb{M}}$ is a finite set containing $\bar{c}$. The formula

$$\chi(\bar{x}, \bar{y}) := (\forall \bar{z}.\psi(\bar{z}))[\varphi(\bar{x}; \bar{z}) \leftrightarrow \varphi(\bar{y}; \bar{z})]$$

defines an equivalence relation with finitely many classes. Therefore, $\chi \in \mathrm{FE}^{\bar{s}}(U)$ and $\mathbb{M} \vDash \chi(\bar{a}, \bar{b})$. Since $\bar{c} \in \psi^{\mathbb{M}}$, it follows that

$$\mathbb{M} \vDash \varphi(\bar{a}; \bar{c}) \quad \text{implies} \quad \mathbb{M} \vDash \varphi(\bar{b}; \bar{c}) \,. \qquad \square$$

4. *Elimination of imaginaries*

In the abstract we can capture the property of $\mathbb{M}^{\mathrm{eq}}$ exhibited in Proposition 2.10 by the following definition.

Definition 4.1. A structure $\mathfrak{M}$ has *uniform elimination of imaginaries* if, for every equivalence formula $\chi(\bar{x}, \bar{y})$ of type $\bar{s}$, there exist sorts $\bar{t}$ and a definable function $f : M^{\bar{s}} \to M^{\bar{t}}$ such that $\ker f = \chi^{\mathfrak{M}}$.

We say that a theory T has *uniform elimination of imaginaries* if every model of T does.

We have shown in Proposition 2.10 that structures of the form $\mathfrak{M}^{\mathrm{eq}}$ have uniform elimination of imaginaries.

Proposition 4.2. *Every structure of the form* $\mathfrak{M}^{\mathrm{eq}}$ *has uniform elimination of imaginaries.*

Exercise 4.1. Show that the structure $\langle \mathbb{N}, +, \cdot \rangle$ has uniform elimination of imaginaries.

Frequently, the following weaker condition is equivalent to having uniform elimination of imaginaries.

Definition 4.3. A structure $\mathfrak{M}$ has *elimination of imaginaries* if, for each equivalence formula $\chi(\bar{x}, \bar{y})$ of type $\bar{s}$ and all tuples $\bar{a} \in M^{\bar{s}}$, the equivalence class $[\bar{a}]_\chi$ has a canonical parameter.

We say that a theory T has *elimination of imaginaries* if every model of T does.

For structures where $\mathrm{dcl}(\varnothing)$ is non-trivial, elimination of imaginaries already implies uniform elimination of imaginaries.

Lemma 4.4. *Let $\mathfrak{M}$ be a structure. The following statements are equivalent:*

(1) *$\mathfrak{M}$ has uniform elimination of imaginaries.*

(2) *$\mathfrak{M}$ has elimination of imaginaries and at least one of the following conditions holds:*
   - *There is some sort u with $|\mathrm{dcl}(\varnothing) \cap M^u| > 1$.*
   - *$|M^s| \leq 1$, for all sorts s.*

Proof. (1) ⇒ (2) To show that $\mathfrak{M}$ has elimination of imaginaries, consider an equivalence formula $\chi(\bar{x}, \bar{y})$ and a tuple $\bar{a}$ in M. By (1), there exists a definable function f with $\ker f = \chi^{\mathfrak{M}}$. Then $[\bar{a}]_\chi$ has the canonical definition

$$\psi(\bar{x}; \bar{b}) := (f(\bar{x}) = \bar{b}) \quad \text{where} \quad \bar{b} := f(\bar{a}) .$$

To conclude the proof, suppose that there is some sort s with $|M^s| > 1$. We have to find a sort u with $|\mathrm{dcl}(\varnothing) \cap M^u| > 1$. Consider the equivalence formula

$$\chi(xx', yy') := (x = x') \leftrightarrow (y = y')$$

of type ss. By (1), there exists a definable function f with $\ker f = \chi^{\mathfrak{M}}$. Fix distinct elements $c, d \in M^s$. It follows that the tuples $\bar{a} := f(c, c)$ and $\bar{b} := f(c, d)$ are definable and distinct. Fixing an index i with $a_i \neq b_i$, we obtain distinct elements a_i and b_i in $\mathrm{dcl}(\varnothing)$ of the same sort.

(2) ⇒ (1) If $|M^s| \leq 1$, for all sorts s, every equivalence formula χ defines the equality relation. Hence, the identity function has kernel $\chi^{\mathfrak{M}}$ and we are done.

It therefore remains to consider the case where $|\mathrm{dcl}(\varnothing) \cap M^u| > 1$, for some sort u. Let $\chi(\bar{x}, \bar{y})$ be an equivalence formula of type $\bar{s}$. For every tuple $\bar{a} \in M^{\bar{s}}$, fix a canonical definition $\delta_{\bar{a}}(\bar{x}; \bar{b}_{\bar{a}})$ of $[\bar{a}]_\chi$. Let $\bar{t}_{\bar{a}}$ be the sorts of $\bar{b}_{\bar{a}}$. We obtain a formula

$$\psi_{\bar{a}}(\bar{x}; \bar{y}) := \delta_{\bar{a}}(\bar{x}, \bar{y}) \wedge \forall \bar{z}[\delta_{\bar{a}}(\bar{z}; \bar{y}) \leftrightarrow \chi(\bar{x}, \bar{z})]$$

that defines a partial function $f_{\bar{a}} : U_{\bar{a}} \to M^{\bar{t}_{\bar{a}}}$ with kernel $\chi^{\mathfrak{M}}|_{U_{\bar{a}}}$. Note that the domain $U_{\bar{a}}$ of $f_{\bar{a}}$ is a union of χ-classes and that it is definable by the formula

$$\vartheta_{\bar{a}}(\bar{x}) := \exists \bar{y} \psi_{\bar{a}}(\bar{x}, \bar{y}) .$$

Hence,

$$M^{\bar{s}} = \bigcup_{\bar{a} \in M^{\bar{s}}} U_{\bar{a}} \quad \text{implies} \quad \mathrm{Th}(\mathfrak{M}) \vDash \bigvee_{\bar{a} \in M^{\bar{s}}} \vartheta_{\bar{a}} .$$

By compactness, there are finitely many tuples $\bar{a}_0, \dots, \bar{a}_n \in M^{\bar{s}}$ such that $M^{\bar{s}} = U_{\bar{a}_0} \cup \dots \cup U_{\bar{a}_n}$. Fix distinct elements $c, d \in \mathrm{dcl}(\varnothing) \cap M^u$. The formula

$$\begin{aligned} &\varphi(\bar{x}; \bar{y}_0, \dots, \bar{y}_n, \bar{z}) := \\ &\quad \bigvee_{i \leq n} \Big[\psi_{\bar{a}_i}(\bar{x}; \bar{y}_i) \wedge \bar{x} \in U_{\bar{a}_i} \setminus (U_{\bar{a}_0} \cup \dots \cup U_{\bar{a}_{i-1}}) \\ &\qquad\qquad \wedge \bigwedge_{j \neq i} \bar{y}_j = \langle c, \dots, c \rangle \\ &\qquad\qquad \wedge \bar{z} = \langle \underbrace{c, \dots, c}_{i \text{ times}}, d \dots, d \rangle \Big] \end{aligned}$$

defines a function $f : M^{\bar{s}} \to M^{\bar{t}_{\bar{a}_0} \dots \bar{t}_{\bar{a}_n} u \dots u}$ with $\ker f = \chi^{\mathfrak{M}}$. □

As an example, we consider o-minimal structures and, in particular, real closed fields. We say that a theory T has *definable Skolem functions* if, for every formula $\varphi(\bar{x}, y)$, there exists a definable function f such that

$$T \vDash \forall \bar{x}[\exists y \varphi(\bar{x}, y) \to \varphi(\bar{x}, f(\bar{x}))] .$$

Proposition 4.5. *Every o-minimal structure $\mathfrak{M}$ with definable Skolem functions has elimination of imaginaries.*

Proof. We start by proving that every parameter-definable set $P \subseteq M$ has a canonical definition. Suppose that $P \subseteq M$ is parameter-definable. By o-minimality, P is of the form

$$P = (a_0, b_0) \cup \dots \cup (a_{m-1}, b_{m-1}) \cup \{c_0, \dots, c_{n-1}\},$$

for elements $a_i, b_i, c_i \in M$ satisfying

$$a_0 < b_0 < a_1 < b_1 < \dots < a_{m-1} < b_{m-1} \quad \text{and} \quad c_0 < \dots < c_{n-1}.$$

Fix such a decomposition of P where m and n are minimal. Then

$$\begin{aligned} \psi(x; \bar{a}, \bar{b}, \bar{c}) := \Big[\bigvee_{i<m} (a_i < x \wedge x < b_i) \vee \bigvee_{i<n} x = c_i \Big] \\ \wedge \Big[\bigwedge_{i<m} a_i < b_i \wedge \bigwedge_{i<m-1} b_i < a_{i+1} \wedge \bigwedge_{i<n-1} c_i < c_{i+1} \Big] \end{aligned}$$

is a canonical definition of P.

To show that $\mathfrak{M}$ has elimination of imaginaries, let $\chi(\bar{x}, \bar{y})$ be an equivalence formula of type $\bar{s}$ and let $\bar{a} \in M^{\bar{s}}$. To find a canonical definition of $[\bar{a}]_\chi$, we define, by induction on $i < n := |\bar{s}|$, a formula $\psi_i(y_i; \bar{z}_i)$, parameters $\bar{b}_i$, and a definable function s_i such that

- $\psi_i(y_i; \bar{b}_i)$ is a canonical definition of the relation defined by

$$\begin{aligned} &\vartheta_i(y_i; \bar{a}, \bar{b}_0, \dots, \bar{b}_{i-1}) := \\ &\quad \exists y_{i+1} \cdots \exists y_{n-1} \chi(\bar{a}, s_0(\bar{b}_0), \dots, s_{i-1}(\bar{b}_{i-1}), \\ &\qquad y_i, y_{i+1}, \dots, y_{n-1}), \end{aligned}$$

- $\mathfrak{M} \vDash \psi_i(s_i(\bar{b}_i); \bar{b}_i)$.

Suppose that we have already defined the formulae $\psi_0(y_0; \bar{b}_0), \dots, \psi_{i-1}(y_{i-1}; \bar{b}_{i-1})$ and the functions $s_0, \dots, s_{i-1}$. Since ϑ_i defines a set, we can use the statement we have proved above to find a canonical definition $\psi_i(y_i;\bar{b}_i)$ of $\vartheta_i^{\mathfrak{M}}(y_i;\bar{a},\bar{b}_0,\dots,\bar{b}_{i-1})$. Let s_i be a definable Skolem function for the formula $\psi_i(y_i;\bar{z}_i)$. This concludes the inductive step.

We claim that the formula

$$\begin{aligned}\psi(\bar{x};\bar{b}_0,\dots,\bar{b}_{n-1}) :=& \\ &\chi(\bar{x},\ s_0(\bar{b}_0),\dots,s_{n-1}(\bar{b}_{n-1})) \\ &\wedge \bigwedge_{i<n} \forall y_i\big[\psi_i(y_i;\bar{b}_i) \leftrightarrow \vartheta_i(y_i;\bar{x},\bar{b}_0,\dots,\bar{b}_{i-1})\big]\end{aligned}$$

is a canonical definition of $[\bar{a}]_\chi$. By construction, we have

$$\psi(\bar{x};\bar{b}_0,\dots,\bar{b}_{n-1})^{\mathfrak{M}} = [\bar{a}]_\chi\,.$$

Suppose that $\bar{b}'_0,\dots,\bar{b}'_{n-1}$ are tuples such that

$$\psi(\bar{x};\bar{b}'_0,\dots,\bar{b}'_{n-1})^{\mathfrak{M}} = [\bar{a}]_\chi\,.$$

Then

$$\psi_i(y_i;\bar{b}'_i)^{\mathfrak{M}} = \vartheta_i(\bar{y}_i;\bar{a},\bar{b}'_0,\dots,b'_{i-1})^{\mathfrak{M}}\,.$$

By choice of ψ_i we can use induction on i to show that $\bar{b}'_i = \bar{b}_i$. □

Corollary 4.6. *The theory* RCF *of real closed fields has uniform elimination of imaginaries.*

Proof. After we have shown that RCF has definable Skolem functions, we can use Proposition 4.5 to show that RCF has elimination of imaginaries. Since $0, 1 \in \mathrm{dcl}(\varnothing)$, it therefore follows by Lemma 4.4 that it even has uniform elimination of imaginaries.

Hence, it remains to show that RCF has definable Skolem functions. Let $\varphi(\bar{x}, y)$ be a formula. By o-minimality, for every choice of values $\bar{c}$ for the variables $\bar{x}$, the relation $\varphi(\bar{a}, y)^{\mathbb{M}}$ is of the form

$$\varphi(\bar{c}, y)^{\mathbb{M}} = (a_0, b_0) \cup \dots \cup (a_{m-1}, b_{m-1}) \cup \{d_0,\dots,d_{n-1}\}\,,$$

for elements $a_i, b_i, c_i \in M$ satisfying

$$a_0 < b_0 < a_1 < b_1 < \cdots < a_{m-1} < b_{m-1} \quad \text{and} \quad d_0 < \cdots < d_{n-1}\,.$$

Furthermore, it follows by Theorem D3.3.11 that there exists a bound $k < \omega$ such that, for every tuple $\bar{c}$, we can choose a decomposition as above where the numbers m and n are less than k.

Let $\psi(\bar{x}; y)$ be a formula stating that, for the given value of $\bar{x}$, there are numbers $m, n < k$ and tuples $\bar{a}, \bar{b}, \bar{d}$ such that

- $\varphi(\bar{x}, y')^{\mathbb{M}} = (a_0, b_0) \cup \cdots \cup (a_m, b_m) \cup \{d_0, \ldots, d_n\}$,
- m and n are the minimal numbers such that $\varphi(\bar{x}, y')^{\mathbb{M}}$ can be written in this form,
- $a_0 < b_0 < a_1 < b_1 < \cdots < a_{m-1} < b_{m-1}$ and $d_0 < \cdots < d_{n-1}$,
- $y = \begin{cases} d_0 & \text{if } n > 0\,, \\ (a_0 + b_0)/2 & \text{if } n = 0\,,\ m > 0\,, \text{ and } -\infty < a_0 < b_0 < \infty\,, \\ b_0 - 1 & \text{if } n = 0\,,\ m > 0\,, \text{ and } -\infty = a_0 < b_0 < \infty\,, \\ a_0 + 1 & \text{if } n = 0\,,\ m > 0\,, \text{ and } -\infty < a_0 < b_0 = \infty\,, \\ 0 & \text{otherwise}\,. \end{cases}$

Then $\psi(\bar{x}, y)$ defines a Skolem function for $\varphi(\bar{x}, y)$. ☐

We can use Galois bases to characterise theories with elimination of imaginaries.

Proposition 4.7. *Let T be a complete first-order theory. The following statements are equivalent:*

(1) *T has elimination of imaginaries.*

(2) *Every parameter-definable relation has a canonical parameter.*

(3) *Every parameter-definable relation has a finite Galois base.*

(4) *For every parameter-definable relation $\mathbb{R}$, there exists a least* $\mathrm{dcl}^{\mathrm{eq}}$-*closed set $B \subseteq \mathbb{M}$ over which $\mathbb{R}$ is definable.*

(5) *For every imaginary element* $e \in \mathbb{M}^{eq}$, *there is a finite set* $B \subseteq \mathbb{M}$ *with* $\mathrm{dcl}^{eq}(e) = \mathrm{dcl}^{eq}(B)$.

Proof. (3) ⇒ (4) ⇔ (2) follows by Proposition 3.5.

(2) ⇒ (1) Let $\chi(\bar{x}, \bar{y})$ be an equivalence formula. If every parameter-definable relation has a canonical parameter then, in particular, this is true for every relation of the form $[\bar{a}]_\chi$.

(1) ⇒ (5) Let $e \in \mathbb{M}^{eq}_\chi$ be an imaginary element and $\mathbb{E} := p_\chi^{-1}(e)$ the corresponding equivalence class. Since T has elimination of imaginaries, there exists a canonical definition $\psi(\bar{x}; \bar{b})$ of $\mathbb{E}$. Obviously, we can choose the tuple $\bar{b}$ to be finite. According to Proposition 3.5, $\bar{b}$ is a Galois base of $\mathbb{E}$. Note that, in the structure $\mathbb{M}^{eq}$, $\{e\}$ is a Galois base of $\mathbb{E}$. Consequently, it follows by Lemmas 3.3 and 3.4 that

$$\mathrm{dcl}^{eq}(e) = \mathrm{dcl}^{eq}(\bar{b}) .$$

(5) ⇒ (3) Let $\mathbb{R}$ be a parameter-definable relation. We fix a formula $\varphi(\bar{x}; \bar{c})$ with parameters $\bar{c}$ defining $\mathbb{R}$. Let $\chi(\bar{y}, \bar{y}')$ be the parameter equivalence for $\varphi(\bar{x}; \bar{y})$ and set $e := [\bar{c}]_\chi$. By assumption, there exists a finite set $B \subseteq \mathbb{M}$ such that $\mathrm{dcl}^{eq}(e) = \mathrm{dcl}^{eq}(B)$. We claim that B is a Galois base of $\mathbb{R}$. Note that, by Lemma 3.3, it is sufficient to prove that B is a Galois base of $\mathbb{R}$ in the structure $\mathbb{M}^{eq}$. Furthermore, it follows by Lemma 2.8 and Proposition 3.5 that e is a Galois base of $\mathbb{R}$. Therefore, Lemma 3.4 (a) implies that B is also a Galois base of $\mathbb{R}$. □

5. *Weak elimination of imaginaries*

In this section we take a look at a weaker condition than elimination of imaginaries.

Definition 5.1. (a) A tuple $\bar{c}$ is a *weak canonical parameter* of a relation $\mathbb{R}$ if there exist a formula $\psi(\bar{x}; \bar{y})$ such that $\bar{c}$ is one of only finitely many tuples satisfying

$$\psi(\bar{x}; \bar{c})^{\mathbb{M}} = \mathbb{R} .$$

In this case, we call the formula $\psi(\bar{x};\bar{c})$ a *weak canonical definition* of $\mathbb{R}$.

(b) A complete first-order theory T has *weak elimination of imaginaries* if, for each equivalence formula $\chi(\bar{x},\bar{y})$ of type $\bar{s}$ and all tuples $\bar{a} \in \mathbb{M}^{\bar{s}}$, the equivalence class $[\bar{a}]_\chi$ has a weak canonical parameter.

We start with an analogue of Proposition 3.5.

Lemma 5.2. *Let $\mathbb{R}$ be a parameter-definable relation and U a set. The following statements are equivalent:*

(1) $\mathbb{R}$ *has a weak canonical parameter $\bar{c}$ with* $\mathrm{acl}(\bar{c}) = \mathrm{acl}(U)$.

(2) $\mathrm{acl}(U)$ *is the least algebraically closed set over which $\mathbb{R}$ is definable.*

Proof. (1) ⇒ (2) Let $\psi(\bar{x};\bar{c})$ be a weak canonical definition of $\mathbb{R}$. We claim that $\mathrm{acl}(\bar{c})$ is the least algebraically closed set over which $\mathbb{R}$ is definable. Obviously, $\mathbb{R}$ is definable over $\mathrm{acl}(\bar{c})$. To show that $\mathrm{acl}(\bar{c})$ is the least such set, let $\varphi(\bar{x};\bar{b})$ be an arbitrary formula defining $\mathbb{R}$. We have to prove that $\mathrm{acl}(\bar{c}) \subseteq \mathrm{acl}(\bar{b})$. The formula

$$\vartheta(\bar{y};\bar{b}) := \forall\bar{x}[\psi(\bar{x};\bar{y}) \leftrightarrow \varphi(\bar{x};\bar{b})]$$

defines the finite set $\{\, \bar{c}' \mid \psi(\bar{x};\bar{c}')^{\mathbb{M}} = \mathbb{R} \,\}$. This implies that $\bar{c} \subseteq \mathrm{acl}(\bar{b})$, as desired.

(2) ⇒ (1) Suppose that $\mathrm{acl}(U)$ is the least algebraically closed set over which $\mathbb{R}$ is definable. Fix a formula $\psi(\bar{x};\bar{c})$ with parameters $\bar{c} \subseteq \mathrm{acl}(U)$ defining $\mathbb{R}$. Note that, by assumption on U, it follows that $\mathrm{acl}(\bar{c}) = \mathrm{acl}(U)$.

We start by proving that there are only finitely many tuples $\bar{c}'$ such that

$$\bar{c}' \equiv_\varnothing \bar{c} \quad \text{and} \quad \psi(\bar{x};\bar{c}')^{\mathbb{M}} = \mathbb{R}\,.$$

For a contradiction, suppose otherwise. By compactness, we can then find a tuple $\bar{c}'$ such that

$$\bar{c}' \nsubseteq \mathrm{acl}(\bar{c})\,, \quad \bar{c}' \equiv_\varnothing \bar{c}\,, \quad \text{and} \quad \psi(\bar{x};\bar{c}')^{\mathbb{M}} = \mathbb{R}\,.$$

Since $\mathbb{R}$ is definable over $\bar{c}'$ it follows by assumption on U that

$$\bar{c} \subseteq \operatorname{acl}(U) \subseteq \operatorname{acl}(\bar{c}') .$$

As $\bar{c}' \equiv_\emptyset \bar{c}$, there exists an automorphism π with $\pi(\bar{c}') = \bar{c}$. Setting $\bar{c}'' := \pi(\bar{c})$ it follows that

$$\bar{c} \nsubseteq \operatorname{acl}(\bar{c}'') \quad \text{and} \quad \bar{c}'' \subseteq \operatorname{acl}(\bar{c}) ,$$

Since, for every tuple $\bar{a}$,

$$\begin{aligned} \mathbb{M} \vDash \varphi(\bar{a}; \bar{c}'') \quad &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}; \pi(\bar{c})) \\ &\text{iff} \quad \mathbb{M} \vDash \varphi(\pi^{-1}(\bar{a}); \bar{c}) \\ &\text{iff} \quad \mathbb{M} \vDash \varphi(\pi^{-1}(\bar{a}); \bar{c}') \\ &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}; \pi(\bar{c}')) \quad \text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}; \bar{c}') , \end{aligned}$$

it furthermore follows that $\psi(\bar{x}; \bar{c}'')^{\mathbb{M}} = \mathbb{R}$. But, by assumption on U, this implies that $\bar{c} \subseteq \operatorname{acl}(U) \subseteq \operatorname{acl}(\bar{c}'')$. A contradiction.

Set $\Phi(\bar{y}) := \operatorname{tp}(\bar{c})$. We have shown that there exists a number $n < \omega$ such that

$$\Phi(\bar{y}_0) \cup \dots \cup \Phi(\bar{y}_n) \cup \{ \forall \bar{x} [\psi(\bar{x}; \bar{y}_i) \leftrightarrow \psi(\bar{x}; \bar{y}_k)] \mid i, k \leq n \}$$

is inconsistent. By compactness, we can find a finite subset $\Phi_0 \subseteq \Phi$ such that

$$\Phi_0(\bar{y}_0) \cup \dots \cup \Phi_0(\bar{y}_n) \cup \{ \forall \bar{x} [\psi(\bar{x}; \bar{y}_i) \leftrightarrow \psi(\bar{x}; \bar{y}_k)] \mid i, k \leq n \}$$

is already inconsistent. Consequently, the formula

$$\psi(\bar{x}; \bar{c}) \wedge \bigwedge \Phi_0(\bar{c})$$

is a weak canonical definition of $\mathbb{R}$ with $\operatorname{acl}(\bar{c}) = \operatorname{acl}(U)$. □

Corollary 5.3. *If $\bar{a}$ and $\bar{b}$ are weak canonical parameters of a relation $\mathbb{R}$, then* $\operatorname{acl}(\bar{a}) = \operatorname{acl}(\bar{b})$.

For relations that do have a Galois base, we can be more precise.

Lemma 5.4. *Let $\mathbb{R}$ be a parameter-definable relation with Galois base $\bar{b}$. A tuple $\bar{c}$ is a weak canonical parameter of $\mathbb{R}$ if, and only if,*

$$\bar{b} \subseteq \operatorname{dcl}(\bar{c}) \quad \textit{and} \quad \bar{c} \subseteq \operatorname{acl}(\bar{b})\,.$$

Proof. By Proposition 3.5, we can fix a canonical definition $\hat{\psi}(\bar{x};\bar{b})$ of $\mathbb{R}$.

($\Rightarrow$) Suppose that $\psi(\bar{x};\bar{c})$ is a weak canonical definition of $\mathbb{R}$. Then $\bar{b} \subseteq \operatorname{dcl}(\bar{c})$ since $\bar{b}$ is the unique tuple satisfying

$$\vartheta(\bar{z};\bar{c}) := \forall \bar{x}[\psi(\bar{x};\bar{c}) \leftrightarrow \hat{\psi}(\bar{x};\bar{z})]\,.$$

Furthermore, $\bar{c} \subseteq \operatorname{acl}(\bar{b})$ since the formula

$$\varphi(\bar{y};\bar{b}) := \forall \bar{x}[\psi(\bar{x};\bar{y}) \leftrightarrow \hat{\psi}(\bar{x};\bar{b})]$$

defines a finite set containing $\bar{c}$.

($\Leftarrow$) Let us first consider the special case where $\mathbb{R} = \varnothing$. Then $\varnothing$ is a Galois base of $\mathbb{R}$ and it follows by Lemma 3.4 that $\bar{b} \subseteq \operatorname{dcl}(\varnothing)$. Hence, $\bar{c} \subseteq \operatorname{acl}(\varnothing)$ and there exists a formula $\vartheta(\bar{y})$ that defines a finite relation containing the tuple $\bar{c}$. It follows that the formula

$$\psi(\bar{x};\bar{c}) := \neg\vartheta(\bar{c})$$

is a weak canonical definition of $\mathbb{R} = \varnothing$.

It remains to consider the case where $\mathbb{R} \neq \varnothing$. Fix formulae $\vartheta(\bar{z};\bar{y})$ and $\varphi(\bar{y};\bar{z})$ such that $\vartheta(\bar{z};\bar{c})^{\mathbb{M}} = \{\bar{b}\}$ and $\varphi(\bar{y};\bar{b})^{\mathbb{M}}$ is a finite set containing $\bar{c}$. We claim that the formula

$$\psi(\bar{x};\bar{c}) := \exists \bar{z}\big[\vartheta(\bar{z};\bar{c}) \wedge \hat{\psi}(\bar{x};\bar{z}) \wedge \varphi(\bar{c};\bar{z})\big]$$

is a weak canonical definition of $\mathbb{R}$. Clearly, $\psi(\bar{x};\bar{c})^{\mathbb{M}} = \mathbb{R}$. Furthermore, suppose that $\bar{c}'$ is a tuple such that $\psi(\bar{x};\bar{c}')^{\mathbb{M}} = \mathbb{R}$. Fix a tuple $\bar{a} \in \mathbb{R}$ and let $\bar{b}'$ be a tuple such that

$$\mathbb{M} \vDash \vartheta(\bar{b}';\bar{c}') \wedge \hat{\psi}(\bar{a};\bar{b}') \wedge \varphi(\bar{c}';\bar{b}')\,.$$

Then $\mathbb{R} = \psi(\bar{x}; \bar{c}')^{\mathbb{M}} = \hat{\psi}(\bar{x}; \bar{b}')^{\mathbb{M}}$ implies that $\bar{b}' = \bar{b}$. Hence, we have $\mathbb{M} \vDash \varphi(\bar{c}'; \bar{b})$. Since there are only finitely many such tuples $\bar{c}'$, it follows that $\psi(\bar{x}; \bar{c})^{\mathbb{M}}$ is a weak canonical definition of $\mathbb{R}$. □

We obtain a characterisation of theories with weak elimination of imaginaries along the same lines as Proposition 4.7.

Proposition 5.5. *Let T be a complete first-order theory. The following statements are equivalent:*

(1) *T has weak elimination of imaginaries.*

(2) *All parameter-definable relations have weak canonical parameters.*

(3) *For every parameter-definable relation $\mathbb{R}$, there is a least algebraically closed set over which $\mathbb{R}$ is definable.*

(4) *For every element $e \in \mathbb{M}^{\mathrm{eq}}$, there is a finite set $B \subseteq \mathbb{M}$ such that*

$$e \in \operatorname{dcl}^{\mathrm{eq}}(B) \quad \textit{and} \quad B \subseteq \operatorname{acl}^{\mathrm{eq}}(e)\,.$$

(5) *For every imaginary element $e \in \mathbb{M}^{\mathrm{eq}}$, there exists a finite tuple $\bar{s}$ of sorts and a finite relation $C \subseteq \mathbb{M}^{\bar{s}}$ such that*

$$\operatorname{dcl}^{\mathrm{eq}}(e) = \operatorname{dcl}^{\mathrm{eq}}(B)\,, \quad \textit{for every Galois base } B \textit{ of } C\,.$$

Proof. (4) ⇒ (1) Let $e \in \mathbb{M}^{\mathrm{eq}}_{\chi}$ be an imaginary element and $\mathbb{E} := p_{\chi}^{-1}(e)$ its equivalence class. By assumption, there exists a finite tuple $\bar{c} \subseteq \mathbb{M}$ such that $e \in \operatorname{dcl}^{\mathrm{eq}}(\bar{c})$ and $\bar{c} \subseteq \operatorname{acl}^{\mathrm{eq}}(e)$. Since e is a Galois base of $\mathbb{E}$ it follows by Lemma 5.4 that $\bar{c}$ is a weak canonical parameter of $\mathbb{E}$.

(1) ⇒ (3) Let $\mathbb{R}$ be a relation defined by the formula $\varphi(\bar{x}; \bar{b})$ and let χ be the parameter equivalence of φ. By assumption, there exists a finite relation C and a formula $\psi(\bar{z}; \bar{y})$ such that

$$\psi(\bar{z}; \bar{c})^{\mathbb{M}} = [\bar{b}]_{\chi} \quad \text{iff} \quad \bar{c} \in C\,.$$

We claim that $\operatorname{acl}(\bigcup C)$ is the desired algebraically closed set.

First, note that $\mathbb{R}$ is defined over $\bar{c} \subseteq \mathrm{acl}(\bigcup C)$ by the formula

$$\vartheta(\bar{x};\bar{c}) := \exists\bar{z}\big[\psi(\bar{z};\bar{c}) \wedge \varphi(\bar{x};\bar{z})\big].$$

Next, suppose that A is an algebraically closed set such that $\mathbb{R}$ is definable over A. For every $\pi \in \mathrm{Aut}\,\mathbb{M}$, it follows that

$$\begin{aligned}
\pi \restriction A = \mathrm{id}_A \quad &\Rightarrow \quad \pi[\mathbb{R}] = \mathbb{R} \\
&\Rightarrow \quad \varphi(\bar{x};\pi(\bar{b}'))^{\mathbb{M}} = \varphi(\bar{x};\bar{b}')^{\mathbb{M}}, \quad \text{for all } \bar{b}' \in [\bar{b}]_\chi \\
&\Rightarrow \quad \pi[\bar{b}]_\chi = [\bar{b}]_\chi \\
&\Rightarrow \quad \pi[\psi(\bar{x};\bar{c})^{\mathbb{M}}] = \psi(\bar{x};\bar{c})^{\mathbb{M}}, \quad \text{for all } \bar{c} \in C \\
&\Rightarrow \quad \pi[C] = C.
\end{aligned}$$

Since C is finite, it follows that every tuple $\bar{c} \in C$ has finitely many conjugates over A. Consequently, Theorem 1.6 implies that $\bigcup C \subseteq \mathrm{acl}(A)$.

(3) ⇒ (2) Let $\mathbb{R}$ be a parameter-definable relation. By assumption, there exists a least algebraically closed set U over which $\mathbb{R}$ is definable. Hence, we can apply Lemma 5.2 to obtain a weak canonical parameter $\bar{c} \subseteq U$ of $\mathbb{R}$.

(2) ⇒ (5) Let $e \in \mathbb{M}^{\mathrm{eq}}_\chi$ be an imaginary element and $\mathbb{E} := p_\chi^{-1}(e)$ its equivalence class. By assumption, $\mathbb{E}$ has a weak canonical definition $\psi(\bar{x};\bar{c})$. Obviously, we may assume that $\bar{c}$ is a finite tuple. Set

$$C := \{\, \bar{c}' \mid \psi(\bar{x};\bar{c}')^{\mathbb{M}} = \mathbb{E} \,\}.$$

For an automorphism $\pi \in \mathrm{Aut}\,\mathbb{M}^{\mathrm{eq}}$, it follows that

$$\begin{aligned}
\pi(e) = e \quad &\text{iff} \quad \pi[\mathbb{E}] = \mathbb{E} \\
&\text{iff} \quad \psi(\bar{x};\pi(\bar{c}))^{\mathbb{M}} = \psi(\bar{x};\bar{c})^{\mathbb{M}}, \quad \text{for all } \bar{c} \in C \\
&\text{iff} \quad \pi[C] = C.
\end{aligned}$$

Hence, e is a Galois base of C. Therefore, it follows by Lemma 3.4 (b) that

$$\mathrm{dcl}^{\mathrm{eq}}(e) = \mathrm{dcl}^{\mathrm{eq}}(B), \quad \text{for every Galois base } B \text{ of } C.$$

(5) ⇒ (4) Suppose that $C = \{\bar{c}_0, \dots, \bar{c}_n\}$ is a finite relation such that $\mathrm{dcl}^{\mathrm{eq}}(e) = \mathrm{dcl}^{\mathrm{eq}}(B)$, for every Galois base B of C. Since $\mathbb{M}^{\mathrm{eq}}$ has elimination of imaginaries, there exists a Galois base $B \subseteq \mathbb{M}^{\mathrm{eq}}$ of C. Consequently, Lemma 3.4 (a) implies that e is also a Galois base of C.

Let π be an automorphisms of $\mathbb{M}^{\mathrm{eq}}$. Then

$$\begin{aligned} \pi(\bar{c}_0 \dots \bar{c}_n) = \bar{c}_0 \dots \bar{c}_n \quad &\text{implies} \quad \pi[C] = C \\ &\text{implies} \quad \pi(e) = e\,. \end{aligned}$$

By Corollary 1.8, it follows that $e \in \mathrm{dcl}^{\mathrm{eq}}(\bar{c}_0 \dots \bar{c}_n)$. Similarly,

$$\begin{aligned} \pi(e) = e \quad &\text{implies} \quad \pi[C] = C \\ &\text{implies} \quad \pi(\bar{c}_0 \dots \bar{c}_n) = \bar{c}_{\sigma(0)} \dots \bar{c}_{\sigma(n)}\,, \\ &\qquad\qquad \text{for some permutation } \sigma. \end{aligned}$$

Therefore, there are only finitely many conjugates of $\bar{c}_0 \dots \bar{c}_n$ over e. According to Theorem 1.6 this implies that $\bar{c}_0 \dots \bar{c}_n \subseteq \mathrm{acl}^{\mathrm{eq}}(e)$. □

In later chapters we will present several conditions implying that a theory has weak elimination of imaginaries. Here, we give only one example.

Lemma 5.6. *A theory T satisfying the following two conditions has weak elimination of imaginaries:*

- *There is no strictly decreasing sequence $A_0 \supset A_1 \supset \dots$ of sets of the form $A_i = \mathrm{acl}(B_i)$ where each B_i is finite.*
- *If A and B are algebraic closures of finite sets, then $\mathfrak{Aut}\,\mathbb{M}_{A \cap B}$ is generated by $\mathrm{Aut}\,\mathbb{M}_A \cup \mathrm{Aut}\,\mathbb{M}_B$.*

Proof. By Proposition 5.5 it is sufficient to show that, for every parameter-definable relation $\mathbb{R}$, there is a least algebraically closed set over which $\mathbb{R}$ is definable.

Hence, let $\mathbb{R}$ be parameter-definable. First, let us show that, if $\mathbb{R}$ is definable over two algebraically closed sets A and B of the form $A = \mathrm{acl}(A_0)$

and $B = \mathrm{acl}(B_0)$, for finite A_0 and B_0, then it is also definable over their intersection $A \cap B$. If $\mathbb{R}$ is definable over both A and B, Lemma 1.10 implies that

$$\mathrm{Aut}\,\mathbb{M}_A \cup \mathrm{Aut}\,\mathbb{M}_B \subseteq \mathrm{Aut}\langle \mathbb{M}, \mathbb{R}\rangle\,.$$

Consequently, the second condition implies that

$$\mathrm{Aut}\,\mathbb{M}_{A\cap B} = \langle\!\langle \mathrm{Aut}\,\mathbb{M}_A \cup \mathrm{Aut}\,\mathbb{M}_B \rangle\!\rangle \subseteq \mathrm{Aut}\langle \mathbb{M}, \mathbb{R}\rangle\,.$$

Hence, it follows by Lemma 1.10 that $\mathbb{R}$ is definable over $A \cap B$.

By the first condition, it therefore follows that there is a least algebraically closed set over which $\mathbb{R}$ is definable. □

The following property is what is missing from weak elimination of imaginaries in order to obtain full elimination of imaginaries.

Definition 5.7. A complete first-order theory T has *elimination of finite imaginaries* if every finite relation has a finite Galois base in $\mathbb{M}$.

As an example, we consider the theory of algebraically closed fields. We will show later in Corollary **??** that this theory actually has uniform elimination of imaginaries.

Lemma 5.8. *The theory of algebraically closed fields of characteristic p has elimination of finite imaginaries.*

Proof. Let $R = \{\bar c^0, \dots, \bar c^{n-1}\}$ be a finite relation consisting of m-tuples $\bar c^i = \langle c^i_0, \dots, c^i_{m-1}\rangle$. We define the polynomial

$$p(x, y_0, \dots, y_{m-1}) := \prod_{i<n} (x - c^i_0 y_0 - \dots - c^i_{m-1} y_{m-1})\,.$$

Let Z be the set of roots of p. Then

$$\pi[Z] = Z \quad \text{iff} \quad \pi[R] = R\,, \quad \text{for every automorphism } \pi\,.$$

Since p is the only polynomial with set of roots Z, it follows that an automorphism fixes p if, and only if, it permutes R. Consequently, the coefficients of p form a Galois base of R. □

Proposition 5.9. *A theory T has elimination of imaginaries if, and only if, it has both, elimination of finite imaginaries and weak elimination of imaginaries.*

Proof. ($\Rightarrow$) Since every canonical parameter is a weak canonical parameter, elimination of imaginaries implies weak elimination of implies. Moreover, it follows by Proposition 4.7 (3) that every theory with elimination of imaginaries has elimination of finite imaginaries.

($\Leftarrow$) Let $e \in \mathbb{M}^{eq}$. By Proposition 5.5, there exists a finite set $C \subseteq \mathbb{M}^{\bar{s}}$ such that

$$\mathrm{dcl}^{eq}(e) = \mathrm{dcl}^{eq}(B)\,, \quad \text{for every Galois base } B \text{ of } C\,.$$

As T has elimination of finite imaginaries, the set C has a finite Galois base $B_0 \subseteq \mathbb{M}$. Hence,

$$\mathrm{dcl}^{eq}(e) = \mathrm{dcl}^{eq}(B_0)\,.$$

By Proposition 4.7, it follows that T has elimination of imaginaries. □

E3. *Prime models*

1. *Isolated types*

The usual way to construct structures in model theory consists in writing down an appropriate theory and proving that it is consistent. In particular, we can reconstruct from the elementary diagram of a structure the structure itself, or we can use it to obtain an elementary extension. If we want to construct rich models realising many types then, as we have seen in Chapter E1, this approach works well.

In the present chapter, on the other hand, we are interested in models realising few types. We start by studying those types that are unavoidable in the sense that they are realised in every model.

Definition 1.1. Let T be a theory.

(a) A formula φ *isolates* a type $\mathfrak{p}$ (w.r.t. T) if $\varphi \vDash \mathfrak{p}$ modulo T. We call a type $\mathfrak{p}$ over U *isolated* if it is isolated by a formula $\varphi(\bar{x}, \bar{c})$ with parameters $\bar{c} \subseteq U$. In particular, a complete type $\mathfrak{p} \in S^{\bar{s}}(U)$ is isolated if and only if $\langle\varphi\rangle = \{\mathfrak{p}\}$, i.e., $\mathfrak{p}$ is an isolated point in the topology of $S^{\bar{s}}(U)$.

(b) A structure $\mathfrak{A}$ is *atomic* if every realised type $\mathfrak{p} \in S^{<\omega}(\varnothing)$ is isolated. More generally, if $B, U \subseteq A$ then we call B *atomic over* U if only isolated types $\mathfrak{p} \in S^{<\omega}(U)$ are realised in B.

Lemma 1.2. *If* $\mathfrak{p}$ *is isolated by* $\varphi(\bar{x})$ *then* $\mathfrak{p}$ *is realised in every model of* $T \cup \{\exists\bar{x}\varphi\}$.

Lemma 1.3. *If* $\bar{a} \subseteq \mathrm{acl}(U)$ *then* $\mathrm{tp}(\bar{a}/U)$ *is isolated.*

Proof. Let $\mathfrak{M}$ be a model containing U. Since $\bar{a}$ is algebraic over U we can choose a formula $\varphi(\bar{x}, \bar{c})$ with parameters $\bar{c} \subseteq U$ such that $\mathfrak{M} \vDash \varphi(\bar{a}, \bar{c})$

and the set $\varphi(\bar{x}, \bar{c})^{\mathfrak{M}}$ is finite and of minimal size. We claim that this formula isolates $\operatorname{tp}(\bar{a}/U)$.

For a contradiction suppose that there is some formula $\psi(\bar{x}, \bar{d}) \in \operatorname{tp}(\bar{a}/U)$ such that $\varphi \not\models \psi$. Then we can find a tuple $\bar{b} \subseteq M$ with

$$\mathfrak{M} \models \varphi(\bar{b}, \bar{c}) \wedge \neg\psi(\bar{b}, \bar{d}) .$$

It follows that

$$[\varphi(\bar{x}, \bar{c}) \wedge \psi(\bar{x}, \bar{d})]^{\mathfrak{M}} \subseteq \varphi(\bar{x}, \bar{c})^{\mathfrak{M}} \setminus \{\bar{b}\} \subset \varphi(\bar{x}, \bar{c})^{\mathfrak{M}},$$

in contradiction to our choice of φ. □

Lemma 1.4. *Every isolated type $\mathfrak{p} \in S^{\bar{s}}(U)$ is definable over a finite subset $U_0 \subseteq U$.*

Proof. Let $\varphi(\bar{x}, \bar{c})$ be a formula over U isolating $\mathfrak{p}$. We claim that $\mathfrak{p}$ is definable over $U_0 := \bar{c}$. Let $\psi(\bar{x}, \bar{y})$ be a formula and $\bar{b} \subseteq U$. Then we have

$$\begin{aligned} \psi(\bar{x}, \bar{b}) \in \mathfrak{p} \quad &\text{iff} \quad T(U) \cup \{\varphi(\bar{x}, \bar{c})\} \models \psi(\bar{x}, \bar{b}) \\ &\text{iff} \quad T(U) \models \forall\bar{x}[\varphi(\bar{x}, \bar{c}) \rightarrow \psi(\bar{x}, \bar{b})] . \end{aligned}$$

Consequently, $\delta_\psi(\bar{y}) := \forall\bar{x}[\varphi(\bar{x}, \bar{c}) \rightarrow \psi(\bar{x}, \bar{y})]$ is a ψ-definition of $\mathfrak{p}$ over U_0. □

Lemma 1.5. $\operatorname{tp}(\bar{a}\bar{b}/U)$ *is isolated if and only if the types* $\operatorname{tp}(\bar{a}/U)$ *and* $\operatorname{tp}(\bar{b}/U \cup \bar{a})$ *are isolated.*

Proof. ($\Leftarrow$) If $\varphi(\bar{x})$ isolates $\operatorname{tp}(\bar{a}/U)$ and $\psi(\bar{y}, \bar{a})$ isolates $\operatorname{tp}(\bar{b}/U \cup \bar{a})$ then the formula $\varphi(\bar{x}) \wedge \psi(\bar{y}, \bar{x})$ isolates $\operatorname{tp}(\bar{a}\bar{b}/U)$.

($\Rightarrow$) Let $\varphi(\bar{x}, \bar{y})$ be a formula isolating $\operatorname{tp}(\bar{a}\bar{b}/U)$. Then the formula $\varphi(\bar{a}, \bar{y})$ isolates $\operatorname{tp}(\bar{b}/U \cup \bar{a})$. Furthermore, we claim that $\exists\bar{y}'\varphi(\bar{x}, \bar{y})$ isolates $\operatorname{tp}(\bar{a}/U)$ where $\bar{y}' \subseteq \bar{y}$ is the finite tuple of those variables that actually appear in φ. Suppose that $\exists\bar{y}'\varphi \in \operatorname{tp}(\bar{c}/U)$. Then there is some tuple $\bar{d}$ with $\varphi \in \operatorname{tp}(\bar{c}\bar{d}/U)$. Consequently, $\operatorname{tp}(\bar{c}\bar{d}/U) = \operatorname{tp}(\bar{a}\bar{b}/U)$ and $\operatorname{tp}(\bar{c}/U) = \operatorname{tp}(\bar{a}/U)$. □

We conclude this section with a collection of basic facts about atomic models.

Lemma 1.6. *If A is atomic over U and $\bar{a} \in A^{<\omega}$ then A is atomic over $U \cup \bar{a}$.*

Proof. For every finite tuple $\bar{b} \in A^{<\omega}$ we know that $\operatorname{tp}(\bar{a}\bar{b}/U)$ is isolated. By Lemma 1.5 it follows that $\operatorname{tp}(\bar{b}/U \cup \bar{a})$ is also isolated. □

Lemma 1.7. *Let $A \subseteq B \subseteq C$. If C is atomic over B and B is atomic over A then C is atomic over A.*

Proof. Let $\bar{c} \subseteq C$ and suppose that $\operatorname{tp}(\bar{c}/B)$ is isolated by $\varphi(\bar{x}, \bar{b})$. Fix some formula $\psi(\bar{y}, \bar{a})$ isolating $\operatorname{tp}(\bar{b}/A)$. We claim that $\operatorname{tp}(\bar{c}/A)$ is isolated by the formula $\chi := \exists \bar{y}[\varphi(\bar{x}, \bar{y}) \wedge \psi(\bar{y}, \bar{a})]$.

Suppose that $\chi \in \operatorname{tp}(\bar{d}/A)$. Then there is some tuple $\bar{e}$ with

$$\varphi(\bar{d}, \bar{e}), \psi(\bar{e}, \bar{a}) \in \operatorname{tp}(\bar{d}\bar{e}/A)\,.$$

Consequently, we have $\operatorname{tp}(\bar{e}/A) = \operatorname{tp}(\bar{b}/A)$ and there exists an A-automorphism π with $\pi(\bar{e}) = \bar{b}$. Let $\bar{d}' := \pi(\bar{d})$. Then $\operatorname{tp}(\bar{d}'/\bar{b}) = \operatorname{tp}(\bar{d}/\bar{e})$ and $\varphi(\bar{x}, \bar{b}) \in \operatorname{tp}(\bar{d}'/\bar{b})$ implies that $\operatorname{tp}(\bar{d}'/B) = \operatorname{tp}(\bar{c}/B)$. It follows that

$$\operatorname{tp}(\bar{d}/A) = \operatorname{tp}(\bar{d}'/A) = \operatorname{tp}(\bar{c}/A)\,. \qquad \square$$

The following two remarks follow immediately from the definition of an atomic model.

Lemma 1.8. (a) *Every elementary substructure of an atomic model is atomic.*

(b) *The union of an elementary chain of atomic models is atomic.*

2. The Omitting Types Theorem

We have seen in Section C2.4 how to build structures from a given set of formulae. In order to find structures realising only certain types we take a closer look at this construction. First, let us determine a minimal set of sorts a model has to realise.

Lemma 2.1. *Let Σ be an S-sorted signature and $T \subseteq \mathrm{FO}^0[\Sigma]$ a first-order theory. There exists a minimal set $S_0 \subseteq S$ such that T has a model $\mathfrak{A}$ with*

$$A_s \neq \varnothing \quad \text{iff} \quad s \in S_0 \,.$$

Proof. Let $\mathcal{S}$ be the class of all sets $S_0 \subseteq S$ such that T has a model $\mathfrak{A}$ with $A = \bigcup_{s \in S_0} A_s$. It is sufficient to show that the partial order $\langle \mathcal{S}, \supseteq \rangle$ is inductively ordered. Let $(S_i)_{i \in I}$ be a decreasing sequence of sets $S_i \in \mathcal{S}$ and set $S_\infty := \bigcap_i S_i$. We claim that $S_\infty \in \mathcal{S}$. Let

$$\Phi := T \cup \{ \eta_s \mid s \in S \setminus S_\infty \} \,,$$

where $\eta_s := \neg \exists x_s (x_s = x_s)$ states that there are no elements of sort s. Every model of Φ witnesses that $S_\infty \in \mathcal{S}$.

To prove that Φ is satisfiable let $\Phi_0 \subseteq \Phi$ be finite. Then there are sorts $s_0, \dots, s_n \in S_\infty$ such that

$$\Phi_0 \subset T \cup \{\eta_{s_0}, \dots, \eta_{s_n}\} \,.$$

Hence, we can find some index $i \in I$ with $s_0, \dots, s_n \in S \setminus S_i$. By assumption there is some S_i-sorted model $\mathfrak{A}$ of T. It follows that $\mathfrak{A} \vDash \Phi_0$. □

We have seen in Section C2.4 how to construct Herbrand models from Hintikka sets. To refine this construction we introduce a special kind of Hintikka set called a *Henkin set.*

Definition 2.2. Let $\Phi \subseteq \mathrm{FO}^0[\Sigma]$ be a set of sentences and $C \subseteq \Sigma$ a set of constant symbols.

(a) Φ has the *Henkin property* with respect to C if, for every formula $\varphi(x) \in \mathrm{FO}^1[\Sigma]$, there is some constant $c \in C$ such that

$$\exists x \varphi(x) \to \varphi(c) \in \Phi \,.$$

(b) We say that Φ is a *Henkin set* for a set $\Phi_0 \subseteq \mathrm{FO}^0[\Sigma]$ with respect to C if $\Phi_0 \subseteq \Phi$, Φ is complete, and Φ has the Henkin property with respect to C.

Lemma 2.3. *Every Henkin set is a Hintikka set.*

Corollary 2.4. *Every Henkin set Φ with respect to C has a Herbrand model $\mathfrak{H}$ where every element is denoted by some constant from C.*

Proof. We have seen in Lemma c2.4.6 that Φ has a Herbrand model $\mathfrak{H}$ where every element is denoted by some term. Since Φ is a Hintikka set, we can find, for every term t a constant $c \in C$ with

$$\exists x(x = t) \to c = t \in \Phi .$$

Therefore, every element is denoted by some constant in C. □

The class of all Henkin sets is in one-to-one correspondence with the class of all Herbrand models. In the next lemma we prove that this class forms a co-meagre set in the type topology.

Lemma 2.5. *Suppose that Σ is a countable signature, $T \subseteq \mathrm{FO}^\circ[\Sigma]$ a theory, and, for every sort s, let C_s be a countably infinite set of constant symbols of sort s with $C_s \cap \Sigma = \varnothing$. Set $C := \bigcup_s C_s$ and*

$$S^\circ_C(T) := S(\mathrm{FO}^\circ[\Sigma_C]/T) .$$

(a) *The complement of the set*

$$H(T) := \{\, \mathfrak{p} \in S^\circ_C(T) \mid \mathfrak{p} \text{ is a Henkin set for } T \text{ w.r.t. } C \,\}$$

is meagre in $S^\circ_C(T)$.

(b) *If $\bar{s}$ is a finite tuple of sorts and $\Phi \subseteq \mathrm{FO}^{\bar{s}}[\Sigma]$ is a set such that $\langle \Phi \rangle_{S^{\bar{s}}(T)}$ is nowhere dense then the complement of*

$$O(\Phi) := \{\, \mathfrak{p} \in S^\circ_C(T) \mid \text{for every } \bar{c} \in C^{<\omega},\ \text{there is some } \varphi \in \Phi \text{ with } \neg\varphi(\bar{c}) \in \mathfrak{p} \,\}$$

is meagre in $S^\circ_C(T)$.

Proof. (a) We have

$$H(T) = \bigcap_{\varphi \in \mathrm{FO}^1[\Sigma_C]} H_\varphi \quad \text{where} \quad H_\varphi = \bigcup_{c \in C} \langle \exists x \varphi(x) \to \varphi(c) \rangle_{S^\circ_C(T)} .$$

Since $\mathrm{FO}^1[\Sigma_C]$ is countable, we can show that the complement of $H(T)$ is meagre by proving that the complement of each H_φ is nowhere dense. Because H_φ is open, it is sufficient to show that its complement has empty interior, that is, that H_φ is dense.

Let $\langle \psi \rangle_{S^\circ_C(T)}$ be a nonempty basic open set and fix some model

$$\mathfrak{M} \vDash T \cup \{\psi\} .$$

Choose some element $a \in M$ with

$$\mathfrak{M} \vDash \exists x \varphi(x) \to \varphi(a) .$$

Let $D \subseteq C$ be the set of constant symbols appearing in ψ or φ. This set is finite and we have

$$\mathfrak{M}|_{\Sigma_D} \vDash T \cup \{\psi, \exists x \varphi(x) \to \varphi(a)\} .$$

Fix some constant symbol $c \in C \smallsetminus D$ of the same sort as a and let $\mathfrak{N}$ be a Σ_C-expansion of $\mathfrak{M}|_{\Sigma_D}$ with $c^{\mathfrak{N}} = a$. Then

$$\mathfrak{N} \vDash T \cup \{\psi, \exists x \varphi(x) \to \varphi(c)\} .$$

Hence, $\mathrm{Th}(\mathfrak{N}) \in \langle \psi \rangle_{S^\circ_C(T)} \cap H_\varphi \neq \emptyset$.

(b) We have

$$O(\Phi) = \bigcap_{\bar{c} \in C^{<\omega}} O_{\bar{c}} \quad \text{where} \quad O_{\bar{c}} = \bigcup_{\varphi \in \Phi} \langle \neg\varphi(\bar{c}) \rangle_{S^\circ_C(T)} .$$

As above it is sufficient to prove that each set $O_{\bar{c}}$ is dense. Consider a nonempty basic open set $\langle \psi(\bar{c}, \bar{d}) \rangle_{S^\circ_C(T)}$ where $\psi \in \mathrm{FO}[\Sigma]$ and $\bar{d} \subseteq C \smallsetminus \bar{c}$. Fix some model $\mathfrak{M} \vDash T \cup \{\psi(\bar{c}, \bar{d})\}$. Then

$$\langle \mathfrak{M}|_\Sigma, \bar{c} \rangle \vDash T \cup \{\exists \bar{y} \psi(\bar{x}, \bar{y})\} .$$

Hence, $\langle \exists \bar{y} \psi(\bar{x}, \bar{y}) \rangle_{S^{\bar{s}}(T)} \neq \varnothing$. Since $\langle \Phi \rangle_{S^{\bar{s}}(T)}$ is nowhere dense it follows that

$$\langle \exists \bar{y} \psi(\bar{x}, \bar{y}) \rangle_{S^{\bar{s}}(T)} \smallsetminus \langle \Phi \rangle_{S^{\bar{s}}(T)} \neq \varnothing .$$

Fix some model $\langle \mathfrak{N}_0, \bar{a} \rangle$ with

$$\mathrm{Th}(\mathfrak{N}_0, \bar{a}) \in \langle \exists \bar{y} \psi(\bar{x}, \bar{y}) \rangle_{S^{\bar{s}}(T)} \smallsetminus \langle \Phi \rangle_{S^{\bar{s}}(T)} .$$

There is some formula $\varphi \in \Phi$ such that

$$\mathfrak{N}_0 \not\models \varphi(\bar{a}) .$$

Furthermore, we can find a tuple $\bar{b} \subseteq N_0$ with

$$\mathfrak{N}_0 \models \psi(\bar{a}, \bar{b}) .$$

Let $\mathfrak{N}$ be a Σ_C-expansion of $\mathfrak{N}_0$ with $\bar{c}^{\mathfrak{N}} = \bar{a}$ and $\bar{d}^{\mathfrak{N}} = \bar{b}$. Then we have

$$\mathrm{Th}(\mathfrak{N}) \in \langle \psi \rangle_{S^0_C(T)} \cap O_{\bar{c}} \neq \varnothing . \qquad \square$$

After these preparations we can prove that every meagre set of types is omitted in some model.

Theorem 2.6 (Omitting Types Theorem). *Let Σ be a countable S-sorted signature and $T \subseteq \mathrm{FO}[\Sigma]$ a countable first-order theory. For every $\bar{s} \in S^{<\omega}$, let $X_{\bar{s}} \subseteq S^{\bar{s}}(T)$ be a meagre set of types. There exists a model of T that omits every type in $\bigcup_{\bar{s}} X_{\bar{s}}$.*

Proof. For every sort s, fix a countably infinite set C_s of constant symbols disjoint from Σ. Each set $X_{\bar{s}}$ can be written as $X_{\bar{s}} = \bigcup_{n<\omega} X^n_{\bar{s}}$, where $X^n_{\bar{s}}$ is nowhere dense. Let $\Phi^n_{\bar{s}}$ be a set of formulae such that $\langle \Phi^n_{\bar{s}} \rangle = \mathrm{cl}(X^n_{\bar{s}})$. By the preceding lemma, we know that

$$Y := H(T) \cap \bigcap_{\bar{s} \in S^{<\omega}} \bigcap_{n<\omega} O(\Phi^n_{\bar{s}})$$

is a countable intersection of sets whose complement is meagre. Hence, the complement of Y is meagre. By Theorem B5.5.8 it follows that Y itself is also dense. Fix some type $\mathfrak{p} \in Y$.

By Corollary 2.4, there exists a Herbrand model $\mathfrak{H}$ of $\mathfrak{p}$ where every element is denoted by some constant in C. If $\bar{a} \in H^{\bar{s}}$ is a finite tuple denoted by the constants $\bar{c} \subseteq C$ then we have

$$\mathrm{tp}(\bar{a}) = \{\, \varphi(\bar{x}) \mid \varphi(\bar{c}) \in \mathfrak{p} \,\} \notin X_{\bar{s}}\,.$$

Hence, no tuple in $\mathfrak{H}$ realises a type in $X_{\bar{s}}$. □

Corollary 2.7. *Let Σ be a countable signature and $T \subseteq \mathrm{FO}[\Sigma]$ a first-order theory. Let $\mathfrak{p}_n$, $n < \omega$, be a sequence of non-isolated partial types over T. There exists a model of T that omits every $\mathfrak{p}_n$, $n < \omega$.*

Let us give a simple example showing that the Omitting Types Theorem fails for uncountable theories.

Example. Let $\Sigma := \{\, c_i \mid i < \omega_1 \,\} \cup \{\, d_n \mid n < \omega \,\}$ be a signature of constant symbols and let

$$T := \{\, c_i \neq c_k \mid i \neq k \,\} \cup \{\, d_i \neq d_k \mid i \neq k \,\}$$

be the theory stating that the values of the c_i are distinct and that the values of the d_n are distinct. Consider the partial 1-type

$$\Phi := \{\, x \neq d_n \mid n < \omega \,\}\,.$$

This type is not isolated since there is no formula $\varphi(x)$ implying that x is different from all constants d_n. On the other hand, every model of T has uncountably many elements and, therefore, realises Φ.

Theorem 2.8. *Let T be a countable complete theory with infinite models. There exists a family $(\mathfrak{M}_\xi)_{\xi<2^{\aleph_0}}$ of models of T such that every complete type that is realised in at least two of the models is isolated.*

Proof. For every sort s, fix a countably infinite set C_s of constant symbols disjoint from Σ. Set $C := \bigcup_s C_s$ and let $(\varphi_n)_n$ be an enumeration of $\mathrm{FO}^1[\Sigma_C]$. We fix an enumeration $\langle u_n, \bar{c}^n, \bar{d}^n \rangle_{n<\omega}$ of all triples in $2^{<\omega} \times C^{<\omega} \times C^{<\omega}$ such that $\bar{c}^n$ and $\bar{d}^n$ have the same length and the same sorts. We assume that the enumeration has been chosen such that every triple appears infinitely often in the sequence.

We construct an increasing chain $T_0 \subseteq T_1 \subseteq \dots$ of finite trees $T_n \subseteq 2^{<\omega}$ and, for each $w \in 2^{<\omega}$, we define a finite set $\Phi_w \subseteq \mathrm{FO}^0[\Sigma_C]$ of formulae such that $\Phi_u \subseteq \Phi_w$, for $u \leq w$.

We start with $T_0 := \{\langle\rangle\}$ and $\Phi_{\langle\rangle} := \varnothing$. For the inductive step, suppose that we have already defined T_n and Φ_w, for $w \in T_n$. To define T_{n+1} we distinguish two cases. If $u_n \notin T_n$ then we simply set

$$T_{n+1} := \{ w0 \mid w \text{ a leaf of } T_n \},$$

and, for every leaf w of T_n,

$$\Phi_{w0} := \Phi_w \cup \{ \exists x \varphi_n \rightarrow \varphi_n(c) \},$$

where $c \in C$ is some new constant symbol not appearing in any formula of Φ_w.

It remains to consider the case that $u_n \in T_n$. Let $v_0, \dots, v_{l-1}$ be an enumeration of all leaves v of T_n with $u_n \leq v$, and let $w_0, \dots, w_{m-1}$ be an enumeration of all leaves w with $u_n \nleq w$. We define sets

$$\begin{aligned} \Phi_{w_i} &= \Psi^i_{-1} \subseteq \Psi^i_0 \subseteq \dots \subseteq \Psi^i_{l-1}, \quad &\text{for } i < m, \\ \Phi_{v_j} &= \Theta^j_{-1} \subseteq \Theta^j_0 \subseteq \dots \subseteq \Theta^j_{m-1}, \quad &\text{for } j < l, \end{aligned}$$

as follows. We start with $\Psi^i_{-1} := \Phi_{w_i}$ and $\Theta^j_{-1} := \Phi_{v_j}$. Suppose that we have already defined Ψ^i_j and Θ^j_i, for all pairs $\langle i, j \rangle$ lexicographically less than $\langle i_0, j_0 \rangle$. To define $\Psi^{i_0}_{j_0}$ and $\Theta^{j_0}_{i_0}$ we set

$$\psi(\bar{c}^n, \bar{e}) := \bigwedge \Psi^{i_0}_{j_0 - 1} \quad \text{and} \quad \vartheta(\bar{d}^n, \bar{f}) := \bigwedge \Theta^{j_0}_{i_0 - 1},$$

where $\bar{e} \subseteq C$ contains all constants in $\Psi^{i_0}_{j_0-1}$ different from $\bar{c}^n$, and $\bar{f} \subseteq C$ contains all constants in $\Theta^{j_0}_{i_0-1}$ different from $\bar{d}^n$. If $\langle \exists \bar{y} \psi(\bar{x}, \bar{y}) \rangle_{S(T)}$ is a singleton then we set

$$\Psi^{i_0}_{j_0} := \Psi^{i_0}_{j_0-1} \quad \text{and} \quad \Theta^{j_0}_{i_0} := \Theta^{j_0}_{i_0-1} .$$

Otherwise, we choose some type $\mathfrak{q} \in \langle \exists \bar{y} \vartheta(\bar{x}, \bar{y}) \rangle_{S(T)}$. By assumption, we can find a type $\mathfrak{p} \in \langle \exists \bar{y} \psi(\bar{x}, \bar{y}) \rangle_{S(T)}$ different from $\mathfrak{q}$. We fix some formula $\eta(\bar{x}) \in \mathfrak{p} \setminus \mathfrak{q}$ and set

$$\Psi^{i_0}_{j_0} := \Psi^{i_0}_{j_0-1} \cup \{\eta(\bar{c}^n)\} \quad \text{and} \quad \Theta^{j_0}_{i_0} := \Theta^{j_0}_{i_0-1} \cup \{\neg\eta(\bar{d}^n)\} .$$

Having defined all Ψ^i_j and Θ^j_i we set

$$\Phi'_{w_i} := \Psi^i_{l-1} \cup \{\exists x \varphi_n \to \varphi_n(c)\} ,$$
$$\Phi'_{v_j} := \Theta^j_{m-1} \cup \{\exists x \varphi_n \to \varphi_n(c)\} ,$$

where $c \in C$ is some constant not appearing in any set Ψ^i_j or Θ^j_i. Let $z_0, \dots, z_{k-1}$ be an enumeration of all leaves z of T_n such that the set $\langle \Phi'_z \rangle_{S(T(C))}$ contains at least two types, and let $u_0, \dots, u_{r-1}$ be an enumeration of all other leaves of T_n. We define

$$T_{n+1} := T_n \cup \{ z_i b \mid i < k,\ b \in [2] \} \cup \{ u_i 0 \mid i < r \} ,$$

and $\Phi_{u_i 0} := \Phi'_{u_i}$, for $i < r$. For each $i < k$, we chose distinct types $\mathfrak{p}_i, \mathfrak{q}_i \in \langle \Phi'_{z_i} \rangle_{S(T(C))}$ and some formula $\eta_i \in \mathfrak{p}_i \setminus \mathfrak{q}_i$. Then we set

$$\Phi_{z_i 0} := \Phi'_{z_i} \cup \{\neg\eta_i\} \quad \text{and} \quad \Phi_{z_i 1} := \Phi'_{z_i} \cup \{\eta_i\} .$$

This completes the construction of T_{n+1}. To define the models $\mathfrak{M}_\xi$ let $T_\omega := \bigcup_n T_n$. A sequence $\beta \in 2^\omega$ is a *branch* of T_ω if $\beta \restriction n \in T_\omega$, for all $n < \omega$. For each branch β of T_ω, we define a sequence $\beta^* \in 2^{<\omega}$ as follows. Let

$$I := \{ n < \omega \mid (\beta \restriction n)0 \in T_\omega \text{ and } (\beta \restriction n)1 \in T_\omega \} ,$$

and let $n_0 < n_1 < \dots$ be an enumeration of I. We define $\beta^* \in 2^{|I|}$ by

$$\beta^*(i) := \beta(n_i), \quad \text{for } i < |I|.$$

For each $\xi \in 2^\omega$, there is a unique branch β_ξ with $\beta^*_\xi \leq \xi$. We define

$$\Psi_\xi := \bigcup_{n<\omega} \Phi_{\beta_\xi \upharpoonright n}.$$

It follows by compactness that each set Ψ_ξ is satisfiable. Furthermore, the above construction ensures that each of these sets has the Henkin property with respect to C. Hence, we can use Corollary 2.4 to find a Herbrand model $\mathfrak{M}_\xi$ of Ψ_ξ.

It remains to prove that every type realised in two different models is isolated. Suppose that

$$\operatorname{tp}(\bar{c}/\mathfrak{M}_\xi) = \operatorname{tp}(\bar{d}/\mathfrak{M}_\zeta) \quad \text{where} \quad \xi \neq \zeta.$$

If β^*_ξ is finite then $\langle \Phi_{\beta^*_\xi} \rangle_{S(T(C))} = \{\mathfrak{p}\}$ is a singleton and every type realised in $\mathfrak{M}_\xi \vDash \Phi_{\beta^*_\xi}$ is isolated. Similarly, if β^*_ζ is finite then $\operatorname{tp}(\bar{d}/\mathfrak{M}_\zeta)$ is isolated.

Hence, suppose that β^*_ξ and β^*_ζ are both infinite. Then there is some $n < \omega$ such that

$$\bar{c}^n = \bar{c}, \quad \bar{d}^n = \bar{d}, \quad u_n \in T_n, \quad \text{and} \quad \beta_\xi \sqcap \beta_\zeta \prec u_n \prec \beta_\zeta.$$

Let w be the leaf of T_n with $w \prec \beta_\xi$ and let v be the leaf with $v \prec \beta_\zeta$. By construction of T_{n+1} it follows that either there is a formula isolating $\operatorname{tp}(\bar{c}/\mathfrak{M}_\xi)$, or there is some formula $\eta(\bar{c}) \in \Phi'_w \subseteq \Psi_\xi$ with $\neg\eta(\bar{d}) \in \Phi'_v \subseteq \Psi_\zeta$. In the first case we are done, whereas in the second case we obtain $\operatorname{tp}(\bar{c}/\mathfrak{M}_\xi) \neq \operatorname{tp}(\bar{d}/\mathfrak{M}_\zeta)$, a contradiction. □

3. *Prime and atomic models*

Not every theory has atomic models, but for countable signatures we can use the Omitting Types Theorem to construct such models.

Theorem 3.1. *Let T be a countable complete theory. If $S^{\bar{s}}(T)$ is countable, for all finite tuples $\bar{s}$, then there exists a countable atomic model of T.*

Proof. For every $\bar{s}$, there are at most countably many non-isolated $\bar{s}$-types. Consequently, they form a meagre set and we can use the Omitting Types Theorem to find a model of T that realises none of them. □

Lemma 3.2. *Let T be a countable complete theory. If $|S^{\bar{s}}(T)| < 2^{\aleph_0}$, for all finite $\bar{s}$, then T has an atomic model over A, for every finite set A of parameters.*

Proof. By Corollary B5.7.5, it follows that each type space $S^{\bar{s}}(T)$ is countable. Let $\bar{a}$ be an enumeration of A. Since $\operatorname{tp}(\bar{b}/\bar{a})$ is determined by $\operatorname{tp}(\bar{b}\bar{a})$ it follows that $S^{\bar{s}}(A)$ is also countable. Hence, according to the preceding theorem $T(A)$ has an atomic model. □

If the type space is too large, atomic models might not exist.

Example. Consider the theory $T := \operatorname{Th}(\mathfrak{C})$ where $\mathfrak{C} := \langle 2^\omega, (P_n)_{n<\omega} \rangle$ and

$$P_n := \{ \alpha \in 2^\omega \mid \alpha(n) = 1 \} .$$

As we have seen in the example on page 534, the type space $S^1(T)$ is homeomorphic to the Cantor discontinuum 2^ω, which does not contain isolated points. Consequently, no type is isolated and T does not have atomic models.

Theorem 3.3. *Let T be a countable complete first-order theory. There exists an atomic model of T if, and only if, the set of isolated $\bar{s}$-types is dense in $S^{\bar{s}}(T)$, for every finite $\bar{s}$.*

Proof. Let $X \subseteq S^{\bar{s}}(T)$ be the set of all isolated $\bar{s}$-types. If T has an atomic model $\mathfrak{M}$ then X is the set of types realised in $\mathfrak{M}$. By Lemma C3.2.6 it follows that X is dense. Conversely, if X is dense then its complement $Y_{\bar{s}} := S^{\bar{s}}(T) \setminus X$ is closed and has empty interior. By the Omitting Types Theorem, there exists a model $\mathfrak{M}$ of T omitting all types in $\bigcup_{\bar{s}} Y_{\bar{s}}$. This model is atomic. □

Corollary 3.4. *Let T be a countable complete theory. If*

$$\mathrm{rk}_{\mathrm{CB}}(S^n(T)) < \infty\,, \quad \textit{for all } n < \omega\,,$$

then T has an atomic model.

Proof. Immediately by Theorem 3.3 and Proposition B5.5.12. □

Intuitively, an atomic model is the opposite of a saturated one. The next lemma shows that these models also behave in the opposite way with respect to the relation $\sqsubseteq^{\aleph_0}_{\mathrm{FO}}$.

Lemma 3.5. (a) *If $\mathfrak{A}$ is atomic then we have $\mathfrak{A} \sqsubseteq^{\aleph_0}_{\mathrm{FO}} \mathfrak{B}$, for all $\mathfrak{B} \equiv \mathfrak{A}$.*

(b) *If $\mathfrak{A}$ is a structure with countable signature such that $\mathfrak{A} \sqsubseteq^{\aleph_0}_{\mathrm{FO}} \mathfrak{B}$, for all $\mathfrak{B} \equiv \mathfrak{A}$, then $\mathfrak{A}$ is atomic.*

Proof. (a) Suppose that

$$\langle \mathfrak{A}, \bar a \rangle \equiv_{\mathrm{FO}} \langle \mathfrak{B}, \bar b \rangle\,.$$

We have to prove the forth property. Let $c \in A$ and choose some formula $\varphi(\bar x, y)$ isolating $\mathfrak{p} := \mathrm{tp}(\bar a c/\mathfrak{A})$. Then

$$\mathfrak{A} \vDash \exists y \varphi(\bar a, y) \quad \text{implies} \quad \mathfrak{B} \vDash \exists y \varphi(\bar b, y)\,.$$

Consequently, there exists some $d \in B$ such that $\mathfrak{B} \vDash \varphi(\bar b, d)$. It follows that $\mathrm{tp}(\bar b d/\mathfrak{B}) = \mathfrak{p}$ and, hence,

$$\langle \mathfrak{A}, \bar a c \rangle \equiv_{\mathrm{FO}} \langle \mathfrak{B}, \bar b d \rangle\,.$$

(b) Suppose that $\mathfrak{A}$ contains a finite tuple $\bar a \subseteq A$ whose type $\mathrm{tp}(\bar a)$ is not isolated. By the Omitting Types Theorem there is a structure $\mathfrak{B} \equiv \mathfrak{A}$ omitting $\mathrm{tp}(\bar a)$. If $\mathfrak{A} \sqsubseteq^{\aleph_0}_{\mathrm{FO}} \mathfrak{B}$ then there would be some tuple $\bar b \subseteq B$ such that $\langle \mathfrak{A}, \bar a \rangle \equiv \langle \mathfrak{B}, \bar b \rangle$. Consequently, $\mathrm{tp}(\bar b/\mathfrak{B}) = \mathrm{tp}(\bar a/\mathfrak{A})$ would be realised in $\mathfrak{B}$. Contradiction. □

Corollary 3.6. *If $\mathfrak{A} \equiv \mathfrak{B}$ are atomic then $\mathfrak{A} \equiv^{\aleph_0}_{\mathrm{FO}} \mathfrak{B}$.*

Corollary 3.7. *Every atomic model is $\aleph_0$-homogeneous.*

Proof. By the preceding corollary we have $\mathfrak{A} \equiv^{\aleph_0}_{\mathrm{FO}} \mathfrak{A}$, for every atomic structure $\mathfrak{A}$. □

If a countable theory T has atomic models then it has a unique countable one. Furthermore, this countable atomic model can be embedded into every other model of T.

Definition 3.8. A structure $\mathfrak{A}$ is a *prime model* of a theory T if, for ever model $\mathfrak{B} \vDash T$, there exists an elementary embedding $\mathfrak{A} \to \mathfrak{B}$. Similarly, we say that $\mathfrak{A}$ is *prime over* a set $U \subseteq A$ if it is a prime model of $T(U)$.

Example. $\mathfrak{N} = \langle \mathbb{N}, +, \cdot, 0, 1 \rangle$ is a prime model of arithmetic.

Remark. Only complete theories can have prime models.

Lemma 3.9. *If $\mathfrak{M}$ is a structure with $M = \mathrm{acl}(\varnothing)$ then $\mathfrak{M}$ is prime.*

Exercise 3.1. Prove the preceding lemma.

Lemma 3.10. *Every prime model with a countable signature is atomic.*

Proof. Let $\mathfrak{M}$ be a model of a theory T that realises a non-isolated type $\mathfrak{p}$. By the Omitting Types Theorem, there exists some model $\mathfrak{N} \vDash T$ in which $\mathfrak{p}$ is not realised. Therefore, there exists no embedding $\mathfrak{M} \to \mathfrak{N}$ and $\mathfrak{M}$ cannot be prime. □

Lemma 3.11. *Every countable atomic model is prime.*

Proof. Let $\mathfrak{A}$ be a countable atomic model and suppose that $\mathfrak{B} \equiv \mathfrak{A}$. Let $(a_i)_{i<\omega}$ be an enumeration of A. Since $\mathfrak{A} \sqsubseteq^{\aleph_0}_{\mathrm{FO}} \mathfrak{B}$ we can find, by Lemma C4.4.9, an enumeration $(b_i)_{i<\omega}$ such that

$$\langle \mathfrak{A}, (a_i)_{i<n} \rangle \equiv_{\mathrm{FO}} \langle \mathfrak{B}, (b_i)_{i<n} \rangle, \qquad \text{for all } n < \omega .$$

Let $p_n : (a_i)_{i<n} \mapsto (b_i)_{i<n} \in I_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B})$ be the corresponding partial isomorphisms. Since $I_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B})$ is $\aleph_1$-complete we have $p := \bigcup_n p_n \in I_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B})$. As $\mathrm{dom}\, p = A$ it follows that p is the desired elementary embedding of $\mathfrak{A}$ into $\mathfrak{B}$. □

The next theorem summarises the relation between prime and atomic models.

Theorem 3.12. *Let T be a countable complete theory.*

(a) *Every prime model of T is countable and atomic.*

(b) *Every countable atomic model of T is prime.*

(c) *T has a prime model if and only if it has an atomic model.*

(d) *All prime models of T are isomorphic.*

Proof. (a) and (b) were proved in Lemmas 3.10 and 3.11, respectively.

(c) By (a), every prime model is atomic. Conversely, if T has an atomic model then it also has a countable one, by the theorem of Löwenheim and Skolem. Hence, the claim follows by (b).

(d) If $\mathfrak{A}$ and $\mathfrak{B}$ are prime models of T then we have $\mathfrak{A} \cong_{\mathrm{FO}}^{\aleph_0} \mathfrak{B}$, by (a) and Corollary 3.6. Since $\mathfrak{A}$ and $\mathfrak{B}$ are countable, Lemma C4.4.10 implies that $\mathfrak{A} \cong \mathfrak{B}$. □

4. Constructible models

For uncountable signatures we cannot use the Omitting Types Theorem to construct prime models. In this section we present an alternative way to obtain such models.

Definition 4.1. Let $\mathfrak{M}$ be a structure and $A, U \subseteq M$.

(a) A *construction* of A over U is an enumeration $(a_i)_{i<\gamma}$ of A such that

$$\mathrm{tp}(a_\alpha / U \cup a[<\alpha]) \text{ is isolated}, \quad \text{for all } \alpha < \gamma,$$

where $a[<\alpha] := \{ a_i \mid i < \alpha \}$.

(b) If there exists a construction of A over U we say that A is *constructible* over U.

Example. Let T_{eq} be the theory of all infinite structures with empty signature. This theory has exactly one model of every infinite cardinality.

The countable model $\mathfrak{M}_{\aleph_0}$ of T_{eq} is constructible. If $(a_n)_{n<\omega}$ is an enumeration of $M_{\aleph_0}$ then $\mathrm{tp}(a_n/a_0 \dots a_{n-1})$ is isolated by the formula

$$x \neq a_0 \wedge \dots \wedge x \neq a_{n-1}\,.$$

Every uncountable model $\mathfrak{M}$ of T_{eq} is not constructible since, for every enumeration $(a_\alpha)_{\alpha<\gamma}$ of M, the type $\mathrm{tp}(a_\omega/a[<\omega])$ is not isolated.

We start by showing that constructible models are prime and atomic.

Lemma 4.2. *If $A \subseteq M$ is constructible over U then A is atomic over U.*

Proof. Let $(a_\alpha)_{\alpha<\gamma}$ be a construction of A over U. We prove by induction on α that $a[<\alpha]$ is atomic over U. For $\alpha = 0$ there is nothing to do. If α is a limit ordinal then any finite tuple in $a[<\alpha] = \bigcup_{\beta<\alpha} a[<\beta]$ belongs to some $a[<\beta]$ with $\beta < \alpha$. Hence, the claim follows immediately by inductive hypothesis.

For the inductive step, note that $a[<\alpha+1] = a[<\alpha] \cup \{a_\alpha\}$ is atomic over $U \cup a[<\alpha]$ and $U \cup a[<\alpha]$ is atomic over U. By Lemma 1.7, it follows that $a[<\alpha+1]$ is atomic over U. □

Proposition 4.3. *Let $\mathfrak{M}$ be a model of a complete theory T and let $U \subseteq M$ be a set such that M is constructible over U.*

(a) *$\mathfrak{M}$ is a prime model over U.*

(b) $|M| \leq |U| \oplus |T|$.

Proof. (a) Let $(a_\alpha)_{\alpha<\gamma}$ be a construction of M over U. Suppose that $\mathfrak{N}$ is a model of $T(U)$. We construct a sequence $(b_\alpha)_{\alpha<\gamma}$ as follows. Suppose that b_i has already been defined for all $i < \alpha$. Since the type $\mathrm{tp}(a_\alpha/U \cup a[<\alpha])$ is isolated, there exists some element $b_\alpha \in N$ with

$$b_\alpha b[<\alpha] \equiv_U a_\alpha a[<\alpha]\,.$$

The mapping $a_\alpha \mapsto b_\alpha$ is the desired elementary embedding $\mathfrak{M} \to \mathfrak{N}$.

(b) By the Theorem of Löwenheim and Skolem, $T(U)$ has a model $\mathfrak{N}$ of size $|N| \leq |U| \oplus |T|$. By (a), there exists an embedding $\mathfrak{M} \to \mathfrak{N}$. Consequently, $|M| \leq |N| \leq |U| \oplus |T|$. ☐

Our next aim is to prove that constructible models are unique, up to isomorphism.

Definition 4.4. Let $(a_\alpha)_{\alpha<\gamma}$ be a construction of A over U. A set $C \subseteq A$ is *closed* (w.r.t. this construction) if, for every $\alpha < \gamma$ with $a_\alpha \in C$, the type $\operatorname{tp}(a_\alpha/U \cup a[<\alpha])$ is isolated by some formula $\varphi(x;\bar{c})$ with parameters $\bar{c} \subseteq U \cup (C \cap a[<\alpha])$.

Lemma 4.5. *Let $(a_\alpha)_{\alpha<\gamma}$ be a construction of A over U.*

(a) *If $C, D \subseteq A$ are closed, then so is $C \cup D$.*

(b) *Every element $a \in A$ is contained in a finite closed set $C \subseteq A$.*

(c) *Every closed subset of A is constructible.*

Proof. (a) is immediate.

(b) By induction on $\alpha < \gamma$, we construct a finite closed set C_α containing a_α. For $\alpha = 0$, we can set $C_0 := \{a_0\}$ since $\operatorname{tp}(a_0/U)$ is isolated by some formula with parameters in U. For the inductive step, suppose that we have already defined C_i, for all $i < \alpha$. Fix a formula $\varphi(x;\bar{c})$ with parameters $\bar{c} \subseteq U \cup a[<\alpha]$ isolating $\operatorname{tp}(a_\alpha/U \cup a[<\alpha])$. Let $I := \{\, i < \alpha \mid a_i \in \bar{c} \,\}$. The set

$$C_\alpha := \{a_\alpha\} \cup \bigcup_{i \in I} C_i$$

is finite and closed.

(c) Let $(a_\alpha)_{\alpha<\gamma}$ be a construction of A over U, $C \subseteq A$ a closed set, and set $C_{<\alpha} := C \cap a[<\alpha]$. For $a_\alpha \in C$, the type $\operatorname{tp}(a_\alpha/U \cup a[<\alpha])$ is isolated by some formula $\varphi_\alpha(x,\bar{c})$ with $\bar{c} \subseteq U \cup (C \cap a[<\alpha]) = U \cup C_{<\alpha}$. Consequently, this formula also isolates the type $\operatorname{tp}(a_\alpha/U \cup C_{<\alpha})$. Hence, $\operatorname{tp}(a_\alpha/U \cup C_{<\alpha})$ is isolated, for all $a_\alpha \in C$, and we obtain a construction of C by omitting form $(a_\alpha)_{\alpha<\gamma}$ all elements that are not in C. ☐

Lemma 4.6. *Let $(a_\alpha)_{\alpha<\gamma}$ be a construction of A over U, C a closed subset of A, $\bar{c}$ an enumeration of C, and, for every $a_\alpha \in C$, let $\varphi_\alpha(x_\alpha; \bar{b}_\alpha)$ be a formula isolating* $\mathrm{tp}(a_\alpha / U \cup a[<\alpha])$. *Then*

$$T(U) \cup \{\, \varphi_\alpha(x_\alpha; \bar{b}_\alpha) \mid a_\alpha \in C \,\} \vDash \mathrm{tp}(\bar{c}/U)\,.$$

Proof. Note that $C_{<\alpha} := C \cap a[<\alpha]$ is closed. Hence, we can prove the claim by induction on α. For $\alpha = 0$ we have $\mathrm{tp}(\langle\rangle/U) = T(U)$. If α is a limit ordinal then the claim follows by inductive hypothesis since every formula refers only to finitely many elements of $C_{<\alpha}$. For the successor step, suppose that $\bar{c} = \bar{c}' a_\alpha$ where $\bar{c}'$ is an enumeration of $C_{<\alpha}$. By inductive hypothesis, we know that

$$T(U) \cup \{\, \varphi_i(x_i; \bar{b}_i) \mid i < \alpha,\ a_i \in C \,\} \vDash \mathrm{tp}(\bar{c}'/U)\,.$$

Furthermore,

$$T(U) \cup \{\varphi_\alpha(x_\alpha; \bar{b}_\alpha)\} \vDash \mathrm{tp}(a_\alpha / U \cup a[<\alpha]) \vDash \mathrm{tp}(a_\alpha / U \cup \bar{c}')\,.$$

Combining these two implications, the claim follows. □

Proposition 4.7. *Let C be a closed subset of a constructible set A. Then A is constructible over C.*

Proof. We start by showing that A is atomic over C. Let $A_0 \subseteq A$ be finite. By Lemma 4.5 (b), we can find a finite closed set D containing A_0. For $X \subseteq A$, set

$$\Phi(X) := \{\, \varphi_\beta(x_\beta; \bar{b}_\beta) \mid a_\beta \in X \,\}\,,$$

where $\varphi_\beta(x_\beta; \bar{b}_\beta)$ is some formula isolating $\mathrm{tp}(a_\beta / a[<\beta])$. According to Lemma 4.6 we have

$$T \cup \Phi(\bar{b}) \vDash \mathrm{tp}(\bar{b})\,, \quad \text{for every closed set } \bar{b} \subseteq A\,.$$

In particular, we have

$$T \cup \Phi(C \cup D) \vDash \mathrm{tp}(\bar{c}\bar{d})\,,$$

where $\bar{c}$ is an enumeration of C and $\bar{d}$ one of D. As $\Phi(C) \subseteq T(C)$, it follows that

$$T(C) \cup \Phi(D) \vDash \operatorname{tp}(\bar{d}/C) .$$

Hence, $\operatorname{tp}(\bar{d}/C)$ is isolated by the formula $\bigwedge \Phi(D)$. In particular, the type of A_0 over C is isolated.

To conclude the proof, let $(a_\alpha)_{\alpha<\gamma}$ be a construction of A. We prove that it is also a construction over C. Let $\alpha < \gamma$. Since $a[<\alpha]$ is closed, so is $C \cup a[<\alpha]$. By the first part of the proof, it follows that a_α is atomic over $C \cup a[<\alpha]$. □

Lemma 4.8. *If $(a_\alpha)_{\alpha<\gamma}$ is a construction of A over U then it is also a construction of A over $U \cup C$, for every finite subset $C \subseteq A$.*

Proof. By Lemma 4.2, A is atomic over $U \cup a[<\alpha]$, for every $\alpha < \gamma$. In particular, $C \cup \{a_\alpha\}$ is atomic over $U \cup a[<\alpha]$. By Lemma 1.5, it follows that a_α is atomic over $U \cup a[<\alpha] \cup C$. □

To prove the uniqueness of constructible models, we employ back-and-forth arguments.

Definition 4.9. Let $\mathfrak{A}$ and $\mathfrak{B}$ be structures such that A and B are constructible over $\varnothing$. We define

$$I_{\mathrm{cl}}(\mathfrak{A}, \mathfrak{B}) := \{ p \in I_{\mathrm{FO}}(\mathfrak{A}, \mathfrak{B}) \mid \operatorname{dom} p \text{ and } \operatorname{rng} p \text{ are closed} \} .$$

Lemma 4.10. *Suppose that $\mathfrak{A}$ and $\mathfrak{B}$ are structures where A and B are constructible over $\varnothing$. Then $I_{\mathrm{cl}}(\mathfrak{A}, \mathfrak{B})$ is $\aleph_1$-bounded and it has the back-and-forth property with respect to itself.*

Proof. By symmetry, we only consider the forth property. Let $\bar{a} \mapsto \bar{b} \in I_{\mathrm{cl}}(\mathfrak{A}, \mathfrak{B})$ and $x \in A$. By induction on n, we construct finite tuples $\bar{c}_n \subseteq A$ and $\bar{d}_n \subseteq B$ such that $\bar{a}\bar{c}_0\bar{c}_1 \cdots \mapsto \bar{b}\bar{d}_0\bar{d}_1 \cdots \in I_{\mathrm{cl}}(\mathfrak{A}, \mathfrak{B})$, $x \in \bar{c}_0$, and

$$\langle \mathfrak{A}, \bar{a}\bar{c}_0 \dots \bar{c}_{n-1} \rangle \equiv \langle \mathfrak{B}, \bar{b}\bar{d}_0 \dots \bar{d}_{n-1} \rangle , \quad \text{for all } n < \omega .$$

We start with some finite closed set $\bar{c}_0$ containing x. For the inductive step, suppose that we have already defined $\bar{c}_0, \dots, \bar{c}_n$ and $\bar{d}_0, \dots, \bar{d}_{n-1}$ such that

$$\langle \mathfrak{A}, \bar{a}\bar{c}_0 \dots \bar{c}_{n-1} \rangle \equiv \langle \mathfrak{B}, \bar{b}\bar{d}_0 \dots \bar{d}_{n-1} \rangle .$$

Since $\mathfrak{A}$ is atomic over $\bar{a}$, we know that the type $\operatorname{tp}(\bar{c}_0 \dots \bar{c}_{n-1}\bar{c}/\bar{a})$ is isolated. By Lemma 1.5, it follows that the type $\operatorname{tp}(\bar{c}_n/\bar{a}\bar{c}_0 \dots \bar{c}_{n-1})$ is also isolated. As

$$\langle \mathfrak{A}, \bar{a}\bar{c}_0 \dots \bar{c}_{n-1} \rangle \equiv \langle \mathfrak{B}, \bar{b}\bar{d}_0 \dots \bar{d}_{n-1} \rangle ,$$

we can therefore find some tuple $\bar{d}_n \subseteq B$ with

$$\langle \mathfrak{A}, \bar{a}\bar{c}_0 \dots \bar{c}_{n-1}\bar{c}_n \rangle \equiv \langle \mathfrak{B}, \bar{b}\bar{d}_0 \dots \bar{d}_{n-1}\bar{d}_n \rangle .$$

If $\bar{b}\bar{d}_0 \dots \bar{d}_n$ is closed then we can stop. Otherwise, let $\bar{d}_{n+1}$ be a finite closed set containing $\bar{d}_n$. Again, since $\bar{b}\bar{d}_0 \dots \bar{d}_{n-1}$ is closed and the type $\operatorname{tp}(\bar{d}_{n+1}/\bar{b}\bar{d}_0 \dots \bar{d}_n)$ is isolated, we can find a tuple $\bar{c}_{n+1} \subseteq A$ such that

$$\langle \mathfrak{A}, \bar{a}\bar{c}_0 \dots \bar{c}_n\bar{c}_{n+1} \rangle \equiv \langle \mathfrak{B}, \bar{b}\bar{d}_0 \dots \bar{d}_n\bar{d}_{n+1} \rangle .$$

If $\bar{a}\bar{c}_0 \dots \bar{c}_{n+1}$ is closed we stop. Otherwise, choose a finite closed set $\bar{c}_{n+2}$ containing $\bar{c}_{n+1}$ and repeat the construction. □

Theorem 4.11 (Ressayre). *All constructible models of a complete theory T are isomorphic and strongly $\aleph_0$-homogeneous.*

Proof. Let $\mathfrak{A}$ and $\mathfrak{B}$ be constructible models of T. First, we show that $\mathfrak{A}$ and $\mathfrak{B}$ are isomorphic. Since constructible models are prime, it follows that we can embed $\mathfrak{A}$ into $\mathfrak{B}$ and vice versa. Hence, $\mathfrak{A}$ and $\mathfrak{B}$ have the same cardinality κ. It follows by Lemma 4.10 that $I_{\mathrm{cl}}(\mathfrak{A}, \mathfrak{B}) : \mathfrak{A} \cong^{\kappa \oplus \aleph_1}_{\mathrm{iso}} \mathfrak{B}$. Consequently, Lemma C4.4.10 implies that $\mathfrak{A} \cong \mathfrak{B}$.

It remains to show that $\mathfrak{A}$ is strongly $\aleph_0$-homogeneous. Suppose that

$$\langle \mathfrak{A}, \bar{a} \rangle \equiv \langle \mathfrak{A}, \bar{b} \rangle ,$$

for finite tuples $\bar{a}, \bar{b} \subseteq A$. By Lemma 4.8, these two expansions of $\mathfrak{A}$ are constructible models of the complete theory $T(\bar{a})$. As we have just shown, this implies that they are isomorphic. Hence, there is an automorphism of $\mathfrak{A}$ mapping $\bar{a}$ to $\bar{b}$. □

We apply these tools to show that $\aleph_0$-stable theories have prime models over all sets of parameters.

Lemma 4.12. *Let T be a totally transcendental theory and U a set of parameters. Then the isolated types are dense in $S^{\bar{s}}(U)$.*

Proof. Since $\mathrm{rk}_{\mathrm{CB}}(S^{\bar{s}}(U)) < \infty$ the statement follows from Proposition B5.5.12 (d). □

Proposition 4.13. *Let T be a totally transcendental theory. For every model $\mathfrak{M}$ of T and all parameters $U \subseteq M$, there exists an elementary substructure $\mathfrak{A} \preceq \mathfrak{M}$ such that A is constructible over U. In particular, $\mathfrak{A}$ is a prime model over U and atomic over U.*

Proof. By induction on α, we construct a sequence $(a_\alpha)_{\alpha<\gamma}$ of elements of M as follows. Suppose that we have already defined $(a_i)_{i<\alpha}$. If there is some $b \in M$ such that $\mathrm{tp}(b/U \cup a[<\alpha])$ is isolated then we select one such element and set $a_\alpha := b$. Otherwise, we stop the construction.

Let $A := a[<\gamma]$ be the set of all elements chosen. Clearly, $U \subseteq A$ and $(a_\alpha)_{\alpha<\gamma}$ is a construction of A over U. Hence, it remains to show that $\mathfrak{A} \preceq \mathfrak{M}$ where $\mathfrak{A}$ is the substructure induced by A.

We apply the Tarski-Vaught Test. Suppose that

$$\mathfrak{M} \vDash \varphi(\bar{b}, c), \quad \text{for } \bar{b} \subseteq A \text{ and } c \in M.$$

By Lemma 4.12, there exists an isolated type $\mathfrak{p} \in \langle \varphi(\bar{b}, y) \rangle \subseteq S^1(A)$. Let $d \in M$ be an element realising $\mathfrak{p}$. Since $\mathfrak{p} = \mathrm{tp}(d/A)$ is isolated, it follows by choice of $a[<\gamma]$ that $d \in a[<\gamma] \subseteq A$. Thus, we have found an element $d \in A$ with $\mathfrak{M} \vDash \varphi(\bar{b}, d)$. □

Combining the preceding proposition with Theorem 4.11, we obtain the following result.

Theorem 4.14. *Let T be a totally transcendental theory and let U be a set of parameters. There exists a prime model over U that is also atomic over U. Furthermore, all prime models over U are isomorphic over U.*

Corollary 4.15. *Let T be a totally transcendental theory and let U be a set of parameters. Every model that is prime over U is also atomic over U.*

E4. $\aleph_0$-categorical theories

1. $\aleph_0$-categorical theories and automorphisms

Model theory investigates axiomatisable classes of structures. One of the most basic question one can ask is how many structures of a given cardinality such a class contains.

Definition 1.1. A class $\mathcal{K}$ is *κ-categorical* if, up to isomorphism, it contains exactly one structure of size κ. Similarly, we call a theory T *κ-categorical* if $\mathrm{Mod}(T)$ is κ-categorical.

Example. (a) According to Theorem C4.1.5, the theory of open dense linear orders is $\aleph_0$-categorical.

(b) We have seen in Corollary B6.5.30 that the theory ACF_p of algebraically closed fields of characteristic p is κ-categorical for all uncountable cardinals κ. It has $\aleph_0$ different models of size $\aleph_0$. Hence, it is not $\aleph_0$-categorical.

(c) By Theorem D1.4.8, the same holds for the theory of divisible torsion-free abelian groups.

In this chapter we study $\aleph_0$-categorical theories. We start by showing that, for models of such theories, there is a tight relationship between definable relations and automorphisms. Recall that the automorphism group $\mathrm{Aut}\,\mathfrak{M}$ of a structure $\mathfrak{M}$ is *oligomorphic* if, for every finite tuple $\bar{s}$ of sorts, there are only finitely many orbits of $\mathrm{Aut}\,\mathfrak{M}$ on the set $M^{\bar{s}}$.

Theorem 1.2 (Engeler, Ryll-Nardzewski, Svenonius)**.** *Let T be a countable complete theory with infinite models. The following statements are equivalent:*

(1) *T is $\aleph_0$-categorical.*

(2) Aut $\mathfrak{M}$ *is oligomorphic, for every countable model $\mathfrak{M}$ of T.*

(3) *T has a countable model $\mathfrak{M}$ such that* Aut $\mathfrak{M}$ *is oligomorphic.*

(4) *There exists a countable model $\mathfrak{M} \vDash T$ in which, for every finite tuple of sorts $\bar{s}$, only finitely many $\bar{s}$-types (over $\varnothing$) are realised.*

(5) $|S^{\bar{s}}(T)| < \aleph_0$, *for all finite $\bar{s}$.*

(6) *For all finite sets $\bar{x}$ of variables, there are only finitely many formulae $\varphi(\bar{x})$ with free variables $\bar{x}$ that are pairwise non-equivalent modulo T.*

(7) *Every type $\mathfrak{p} \in S^{<\omega}(T)$ is isolated.*

(8) *T has a model that is atomic and $\aleph_0$-saturated.*

(9) *Every model of T is atomic.*

(10) *Every model of T is $\aleph_0$-saturated.*

(11) $\mathfrak{A} \cong^{\aleph_0}_{\mathrm{FO}} \mathfrak{B}$, *for all models $\mathfrak{A}$ and $\mathfrak{B}$ of T.*

(12) $\mathfrak{A} \cong_\infty \mathfrak{B}$, *for all models $\mathfrak{A}$ and $\mathfrak{B}$ of T.*

Proof. (5) $\Rightarrow$ (6) If $\langle\varphi\rangle = \langle\psi\rangle$ then $\varphi \equiv \psi$ modulo T. If $|S^{\bar{s}}(T)| = k < \aleph_0$ then there are at most 2^k sets of the form $\langle\varphi\rangle$ and, hence, at most that many non-equivalent formulae.

(6) $\Rightarrow$ (7) For all finite tuples of sorts $\bar{s}$, fix a tuple of variables $\bar{x}$ of sort $\bar{s}$ and a maximal family $\Phi_{\bar{s}}$ of pairwise non-equivalent formulae with free variables $\bar{x}$. For $\mathfrak{p} \in S^{\bar{s}}(T)$, let

$$\psi_{\mathfrak{p}} := \bigwedge \{ \varphi \in \Phi_{\bar{s}} \mid \mathfrak{p} \in \langle\varphi\rangle \} .$$

Then $T \cup \{\psi_{\mathfrak{p}}\} \vDash \mathfrak{p}$ and $\mathfrak{p}$ is isolated.

(7) $\Rightarrow$ (5) If every type in $S^{\bar{s}}(T)$ is isolated then $S^{\bar{s}}(T)$ is finite, by Lemma B5.5.10.

(7) $\Rightarrow$ (9) Each model can only realise isolated types since there are no non-isolated ones.

(9) $\Rightarrow$ (8) Every consistent theory has $\aleph_0$-saturated models.

(8) ⇒ (7) If there is a non-isolated type $\mathfrak{p} \in S^{<\omega}(T)$ then it is realised in all $\aleph_0$-saturated models. Consequently, none of them can be atomic.

(7) ⇒ (10) Suppose that $\mathfrak{M} \vDash T$ is a model, $\bar{a} \in M^m$ a finite tuple, and $\mathfrak{p} \in S^n(\bar{a})$. There is an elementary extension $\mathfrak{N} \geq \mathfrak{M}$ in which $\mathfrak{p}$ is realised by some tuple $\bar{c} \in N^n$. Set $\mathfrak{q} := \mathrm{tp}(\bar{a}\bar{c}/\mathfrak{N})$. Then $\mathfrak{q} \in S^{m+n}(T)$ and, by hypothesis, there is some formula $\varphi(\bar{x}, \bar{y})$ isolating $\mathfrak{q}$. Let $\psi(\bar{x})$ be the formula isolating $\mathfrak{r} := \mathrm{tp}(\bar{a}/\mathfrak{M})$. We claim that

$$T \vDash \psi(\bar{x}) \to \exists \bar{y} \varphi(\bar{x}, \bar{y}) .$$

Then it follows that $\mathfrak{M} \vDash \exists \bar{y} \varphi(\bar{a}, \bar{y})$ and we can find some tuple $\bar{b} \in M^n$ realising $\mathfrak{p}$.

It remains to prove the claim. For a contradiction, suppose it does not hold. Since $\mathfrak{r}$ is complete it follows that $\neg \exists \bar{y} \varphi \in \mathfrak{r}$ and, therefore,

$$T \vDash \psi(\bar{x}) \to \forall \bar{y} \neg \varphi(\bar{x}, \bar{y}) .$$

On the other hand, $\mathfrak{r} \subseteq \mathfrak{q}$ implies that

$$T \vDash \varphi(\bar{x}, \bar{y}) \to \psi(\bar{x}) .$$

Consequently, $T \cup \{\varphi(\bar{x}, \bar{y})\}$ is inconsistent. But this contradicts the fact that $\mathfrak{q} \in S^{m+n}(T)$.

(10) ⇒ (11) follows from Corollary E1.2.3.

(11) ⇒ (12) immediately, since $\mathfrak{A} \cong^{\aleph_0}_{\mathrm{FO}} \mathfrak{B}$ implies $\mathfrak{A} \cong_\infty \mathfrak{B}$.

(12) ⇒ (1) Since T is a countable theory with infinite models it follows that T has a model of cardinality $\aleph_0$. Furthermore, by (12) and Lemma C4.4.10, all such models are isomorphic.

(1) ⇒ (7) Suppose that there exists a type $\mathfrak{p} \in S^{<\omega}(T)$ that is not isolated. T has a model $\mathfrak{A}$ in which $\mathfrak{p}$ is not realised, and it has a model $\mathfrak{B}$ in which $\mathfrak{p}$ is realised. By the Theorem of Löwenheim and Skolem, we can assume that $|A| = |B| = \aleph_0$. Since $\mathfrak{A} \ncong \mathfrak{B}$ T cannot be $\aleph_0$-categorical.

(5) ⇒ (2) Let $\mathfrak{A}$ be a countable model of T and let $\mathfrak{p} \in S^n(T)$. We claim that all tuples realising $\mathfrak{p}$ are in the same orbit of $\mathrm{Aut}\,\mathfrak{A}$. Hence, the number of orbits is bounded by the number of types which, by (5), is finite.

Suppose that $\bar{a}, \bar{b} \in A^n$ realise $\mathfrak{p}$. We have already seen that (5) implies (11). Hence, we have $\mathfrak{A} \cong_{\mathrm{FO}}^{\aleph_0} \mathfrak{A}$, and $\bar{a} \mapsto \bar{b} \in I_{\mathrm{FO}}^{\aleph_0}(\mathfrak{A}, \mathfrak{A})$ implies that $\langle \mathfrak{A}, \bar{a} \rangle \cong_{\mathrm{FO}}^{\aleph_0} \langle \mathfrak{A}, \bar{b} \rangle$. By Corollary E1.2.3, it follows that there exists an automorphism π with $\pi(\bar{a}) = \bar{b}$.

(2) ⇒ (3) is trivial since T is satisfiable.

(3) ⇒ (4) We have $\mathrm{tp}(\pi\bar{a}) = \mathrm{tp}(\bar{a})$, for all $\pi \in \mathrm{Aut}\,\mathfrak{M}$. Hence, the number of realised types is bounded by the number of orbits.

(4) ⇒ (5) Fix a countable model $\mathfrak{M} \vDash T$ in which only finitely many $\bar{s}$-types are realised, for all finite $\bar{s}$. For a given $\bar{s}$, let $\mathfrak{p}_0, \dots, \mathfrak{p}_{k-1}$ be an enumeration of these $\bar{s}$-types. By Lemma C3.2.6, the set $\{\mathfrak{p}_0, \dots, \mathfrak{p}_{k-1}\}$ is dense in $S^{\bar{s}}(T)$. Consequently, it follows by Lemma B5.5.10 that $S^{\bar{s}}(T)$ is finite. □

Let us also mention a necessary condition for $\aleph_0$-categoricity that deals with the size of the algebraic closure of finite sets.

Lemma 1.3. *Let T be a countable $\aleph_0$-categorical theory with finitely many sorts. There exists a function $s : \omega \to \omega$ such that, for every model $\mathfrak{M}$ of T and every finite set $U \subseteq M$ of parameters, we have*

$$|\mathrm{acl}(U)| \leq s(|U|)\,.$$

In particular, $\mathrm{acl}(U)$ *is finite for finite sets U.*

Proof. Let $n := |U|$. By Theorem 1.2, $S^{n+1}(T)$ is finite. Let $\mathfrak{p}_0, \dots, \mathfrak{p}_{k-1}$ be an enumeration of $S^{n+1}(T)$ and set

$$I := \Big\{ i < k \;\Big|\; \text{there are } \varphi(x, \bar{y}) \in \mathfrak{p}_i \text{ and } m < \omega \text{ such that } T \vDash \neg\exists^m x \varphi(x, \bar{y}) \Big\}\,.$$

For $i \in I$, let $m_i < \omega$ be the least number such that

$$\neg\exists^{m_i} x \varphi(x, \bar{y}) \in \mathfrak{p}_i\,, \quad \text{for some formula } \varphi(x, \bar{y})\,.$$

We set $s(n) := \sum_{i \in I} m_i$. Let $a \in \mathrm{acl}(U)$ and let $\bar{b} \in M^n$ be an enumeration of U. The tuple $a\bar{b}$ realises some type $\mathfrak{p}_i$ with $i \in I$. Since there are at most m_i elements c such that $c\bar{b}$ realises $\mathfrak{p}_i$, it follows that

$$|\mathrm{acl}(U)| \leq \sum_{i \in I} m_i = s(n)\,. \qquad \square$$

As an application, we consider fields and groups.

Lemma 1.4. *No infinite field has an $\aleph_0$-categorical theory.*

Proof. Let $\mathfrak{K}$ be an infinite field. By compactness, there exists an elementary extension $\mathfrak{K}_+ \succeq \mathfrak{K}$ that contains a transcendental element c. The algebraic closure $\mathrm{acl}(c)$ is infinite since it contains the elements $c, c^2, c^3, \dots$, which are all distinct. By Lemma 1.3, it follows that $\mathrm{Th}(\mathfrak{K})$ is not $\aleph_0$-categorical. $\square$

Lemma 1.5. *Let $\mathfrak{G}$ be an infinite group.*

(a) *If $\mathrm{Th}(\mathfrak{G})$ is $\aleph_0$-categorical then $\mathfrak{G}$ is locally finite and there exists a number $n < \omega$ such that $g^n = 1$, for all $g \in G$.*

(b) *Conversely, if $\mathfrak{G}$ is abelian and there exists a number n as in* (a), *then $\mathrm{Th}(\mathfrak{G})$ is $\aleph_0$-categorical.*

Proof. (a) Fix an element $g \in G$ and let $s : \omega \to \omega$ be the function from Lemma 1.3. Since $g^n \in \mathrm{acl}(g)$, for all $n < \omega$, and $|\mathrm{acl}(g)| \leq s(1)$, there is some $n < s(1)$ such that $g^{s(1)} = g^n$. Consequently, $g^{s(1)-n} = 1$. Setting $m := s(1)!$ it follows that $g^m = 1$ for all $g \in G$.

(b) Let $\mathfrak{G}$ be a countable abelian group such that $g^n = 1$, for all $g \in G$. There are prime numbers $p_0, \dots, p_{m-1}$, numbers $k_0, \dots, k_{m-1} < \omega$, and cardinals $\lambda_0, \dots, \lambda_{m-1} \leq \aleph_0$ such that

$$\mathfrak{G} \cong \bigoplus_{i<m} (\mathbb{Z}/p_i^{k_i}\mathbb{Z})^{(\lambda_i)}.$$

Set $q_i := p_i^{k_i}$. Note that, for $\lambda_i < \aleph_0$, the group $(\mathbb{Z}/q_i\mathbb{Z})^{(\lambda_i)}$ has

$$r_i := p_i^{\lambda_i k_i} - p_i^{\lambda_i(k_i-1)}$$

elements of order exactly q_i, and, for each element $g \in (\mathbb{Z}/q_i\mathbb{Z})^{(\lambda_i)}$ of order less than q_i, there exists some element h such that $g = h^{p_i}$.

It follows that $\mathfrak{G}$ satisfies the following formula:

- the axioms of an abelian group;
- $\forall x(x^{q_0 \cdots q_{m-1}} = 1)$
- for each $i < m$ such that $\lambda_i < \aleph_0$, the statement that there are exactly r_i elements of order exactly q_i that cannot be written in the form h^{p_i}, for some $h \in G$;
- for each $i < m$ such that $\lambda_i = \aleph_0$, the statement that there are infinitely many elements of order exactly q_i that cannot be written in the form h^{p_i}, for some $h \in G$.

Furthermore, every countable structure $\mathfrak{H}$ satisfying these formulae is isomorphic to $\mathfrak{G}$. Consequently, $\mathrm{Th}(\mathfrak{G})$ is $\aleph_0$-categorical. □

Having characterised the countable theories with exactly one countable model we turn to countable theories with several countable models.

Lemma 1.6. *If T is a countable complete theory with less than $2^{\aleph_0}$ countable models, up to isomorphism, then $|S^{\bar{s}}(T)| \leq \aleph_0$, for all finite $\bar{s}$.*

Proof. Assume that $S^{\bar{s}}(T)$ is uncountable. Then we have $|S^{\bar{s}}(T)| = 2^{\aleph_0}$, by Corollary B5.7.5. Each type $\mathfrak{p} \in S^{\bar{s}}(T)$ is realised in some countable model of T. Since each countable model of T realises only countably many types it follows that T has $2^{\aleph_0}$ models. □

Surprisingly there are no theories with exactly two countable models.

Theorem 1.7. *Let T be a countable complete theory. If T is not $\aleph_0$-categorical then it has at least 3 countable models.*

Proof. If there is a finite tuple $\bar{s}$ of sorts such that $S^{\bar{s}}(T)$ is uncountable then it follows by Lemma 1.6 that T has uncountably many countable models. Hence, we may assume that $S^{\bar{s}}(T)$ is countable, for all $\bar{s}$. By Theorem E3.3.1 and Proposition E1.2.15, it follows that T has a prime model $\mathfrak{A}$ and a countable saturated model $\mathfrak{B}$. If T is not $\aleph_0$-categorical then there is some $\bar{s}$ such that $S^{\bar{s}}(T)$ is infinite and there exists a non-isolated type $\mathfrak{p} \in S^{\bar{s}}(T)$. This type is realised in $\mathfrak{B}$ but not in $\mathfrak{A}$ which implies that $\mathfrak{A} \ncong \mathfrak{B}$.

Let $\bar{a} \in B^{\bar{s}}$ be a tuple of type $\mathfrak{p}$. We know that, for some $k < \omega$, there are infinitely many pairwise non-equivalent formulae with free variables $x_0, \ldots, x_{k-1}$. These formulae are still non-equivalent modulo the theory $\mathrm{Th}(\mathfrak{B}_{\bar{a}})$. Hence, $\mathrm{Th}(\mathfrak{B}_{\bar{a}})$ is not $\aleph_0$-categorical and there exists a prime model $\mathfrak{C}$ of this theory. We have $\mathfrak{C} \ncong \mathfrak{A}$ since $\mathfrak{p}$ is realised in $\mathfrak{C}$. As $\mathfrak{C}$ is not $\aleph_0$-saturated there is a non-isolated type $\mathfrak{q} \in S^{<\omega}(\bar{a})$. Since $\mathfrak{B}$ realises $\mathfrak{q}$ and $\mathfrak{C}$ does not, it follows that $\mathfrak{C} \ncong \mathfrak{B}$. Thus, we have found three non-isomorphic models $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$. □

Lemma 1.8. *There is a countable complete theory T which has exactly three countable models.*

Proof. Let T be the theory of open dense linear orders augmented by the sentences $c_i < c_k$, for all $i < k < \omega$. This theory is complete, it admits quantifier elimination, and the only non-isolated type $\mathfrak{p}$ is the one containing all formulae $x > c_i$, $i < \omega$. There are three models.

(i) The prime model is $\mathfrak{M}_0 \cong \langle \mathbb{Q}, <, (n)_{n<\omega} \rangle$ where the type $\mathfrak{p}$ is not realised since the sequence $(c_i)_i$ is unbounded.

(ii) In $\mathfrak{M}_1 \cong \left\langle \mathbb{Q}, <, \left((1+\frac{1}{n})^n\right)_{n<\omega} \right\rangle$ the sequence $(c_i)_i$ is bounded but it has no least upper bound.

(iii) In $\mathfrak{M}_2 \cong \left\langle \mathbb{Q}, <, \left(-\frac{1}{n}\right)_{n<\omega} \right\rangle$ the sequence $(c_i)_i$ has a least upper bound. □

Exercise 1.1. For every $3 < n < \omega$, find a countable complete first-order theory with exactly n models.

All possibilities for the number of countable models of a countable theory are listed in the following theorem. Each of them is realised by some theory. The question of whether there are really countable theories with exactly $\aleph_1$ countable models was open for a long time. An affirmative answer was recently given by Knight.

Theorem 1.9 (Morley). *The number of nonisomorphic countably infinite models of a countable complete theory is either a finite number $n \neq 2$, or it is one of $\aleph_0$, $\aleph_1$, or $2^{\aleph_0}$.*

We will not give the complete proof of this result. The next lemma characterises those theories with at most $\aleph_1$ countable models. Morley has shown that all theories that do not satisfy the conditions of the lemma have $2^{\aleph_0}$ countable models.

Lemma 1.10. *Let T be a countable complete theory and let $\mathcal{K}$ be the class of all countable models of T. If we have*

$$|\mathcal{K}/{\equiv_\alpha}| \leq \aleph_0\,, \quad \textit{for every } \alpha < \omega_1\,,$$

then, up to isomorphism, T has at most $\aleph_1$ countable models.

Proof. For $\mathfrak{A} \in \mathcal{K}$, let $\chi(\mathfrak{A}) := \langle \alpha, [\mathfrak{A}]_\alpha \rangle$ where α is the Scott height of $\mathfrak{A}$ and $[\mathfrak{A}]_\alpha \in \mathcal{K}/{\equiv_{\alpha+\omega}}$ is the $\equiv_{\alpha+\omega}$-class of $\mathfrak{A}$. By Corollary C4.4.11, it follows that we have

$$\chi(\mathfrak{A}) = \chi(\mathfrak{B}) \quad \text{iff} \quad \mathfrak{A} \cong \mathfrak{B}\,, \quad \text{for all } \mathfrak{A}, \mathfrak{B} \in \mathcal{K}\,.$$

Consequently, the number of countable models of T is at most

$$|\operatorname{rng} \chi| \leq \aleph_1 \otimes \sup \{ |\mathcal{K}/{\equiv_\alpha}| \mid \alpha < \omega_1 \} \leq \aleph_1 \otimes \aleph_1 = \aleph_1\,. \qquad \square$$

We conclude this section by an investigation of definable relations in countable models of $\aleph_0$-categorical theories.

Lemma 1.11. *Let $\mathfrak{M}$ be a countable model of a countable $\aleph_0$-categorical theory T.*

(a) *Let $\bar{s}$ be a finite tuple of sorts. A relation $R \subseteq M^{\bar{s}}$ is definable in $\mathfrak{M}$ if and only if*

$$\pi[R] = R\,, \quad \textit{for all } \pi \in \operatorname{Aut} \mathfrak{M}\,.$$

(b) *A partial function $f : M_s \to M_t$ is definable in $\mathfrak{M}$ if and only if*

$$\pi \circ f = f \circ \pi\,, \quad \textit{for all } \pi \in \operatorname{Aut} \mathfrak{M}\,.$$

Proof. (a) For the nontrivial direction suppose that $\pi[R] = R$ for all automorphisms π. Since T is $\aleph_0$-categorical there are only finitely many orbits of $\operatorname{Aut}\mathfrak{M}$ on $M^{\bar{s}}$. Hence, R is a finite union of such orbits and it is sufficient to prove that every orbit S is definable.

Fix some tuple $\bar{a} \in S$. We have seen in Theorem 1.2 that $\mathfrak{M}$ is saturated. Hence, it follows by Lemma E1.4.2 and Proposition E1.4.7 that $\mathfrak{M}$ is strongly homogeneous. Consequently, $\operatorname{tp}(\bar{a}) = \operatorname{tp}(\bar{b})$ implies that there is some automorphism π mapping $\bar{a}$ to $\bar{b}$. It follows that

$$S = \{ \bar{b} \in M^{\bar{s}} \mid \operatorname{tp}(\bar{b}) = \operatorname{tp}(\bar{a}) \}\,.$$

Since every type is isolated there is some formula $\varphi(\bar{x})$ with

$$\mathfrak{M} \vDash \varphi(\bar{b}) \quad \text{iff} \quad \operatorname{tp}(\bar{b}) = \operatorname{tp}(\bar{a})\,.$$

It follows that $S = \varphi^{\mathfrak{M}}$.

(b) By (a), a function f is definable if and only if it is invariant under automorphisms, i.e., if and only if

$$b = f(a) \quad \text{iff} \quad \pi(b) = f(\pi(a))\,, \quad \text{for all } \pi \in \operatorname{Aut}\mathfrak{M}\,.$$

We can rewrite this condition as $\pi(f(a)) = f(\pi(a))$. □

We can use these results to relate interpretations and automorphism groups.

Definition 1.12. (a) Let $\mathfrak{A}$ and $\mathfrak{B}$ be structures. $\mathfrak{B}$ is *definable* in $\mathfrak{A}$ if it is isomorphic to a structure $\mathfrak{C}$ each domain C_s of which is a definable subset of A such that all relations $R^{\mathfrak{C}}$ and functions $f^{\mathfrak{C}}$ are definable in $\mathfrak{A}$. We call $\mathfrak{A}$ and $\mathfrak{B}$ *bidefinable* if each of them is definable in the other one and the corresponding isomorphisms are inverses of each other.

Definition 1.13. Suppose that $\mathfrak{G}$ and $\mathfrak{H}$ are permutation groups with actions $\alpha : \mathfrak{G} \to \mathfrak{Sym}\,\Omega$ and $\beta : \mathfrak{H} \to \mathfrak{Sym}\,\Delta$, respectively.

(a) A *morphism* $\mathfrak{G} \to \mathfrak{H}$ (or, more precisely, $\alpha \to \beta$) is a pair $\langle h, i \rangle$ where $h : \mathfrak{G} \to \mathfrak{H}$ is a group homomorphism and $i : \Delta \to \Omega$ is a function such that

$$\alpha(g) \circ i = i \circ \beta(h(g)), \quad \text{for all } g \in G.$$

(b) An *embedding* of permutation groups is a morphism $\langle h, i \rangle : \mathfrak{G} \to \mathfrak{H}$ where h and i are both injective.

Theorem 1.14. *Let $\mathfrak{A}$ be a countable model of a countable $\aleph_0$-categorical theory. A structure $\mathfrak{B}$ is definable in $\mathfrak{A}$ if and only if there exists an embedding $\mathfrak{Aut}\,\mathfrak{A} \to \mathfrak{Aut}\,\mathfrak{B}$.*

Proof. The claim follows from Lemma 1.11. If $\mathfrak{B}$ is definable in $\mathfrak{A}$ then every relation $R^{\mathfrak{B}}$ of $\mathfrak{B}$ is closed under $\mathfrak{Aut}\,\mathfrak{A}$. This implies that every automorphism of $\mathfrak{A}$ is also an automorphism of $\mathfrak{B}$. Conversely, each relation $R^{\mathfrak{B}}$ of $\mathfrak{B}$ is closed under all automorphisms of $\mathfrak{B}$. If $\mathfrak{Aut}\,\mathfrak{A} \leq \mathfrak{Aut}\,\mathfrak{B}$ then it also closed under all automorphisms of $\mathfrak{A}$ and, hence, it is definable in $\mathfrak{A}$. □

Corollary 1.15. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be countable models of countable $\aleph_0$-categorical theories. Then $\mathfrak{A}$ and $\mathfrak{B}$ are bidefinable if and only if $\mathfrak{Aut}\,\mathfrak{A}$ and $\mathfrak{Aut}\,\mathfrak{B}$ are isomorphic as permutation groups.*

Corollary 1.16. *Let $\mathfrak{A}$ be a countable model of a countable $\aleph_0$-categorical theory. If $\mathfrak{B}$ is a structure with countable signature that is definable in $\mathfrak{A}$ then $\mathrm{Th}(\mathfrak{B})$ is also $\aleph_0$-categorical.*

Proof. If $\mathfrak{Aut}\,\mathfrak{A}$ is oligomorphic and $\mathfrak{Aut}\,\mathfrak{B} \geq \mathfrak{Aut}\,\mathfrak{A}$ then $\mathfrak{Aut}\,\mathfrak{B}$ is also oligomorphic. □

A similar characterisation holds for interpretations. Recall that every structure interpretable in $\mathfrak{M}$ can be seen as a definable substructure of $\mathfrak{M}^{\mathrm{eq}}$.

Definition 1.17. Let $\mathcal{I} = \langle \alpha, (\delta_s)_{s \in S}, (\varepsilon_s)_{s \in S}, (\varphi_\xi)_{\xi \in \Gamma} \rangle$ be a first-order interpretation and $\pi : \mathfrak{A} \to \mathfrak{B}$ an isomorphism.

(a) We denote by $\pi^{\mathrm{eq}} : \mathfrak{A}^{\mathrm{eq}} \to \mathfrak{B}^{\mathrm{eq}}$ the unique isomorphism with $\pi^{\mathrm{eq}} \restriction A = \pi$.

(b) Set $\mathfrak{C} := \mathcal{I}(\mathfrak{A})$. For every sort s, the coordinate map of $\mathcal{I}$ induces a bijection $\mathcal{I}_s : D_s \to C_s$ where

$$D_s := \{ [\bar{a}]_{\varepsilon_s} \mid \bar{a} \in \delta_s^{\mathfrak{A}} \} \subseteq A^{\mathrm{eq}}_{\varepsilon_s} .$$

(c) We define

$$\pi^{\mathcal{I}} := \bigcup_s \mathcal{I}_s \circ \pi^{\mathrm{eq}} \circ \mathcal{I}_s^{-1} ,$$

where s ranges over all sorts of $\mathcal{I}(\mathfrak{A})$. We denote the induced map on automorphism groups by $\operatorname{Aut}\mathcal{I} : \mathfrak{Aut}\,\mathfrak{A} \to \mathfrak{Aut}\,\mathcal{I}(\mathfrak{A}) : \pi \mapsto \pi^{\mathcal{I}}$.

Lemma 1.18. *Let $\mathcal{I}$ be a first-order interpretation. $\pi^{\mathcal{I}} : \mathcal{I}(\mathfrak{A}) \to \mathcal{I}(\mathfrak{B})$ is an isomorphism, for every isomorphism $\pi : \mathfrak{A} \to \mathfrak{B}$.*

Lemma 1.19. *Every isomorphism $h : \mathfrak{A} \to \mathfrak{B}$ induces an isomorphism $\mathfrak{Aut}\,h : \mathfrak{Aut}\,\mathfrak{A} \to \mathfrak{Aut}\,\mathfrak{B}$ where*

$$(\mathfrak{Aut}\,h)(\pi) := h \circ \pi \circ h^{-1}.$$

Lemma 1.20. *For every first-order interpretation $\mathcal{I}$, the map $\mathfrak{Aut}\,\mathcal{I}$ is a continuous homomorphism*

$$\mathfrak{Aut}\,\mathcal{I} : \mathfrak{Aut}\,\mathfrak{M} \to \mathfrak{Aut}\,\mathcal{I}(\mathfrak{M}) .$$

Proof. It is straightforward to verify that $\mathfrak{Aut}\,\mathcal{I} : \mathfrak{Aut}\,\mathfrak{M} \to \mathfrak{Aut}\,\mathcal{I}(\mathfrak{M})$ is a homomorphism. To see that it is continuous let $S \subseteq \operatorname{Aut}\mathcal{I}(\mathfrak{M})$ be a basic open neighbourhood of 1. Then there is some finite tuple $\bar{a}$ in $\mathcal{I}(\mathfrak{M})$ such that

$$S = (\operatorname{Aut}\mathcal{I}(\mathfrak{M}))_{(\bar{a})} .$$

Suppose that the sorts of $\bar{a}$ are $\bar{s}$. We fix elements $c_i \in D_{s_i}$ with $\mathcal{I}(c_i) = a_i$. There are finite tuples $\bar{c}_i^* \subseteq M$ such that

$$\mathrm{dcl}^{\mathrm{eq}}(c_i) = \mathrm{dcl}^{\mathrm{eq}}(\bar{c}_i^*) .$$

Setting $S' := (\mathfrak{Aut}\,\mathcal{I})^{-1}[S]$ we have

$$\begin{aligned} \pi \in S' \quad &\text{iff} \quad \mathfrak{Aut}\,\mathcal{I}(\pi)(\bar{a}) = \bar{a} \\ &\text{iff} \quad (\mathcal{I}_{s_i} \circ \pi^{\mathrm{eq}} \circ \mathcal{I}_{s_i}^{-1})(a_i) = a_i\,, \quad \text{for all } i \\ &\text{iff} \quad \pi^{\mathrm{eq}}(c_i) = c_i\,, \quad \text{for all } i \\ &\text{iff} \quad \pi(\bar{c}_i^*) = \bar{c}_i^*\,, \quad \text{for all } i\,. \end{aligned}$$

Consequently, $S' = (\mathrm{Aut}\,\mathfrak{M})_{(\bar{c}_0^* \dots \bar{c}_{m-1}^*)}$ is open. □

Let us call a function $f : M \to M$ *definable* in the structure $\mathfrak{M}$ if each restriction $f \restriction M_s$ is definable, where s ranges over all sorts of $\mathfrak{M}$.

Lemma 1.21. *Let $\varphi : \mathfrak{Aut}\,\mathfrak{A} \to \mathfrak{Aut}\,\mathfrak{B}$ be a continuous homomorphism and suppose that $\mathfrak{A}$ is a countable model of an $\aleph_0$-categorical theory. The following statements are equivalent:*

(1) $\varphi = \mathrm{Aut}\,\pi \circ \mathrm{Aut}\,\mathcal{I}$, *for some interpretation $\mathcal{I}$ and some isomorphism* $\pi : \mathcal{I}(\mathfrak{A}) \to \mathfrak{B}$.

(2) *The subgroup* $\mathrm{rng}\,\varphi \leq \mathfrak{Aut}\,\mathfrak{B}$ *is oligomorphic.*

Proof. (1) ⇒ (2) For every finite tuple $\bar{s}$ of sorts and every orbit S of $\mathrm{rng}\,\varphi$ on $B^{\bar{s}}$, we introduce a new relation R_S of type $\bar{s}$ containing all tuples in the orbit S. Let $\mathfrak{B}^+$ be the expansion of $\mathfrak{B}$ by all these relations R_S. Every automorphism $\sigma \in \mathrm{rng}\,\varphi$ is still an automorphism of the expansion $\mathfrak{B}^+$. Hence, $\mathrm{rng}\,\varphi \leq \mathfrak{Aut}\,\mathfrak{B}^+$. We claim that $\mathrm{rng}\,\varphi$ and $\mathfrak{Aut}\,\mathfrak{B}^+$ have the same orbits.

Since $\mathrm{rng}\,\varphi \leq \mathfrak{Aut}\,\mathfrak{B}^+$ it is sufficient to check that tuples $\bar{a}, \bar{b} \in B^{\bar{s}}$ in different orbits of $\mathrm{rng}\,\varphi$ belong to different orbits of $\mathfrak{Aut}\,\mathfrak{B}^+$. Let S and S' be the orbits under $\mathrm{rng}\,\varphi$ of $\bar{a}$ and $\bar{b}$, respectively. Then $\bar{a} \in R_S$ and $\bar{b} \in R_{S'}$. If $S \neq S'$ then R_S and $R_{S'}$ are disjoint and there is no automorphism of $\mathfrak{B}^+$ mapping $\bar{a}$ to $\bar{b}$.

Consequently, $\mathrm{rng}\,\varphi$ and $\mathfrak{Aut}\,\mathfrak{B}^+$ have the same orbits. To prove (2) it is therefore sufficient to show that $\mathfrak{Aut}\,\mathfrak{B}^+$ is oligomorphic. For a contradiction, suppose that some set $B^{\bar{s}}$ contains tuples $\bar{b}^n$, $n < \omega$, from pairwise distinct orbits. Fix tuples $\bar{a}^n \subseteq A$ such that $(\pi \circ \mathcal{I})(\bar{a}^n) = \bar{b}^n$.

Since $\mathfrak{A}$ is $\aleph_0$-categorical there are indices $k < n$ such that $\bar{a}^k$ and $\bar{a}^n$ belong to the same orbit under $\mathfrak{Aut}\,\mathfrak{A}$. Fix an automorphism $\sigma \in \mathfrak{Aut}\,\mathfrak{A}$ with $\sigma(\bar{a}^k) = \bar{a}^n$. Then

$$\begin{aligned}\varphi(\sigma)(\bar{b}^k) &= (\operatorname{Aut}\pi \circ \operatorname{Aut}\mathcal{I})(\sigma)(\bar{b}^k) \\ &= (\pi \circ \mathcal{I}_{\bar{s}} \circ \sigma^{\mathrm{eq}} \circ \mathcal{I}_{\bar{s}}^{-1} \circ \pi^{-1})(\bar{b}^k) \\ &= (\pi \circ \mathcal{I}_{\bar{s}} \circ \sigma^{\mathrm{eq}})(\bar{a}^k) \\ &= (\pi \circ \mathcal{I}_{\bar{s}})(\bar{a}^n) = \bar{b}^n .\end{aligned}$$

Hence, the automorphism $\varphi(\sigma)$ maps $\bar{b}^k$ to $\bar{b}^n$. Contradiction.

(2) $\Rightarrow$ (1) Let $\mathfrak{G} := \mathfrak{Aut}\,\mathfrak{A}$ and $\mathfrak{H} := \mathfrak{Aut}\,\mathfrak{B}$. For each sort s, fix representatives $b^s_0, b^s_1, \dots$ of the orbits of B_s under $\operatorname{rng}\varphi$. The stabiliser $\mathfrak{H}_{(b^s_i)}$ of b^s_i is a basic open neighbourhood of 1 in $\mathfrak{H}$. Since φ is continuous we can find, for each b^s_i, a basic open neighbourhood U^s_i of 1 in $\mathfrak{G}$ with

$$U^s_i \subseteq \varphi^{-1}[\mathfrak{H}_{(b^s_i)}] .$$

Every such neighbourhood is of the form $U^s_i = \mathfrak{G}_{(\bar{a}^s_i)}$, for some $\bar{a}^s_i \subseteq A$. Let O^s_i be the orbit of $\bar{a}^s_i$. We define a map $\pi^s_i : O^s_i \to B_s$ by

$$\pi^s_i(\sigma(\bar{a}^s_i)) := \varphi(\sigma)(b^s_i) , \quad \text{for } \sigma \in \mathfrak{G} .$$

It follows that $\operatorname{rng}\pi^s_i$ is the orbit of b^s_i under $\operatorname{rng}\varphi$. Note that $\ker\pi^s_i$ is invariant under automorphisms since

$$\pi^s_i(\sigma_0(\bar{a}^s_i)) = \pi^s_i(\sigma_1(\bar{a}^s_i))$$

implies

$$\begin{aligned}\pi^s_i((\rho \circ \sigma_0)(\bar{a}^s_i)) &= \varphi(\rho \circ \sigma_0)(b^s_i) \\ &= \varphi(\rho \circ \sigma_1)(b^s_i) = \pi^s_i((\rho \circ \sigma_1)(\bar{a}^s_i)) .\end{aligned}$$

By Lemma 1.11 it follows that $\ker\pi^s_i$ is definable. We obtain a definable subset $U^s_i := O^s_i / \ker\pi^s_i \subseteq A^{\mathrm{eq}}$ and an injective function

$$\pi^s : \textstyle\bigcup_i U^s_i \to B_s .$$

This map is also surjective since its range contains every orbit of B_s under $\operatorname{rng}\varphi$. Setting $\pi := \bigcup_s \pi^s$ we obtain a bijection $\pi : \bigcup_s U_s \to B$. We claim that this bijection is an isomorphism between $\mathfrak{B}$ and a structure of the form $\mathcal{I}(\mathfrak{A})$, for a suitable interpretation $\mathcal{I}$.

If R is a definable relation in $\mathfrak{B}$ then its preimage $\pi^{-1}[R]$ is invariant under automorphisms. Hence, $\pi^{-1}[R]$ is definable in $\mathfrak{A}^{\mathrm{eq}}$. It follows that there exists an interpretation $\mathcal{I}$ such that $\pi : \mathcal{I}(\mathfrak{A}) \cong \mathfrak{B}$.

It remains to check that $\varphi = \operatorname{Aut}\pi \circ \operatorname{Aut}\mathcal{I}$. For every $\sigma \in \mathfrak{G}$ we have

$$\begin{aligned}(\operatorname{Aut}\pi \circ \operatorname{Aut}\mathcal{I})(\sigma)(b^s_i) &= (\pi \circ \mathcal{I}_s \circ \sigma^{\mathrm{eq}} \circ \mathcal{I}_s^{-1} \circ \pi^{-1})(b^s_i) \\ &= (\pi \circ \mathcal{I}_s \circ \sigma^{\mathrm{eq}})(\bar{a}^s_i) \\ &= (\pi \circ \mathcal{I}_s)(\sigma(\bar{a}^s_i)) = \varphi(\sigma)(b^s_i) \qquad \square\end{aligned}$$

Corollary 1.22. *Let Σ and Γ be countable signatures and $\mathcal{I}$ a first-order interpretation from Σ to Γ. If $\mathfrak{A}$ is a countable Σ-structure with $\aleph_0$-categorical theory then the theory of $\mathcal{I}(\mathfrak{A})$ is also $\aleph_0$-categorical.*

Proof. $\operatorname{Aut}\mathcal{I} : \mathfrak{Aut}\,\mathfrak{A} \to \mathfrak{Aut}\,\mathcal{I}(\mathfrak{A})$ is a continuous homomorphism. By the preceding lemma it follows that $\operatorname{rng}(\operatorname{Aut}\mathcal{I})$ is oligomorphic. Since $\operatorname{rng}(\operatorname{Aut}\mathcal{I}) \leq \mathfrak{Aut}\,\mathcal{I}(\mathfrak{A})$ it follows that $\mathfrak{Aut}\,\mathcal{I}(\mathfrak{A})$ is also oligomorphic. $\square$

Definition 1.23. Let $\mathfrak{M}$ be a structure and suppose that $\mathcal{I}$ and $\mathcal{J}$ are interpretations such that there exists an isomorphism $\pi : \mathcal{I}(\mathfrak{M}) \cong \mathcal{J}(\mathfrak{M})$. We call $\mathcal{I}$ and $\mathcal{J}$ *homotopic* (via π) if there exists a definable function $\rho : M \to M$ such that $\pi \circ \mathcal{I} = \mathcal{J} \circ \rho$.

$$\begin{array}{ccc} \mathfrak{M} & \xrightarrow{\ \mathcal{I}\ } & \mathcal{I}(\mathfrak{M}) \\ {\scriptstyle\rho}\downarrow & & \downarrow{\scriptstyle\pi} \\ \mathfrak{M} & \xrightarrow[\ \mathcal{J}\ ]{} & \mathcal{J}(\mathfrak{M}) \end{array}$$

Lemma 1.24. *Let $\mathfrak{M}$ be a countable structure with $\aleph_0$-categorical theory and suppose that $\mathcal{I}$ and $\mathcal{J}$ are interpretations with $\mathcal{I}(\mathfrak{M}) \cong \mathcal{J}(\mathfrak{M})$. Let*

$\pi : \mathcal{I}(\mathfrak{M}) \to \mathcal{J}(\mathfrak{M})$ be an isomorphism. Then $\mathcal{I}$ and $\mathcal{J}$ are homotopic via π if and only if $\operatorname{Aut}\mathcal{J} = \operatorname{Aut}\pi \circ \operatorname{Aut}\mathcal{I}$.

Proof. ($\Rightarrow$) Let $\rho : M \to M$ be a definable function such that $\pi \circ \mathcal{I} = \mathcal{J} \circ \rho$. For every element b of $\mathcal{J}(\mathfrak{M})$ and every automorphism $\sigma \in \operatorname{Aut}\mathfrak{M}$, we have

$$\begin{aligned}
(\operatorname{Aut}\pi \circ \operatorname{Aut}\mathcal{I})(\sigma)(b) &= (\pi \circ \mathcal{I}_s \circ \sigma^{\mathrm{eq}} \circ \mathcal{I}_s^{-1} \circ \pi^{-1})(b) \\
&= (\mathcal{J}_s \circ \rho \circ \sigma^{\mathrm{eq}} \circ \rho^{-1} \circ \mathcal{J}_s^{-1})(b) \\
&= (\mathcal{J}_s \circ \sigma^{\mathrm{eq}} \circ \rho \circ \rho^{-1} \circ \mathcal{J}_s^{-1})(b) \\
&= (\operatorname{Aut}\mathcal{J})(\sigma)(b)
\end{aligned}$$

Hence, $\operatorname{Aut}\pi \circ \operatorname{Aut}\mathcal{I} = \operatorname{Aut}\mathcal{J}$.

($\Leftarrow$) For $a \in M$, we define

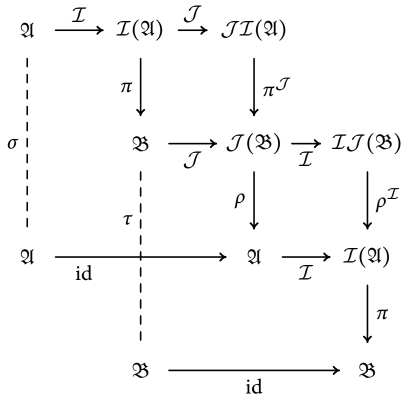
$$\rho(a) := (\mathcal{J}_s^{-1} \circ \pi \circ \mathcal{I}_s)(a)\,.$$

We claim that ρ is definable. For $\sigma \in \operatorname{Aut}\mathfrak{M}$ and $a \in M$, we have

$$\begin{aligned}
\rho(\sigma(a)) &= (\mathcal{J}_s^{-1} \circ \pi \circ \mathcal{I}_s \circ \sigma)(a) \\
&= (\mathcal{J}_s^{-1} \circ \pi \circ \mathcal{I}_s \circ \sigma \circ \mathcal{I}_s^{-1} \circ \pi^{-1} \circ \pi \circ \mathcal{I}_s)(a) \\
&= (\mathcal{J}_s^{-1} \circ (\operatorname{Aut}\pi \circ \operatorname{Aut}\mathcal{I})(\sigma) \circ \pi \circ \mathcal{I}_s)(a) \\
&= (\mathcal{J}_s^{-1} \circ (\operatorname{Aut}\mathcal{J})(\sigma) \circ \pi \circ \mathcal{I}_s)(a) \\
&= (\sigma \circ \mathcal{J}_s^{-1} \circ \pi \circ \mathcal{I}_s)(a) \\
&= \sigma(\rho(a))\,.
\end{aligned}$$

Hence, ρ is invariant under automorphisms and, thus, definable. □

Definition 1.25. Two structures $\mathfrak{A}$ and $\mathfrak{B}$ are *biinterpretable* if there exist first-order interpretations $\mathcal{I}, \mathcal{J}$ and isomorphisms $\pi : \mathcal{I}(\mathfrak{A}) \to \mathfrak{B}$ and $\rho : \mathcal{J}(\mathfrak{B}) \to \mathfrak{A}$ such that $\mathcal{J} \circ \mathcal{I}$ is homotopic to $\mathrm{id}_{\mathfrak{A}}$ via $\rho \circ \pi^{\mathcal{J}}$ and $\mathcal{I} \circ \mathcal{J}$ is homotopic to $\mathrm{id}_{\mathfrak{B}}$ via $\pi \circ \rho^{\mathcal{I}}$.



Theorem 1.26. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be countable models of countable $\aleph_0$-categorical theories. Then $\mathfrak{A}$ and $\mathfrak{B}$ are biinterpretable if and only if $\mathfrak{Aut}\,\mathfrak{A}$ and $\mathfrak{Aut}\,\mathfrak{B}$ are isomorphic as topological groups.*

Proof. ($\Rightarrow$) Let $\mathcal{I}, \mathcal{J}$ and π, ρ witness that $\mathfrak{A}$ and $\mathfrak{B}$ are biinterpretable. There exist definable maps $\sigma : A \to A$ and $\tau : B \to B$ such that

$$\rho \circ \pi^{\mathcal{J}} \circ \mathcal{J} \circ \mathcal{I} = \sigma \quad \text{and} \quad \pi \circ \rho^{\mathcal{I}} \circ \mathcal{I} \circ \mathcal{J} = \tau\,.$$

Set $\varphi := \operatorname{Aut}\pi \circ \operatorname{Aut}\mathcal{I}$ and $\psi := \operatorname{Aut}\rho \circ \operatorname{Aut}\mathcal{J}$. Since σ and τ are definable we have

$$\operatorname{Aut}\sigma = \mathrm{id} \quad \text{and} \quad \operatorname{Aut}\tau = \mathrm{id}\,.$$

It follows that

$$\begin{aligned}
\varphi \circ \psi &= \operatorname{Aut}\rho \circ \operatorname{Aut}\mathcal{J} \circ \operatorname{Aut}\pi \circ \operatorname{Aut}\mathcal{I} \\
&= \operatorname{Aut}(\rho \circ \mathcal{J} \circ \pi \circ \mathcal{I}) \\
&= \operatorname{Aut}(\rho \circ \pi^{\mathcal{J}} \circ \mathcal{J} \circ \mathcal{I}) \\
&= \operatorname{Aut}\sigma \\
&= \mathrm{id}\,,
\end{aligned}$$

and, analogously,

$$\psi \circ \varphi = \mathrm{id}\,.$$

Hence, $\psi = \varphi^{-1}$ and $\varphi : \mathfrak{Aut}\,\mathfrak{A} \to \mathfrak{Aut}\,\mathfrak{B}$ is the desired isomorphism.

($\Leftarrow$) Let $\varphi : \mathfrak{Aut}\,\mathfrak{A} \to \mathfrak{Aut}\,\mathfrak{B}$ be an isomorphism. Since $\operatorname{rng}\varphi = \mathfrak{Aut}\,\mathfrak{B}$ is oligomorphic it follows by Lemma 1.21 that $\varphi = \operatorname{Aut}\pi \circ \operatorname{Aut}\mathcal{I}$, for some interpretation $\mathcal{I}$ and some isomorphism $\pi : \mathcal{I}(\mathfrak{A}) \to \mathfrak{B}$. Similarly, $\operatorname{rng}\varphi^{-1}$ is oligomorphic and we have $\varphi^{-1} = \operatorname{Aut}\rho \circ \operatorname{Aut}\mathcal{J}$, for some $\mathcal{J}$ and ρ. It follows that

$$\begin{aligned}\operatorname{Aut}(\rho \circ \mathcal{J} \circ \pi \circ \mathcal{I}) &= \operatorname{Aut}\rho \circ \operatorname{Aut}\mathcal{J} \circ \operatorname{Aut}\pi \circ \operatorname{Aut}\mathcal{I} \\ &= \varphi^{-1} \circ \varphi = \mathrm{id}\,.\end{aligned}$$

By Lemma 1.24, there exists a definable map $\sigma : A \to A$ such that

$$\pi \circ \mathcal{I} \circ \rho \circ \mathcal{J} = \sigma\,.$$

Analogously, we obtain a definable map $\tau : B \to B$ such that

$$\rho \circ \mathcal{J} \circ \pi \circ \mathcal{I} = \tau\,.$$

Hence, $\mathcal{J} \circ \mathcal{I}$ and id are homotopic via $\rho \circ \pi^{\mathcal{J}}$ and $\mathcal{I} \circ \mathcal{J}$ and id are homotopic via $\pi \circ \rho^{\mathcal{I}}$. ☐

2. *Back-and-forth arguments in accessible categories*

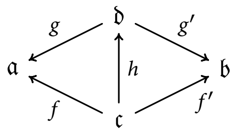
In the next section, we will prove a result about accessible categories using back-and-forth arguments. The necessary machinery for such arguments is developed in the present section. We start by generalising the notion of a partial isomorphism and the forth-property.

Definition 2.1. Let $\mathcal{C}$ be a category, $\mathcal{K} \subseteq \mathcal{C}^{\mathrm{obj}}$ a class of objects, and $\mathfrak{a}, \mathfrak{b} \in \mathcal{C}$.

(a) A *partial morphism* from $\mathfrak{a}$ to $\mathfrak{b}$ is a pair $p = \langle f, f' \rangle$ of morphisms $f : \mathfrak{c} \to \mathfrak{a}$ and $f' : \mathfrak{c} \to \mathfrak{b}$, for some object $\mathfrak{c} \in \mathcal{C}$. We call $\mathfrak{a}$ the *domain* of p, $\mathfrak{b}$ its *codomain,* and $\mathfrak{c}$ is its *base.*

(b) Let $p = \langle f, f' \rangle$ and $q = \langle g, g' \rangle$ be partial morphisms with bases $\mathfrak{c}$ and $\mathfrak{d}$, respectively. A *morphism* $p \to q$ is a morphism $h : \mathfrak{c} \to \mathfrak{d}$ such that

$$f = g \circ h \quad \text{and} \quad f' = g' \circ h\,.$$



(c) We denote by $\mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$ the category of all partial morphisms p from $\mathfrak{a}$ to $\mathfrak{b}$ whose base belongs to $\mathcal{K}$. If $\mathcal{K}$ is the class of all κ-presentable objects, we will write $\mathfrak{pMor}_{\kappa}(\mathfrak{a}, \mathfrak{b})$ instead.

(d) The *domain projection* is the functor

$$P : \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b}) \to \mathfrak{Sub}_{\mathcal{K}}(\mathfrak{a})$$

that maps a partial morphism $p = \langle f, f' \rangle$ to its first component f and a morphism $h : \langle f, f' \rangle \to \langle g, g' \rangle$ of $\mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$ to the underlying morphism $h : f \to g$ of $\mathfrak{Sub}_{\mathcal{K}}(\mathfrak{a})$.

Analogously, the *codomain projection* is the functor

$$Q : \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b}) \to \mathfrak{Sub}_{\mathcal{K}}(\mathfrak{b})$$

mapping $\langle f, f' \rangle$ to f' and $h : \langle f, f' \rangle \to \langle g, g' \rangle$ to $h : f' \to g'$.

Finally, the *base projection* is the functor

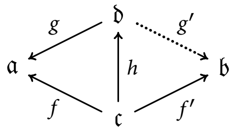
$$B : \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b}) \to \mathcal{C}$$

mapping a partial morphism p to its base and a morphism $h : p \to q$ to the corresponding morphism $h : B(p) \to B(q)$ between the bases.

Definition 2.2. Let $\mathcal{C}$ be a category, $\mathcal{K} \subseteq \mathcal{C}^{\text{obj}}$ a class of objects, and $\mathfrak{a}, \mathfrak{b} \in \mathcal{C}$.

(a) A set I of partial morphisms from $\mathfrak{a}$ to $\mathfrak{b}$ has the *forth property* with respect to $\mathcal{K}$ if, for every $p = \langle f, f' \rangle \in I$ with base $\mathfrak{c}$, every $\mathfrak{d} \in \mathcal{K}$, and every pair of morphisms $g : \mathfrak{d} \to \mathfrak{a}$ and $h : \mathfrak{c} \to \mathfrak{d}$ with $f = g \circ h$, there exists a morphism $g' : \mathfrak{d} \to \mathfrak{b}$ such that $\langle g, g' \rangle \in I$ and

$$h : \langle f, f' \rangle \to \langle g, g' \rangle .$$



(b) We write

$$\mathfrak{a} \sqsubseteq_{\mathcal{K}} \mathfrak{b} \quad :\text{iff} \quad \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b}) \text{ is nonempty and it has the forth property with respect to } \mathcal{K}.$$

Furthermore, we write

$$\mathfrak{a} \sqsubseteq^{\kappa}_{\text{pres}} \mathfrak{b} \quad :\text{iff} \quad \mathfrak{a} \sqsubseteq_{\mathcal{K}_\kappa} \mathfrak{b} ,$$

where $\mathcal{K}_\kappa \subseteq \mathcal{C}$ is the class of all κ-presentable objects. The corresponding equivalence relations are

$$\mathfrak{a} \equiv_{\mathcal{K}} \mathfrak{b} \quad :\text{iff} \quad \mathfrak{a} \sqsubseteq_{\mathcal{K}} \mathfrak{b} \quad \text{and} \quad \mathfrak{b} \sqsubseteq_{\mathcal{K}} \mathfrak{a} ,$$
$$\mathfrak{a} \equiv^{\kappa}_{\text{pres}} \mathfrak{b} \quad :\text{iff} \quad \mathfrak{a} \sqsubseteq^{\kappa}_{\text{pres}} \mathfrak{b} \quad \text{and} \quad \mathfrak{b} \sqsubseteq^{\kappa}_{\text{pres}} \mathfrak{a} .$$

Remark. In the category $\mathfrak{Emb}(\Sigma)$ we have

$$\mathfrak{A} \sqsubseteq^{\kappa}_{\text{pres}} \mathfrak{B} \quad \text{iff} \quad \mathfrak{A} \sqsubseteq^{\kappa}_{0} \mathfrak{B} .$$

Note that, for an arbitrary category, the relation $\sqsubseteq_{\mathcal{K}}$ is not very well-behaved. For instance, in general it is neither reflexive nor transitive. The next lemma collects some basic properties that hold in every category.

Lemma 2.3. *Let $\mathcal{C}$ be a category and $\mathcal{K} \subseteq \mathcal{C}^{\text{obj}}$.*

(a) *If there exists a morphism $\varphi : \mathfrak{a}_0 \to \mathfrak{a}$, then*

$$\mathfrak{a} \sqsubseteq_{\mathcal{K}} \mathfrak{b} \quad \textit{implies} \quad \mathfrak{a}_0 \sqsubseteq_{\mathcal{K}} \mathfrak{b} .$$

(b) *If* $\mathfrak{a} \in \mathcal{K}$ *and* $\mathfrak{a} \sqsubseteq_{\mathcal{K}} \mathfrak{b}$, *then there exists a morphism* $\mathfrak{a} \to \mathfrak{b}$.

(c) *If* $\mathfrak{a}, \mathfrak{b} \in \mathcal{K}$ *and* $\mathfrak{a} \equiv_{\mathcal{K}} \mathfrak{b}$, *then* $\mathfrak{a} \cong \mathfrak{b}$.

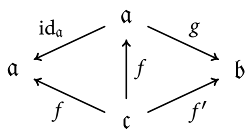
Proof. (a) Let $\langle f, f' \rangle \in \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}_0, \mathfrak{b})$ be a partial morphism with base $\mathfrak{c}$ and let $h : \mathfrak{c} \to \mathfrak{d}$ and $g : \mathfrak{d} \to \mathfrak{a}_0$ be morphisms with $f = g \circ h$ and $\mathfrak{d} \in \mathcal{K}$. Then $\langle \varphi \circ f, f' \rangle \in \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$ and $h : \mathfrak{c} \to \mathfrak{d}$ and $\varphi \circ g : \mathfrak{d} \to \mathfrak{a}$ are morphisms such that $\varphi \circ f = \varphi \circ g \circ h$ and $\mathfrak{d} \in \mathcal{K}$. Consequently, $\mathfrak{a} \sqsubseteq_{\mathcal{K}} \mathfrak{b}$ implies that there exists a morphism $g' : \mathfrak{d} \to \mathfrak{b}$ such that

$$\langle \varphi \circ g, g' \rangle \in \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b}) \quad \text{and} \quad h : \langle \varphi \circ f, f' \rangle \to \langle \varphi \circ g, g' \rangle .$$

It follows that $\langle g, g' \rangle \in \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}_0, \mathfrak{b})$ and $h : \langle f, f' \rangle \to \langle g, g' \rangle$.

(b) As $\mathfrak{a} \sqsubseteq_{\mathcal{K}} \mathfrak{b}$, there exists a partial morphism $\langle f, f' \rangle \in \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$. Since $\mathfrak{a} \in \mathcal{K}$, we can use the forth-property to find a morphism $g : \mathfrak{a} \to \mathfrak{b}$ such that

$$\langle \mathrm{id}_{\mathfrak{a}}, g \rangle \in \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$$
and $$f : \langle f, f' \rangle \to \langle \mathrm{id}_{\mathfrak{a}}, g \rangle .$$

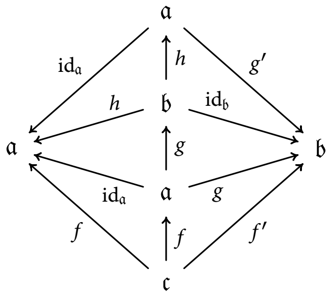


(c) As $a \equiv_{\mathcal{K}} \mathfrak{b}$, there exists a partial morphism $\langle f, f' \rangle \in \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$. As in (b), we can use the forth-property to find a morphism $g : \mathfrak{a} \to \mathfrak{b}$ such that

$$\langle \mathrm{id}_{\mathfrak{a}}, g \rangle \in \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b}) \quad \text{and} \quad f : \langle f, f' \rangle \to \langle \mathrm{id}_{\mathfrak{a}}, g \rangle .$$

Similarly, we can use the back-property to find a morphism $h : \mathfrak{b} \to \mathfrak{a}$ such that

$$\langle h, \mathrm{id}_{\mathfrak{b}} \rangle \in \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$$
and $\quad g : \langle \mathrm{id}_{\mathfrak{a}}, g \rangle \to \langle h, \mathrm{id}_{\mathfrak{b}} \rangle\,.$



Using the forth-property again, we obtain a morphism $g' : \mathfrak{a} \to \mathfrak{b}$ such that

$$\langle \mathrm{id}_{\mathfrak{a}}, g' \rangle \in \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$$
and $\quad h : \langle h, \mathrm{id}_{\mathfrak{b}} \rangle \to \langle \mathrm{id}_{\mathfrak{a}}, g' \rangle\,.$

In particular, $h \circ g = \mathrm{id}_{\mathfrak{a}}$ and $g' \circ h = \mathrm{id}_{\mathfrak{b}}$. By Lemma B1.3.4, it follows that $g = g'$ and $h : \mathfrak{b} \cong \mathfrak{a}$ is an isomorphism. □

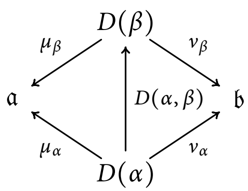
Our goal is to generalise Lemma C4.4.10 to relations of the form $\sqsubseteq_{\mathcal{K}}$. We start with the forth-property.

Proposition 2.4. *Let κ be an infinite cardinal or $\kappa = \infty$, $\mathcal{C}$ a category with colimits of nonempty chains of length less than κ, and let $\mathcal{K} \subseteq \mathcal{C}^{\mathrm{obj}}$ be a class of objects that is closed under colimits of nonempty chains of length less than κ. Let $D : \gamma \to \mathcal{K}$ be a chain of length $0 < \gamma \leq \kappa$ with limiting cocone $\mu \in \mathrm{Cone}(D, \mathfrak{a})$. Suppose that every morphism from some object in $\mathcal{K}$ to $\mathfrak{a}$ factorises essentially uniquely through μ.*

If $\mathfrak{a} \sqsubseteq_{\mathcal{K}} \mathfrak{b}$, then there exists a chain $E : \gamma \to \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$ such that $D = B \circ E$, where B is the base projection functor.

Proof. By induction on $\alpha < \gamma$, we define morphisms $\nu_\alpha : D(\alpha) \to \mathfrak{b}$ such that

$$\langle \mu_\alpha, \nu_\alpha \rangle \in \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$$
and $\quad D(\alpha, \beta) : \langle \mu_\alpha, \nu_\alpha \rangle \to \langle \mu_\beta, \nu_\beta \rangle\,,$



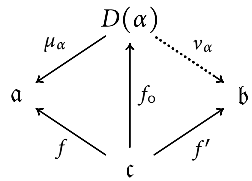
for $\alpha \leq \beta < \gamma$.

Then we can set

$$E(\alpha) := \langle \mu_\alpha, \nu_\alpha \rangle \quad \text{and} \quad E(\alpha, \beta) := D(\alpha, \beta), \quad \text{for } \alpha \le \beta < \gamma.$$

For $\alpha = 0$, we define ν_α as follows. Since $\mathfrak{a} \sqsubseteq_{\mathcal{K}} \mathfrak{b}$, there exists a partial morphism $\langle f, f' \rangle \in \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$. Let $\mathfrak{c}$ be its base. By assumption on D, f factorises as $f = \mu_\alpha \circ f_0$, for some index $\alpha < \gamma$ and some morphism $f_0 : \mathfrak{c} \to D(\alpha)$. As $\mathfrak{a} \sqsubseteq_{\mathcal{K}} \mathfrak{b}$, there exists a morphism $\nu_\alpha : D(\alpha) \to \mathfrak{b}$ such that $\langle \mu_\alpha, \nu_\alpha \rangle \in \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$ and

$$f_0 : \langle f, f' \rangle \to \langle \mu_\alpha, \nu_\alpha \rangle.$$



Setting $\nu_0 := \nu_\alpha \circ D(0, \alpha)$ we obtain the desired morphism $D(0) \to \mathfrak{b}$.

For the inductive step, suppose that we have already defined ν_α for all $\alpha < \beta$. Let λ^β be a limiting cocone from $D \upharpoonright \beta$ to some object $\mathfrak{d}_\beta$. As $\mathcal{K}$ is closed under colimits of chains of length β, we have $\mathfrak{d}_\beta \in \mathcal{K}$. Since $(\mu_\alpha)_{\alpha<\beta}$ and $(\nu_\alpha)_{\alpha<\beta}$ are cocones of $D \upharpoonright \beta$, there exist unique morphisms $\varphi : \mathfrak{d}_\beta \to \mathfrak{a}$ and $\varphi' : \mathfrak{d}_\beta \to \mathfrak{b}$ such that

$$(\mu_\alpha)_{\alpha<\beta} = \varphi * \lambda^\beta \quad \text{and} \quad (\nu_\alpha)_{\alpha<\beta} = \varphi' * \lambda^\beta.$$

Similarly, $(D(\alpha, \beta))_{\alpha<\beta}$ is a cocone from $D \upharpoonright \beta$ to $D(\beta)$ and there exists a unique morphism $\psi : \mathfrak{d}_\beta \to D(\beta)$ such that

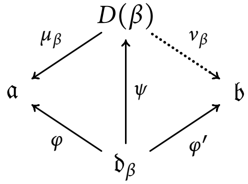
$$(D(\alpha, \beta))_{\alpha<\beta} = \psi * \lambda^\beta.$$

Since

$$\mu_\beta \circ \psi \circ \lambda^\beta_\alpha = \mu_\beta \circ D(\alpha, \beta) = \mu_\alpha = \varphi \circ \lambda^\beta_\alpha, \quad \text{for all } \alpha < \beta,$$

it follows by Lemma B3.4.2 that

$$\mu_\beta \circ \psi = \varphi \,.$$



Therefore, $\mathfrak{a} \sqsubseteq_{\mathcal{K}} \mathfrak{b}$ implies that there exists a morphism $\nu_\beta : D(\beta) \to \mathfrak{b}$ such that

$$\langle \mu_\beta, \nu_\beta \rangle \in \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b}) \quad \text{and} \quad \psi : \langle \varphi, \varphi' \rangle \to \langle \mu_\beta, \nu_\beta \rangle \,.$$

For $\alpha < \beta$ it follows that $D(\alpha, \beta) = \psi \circ \lambda^\beta_\alpha$ is a morphism

$$D(\alpha, \beta) : \langle \mu_\alpha, \nu_\alpha \rangle \to \langle \mu_\beta, \nu_\beta \rangle \,. \qquad \square$$

Proposition 2.5. *Let κ be an infinite cardinal or $\kappa = \infty$, $\mathcal{C}$ a category with colimits of nonempty chains of length at most κ, and let $\mathcal{K} \subseteq \mathcal{C}^{\mathrm{obj}}$ be a class of objects that is closed under colimits of nonempty chains of length less than κ. Let $D : \gamma^D \to \mathcal{K}$ and $E : \gamma^E \to \mathcal{K}$ be chains of length $0 < \gamma^D, \gamma^E \leq \kappa$ with limiting cocones $\lambda^D \in \mathrm{Cone}(D, \mathfrak{a})$ and $\lambda^E \in \mathrm{Cone}(E, \mathfrak{b})$. Suppose that every morphism from some object in $\mathcal{K}$ to $\mathfrak{a}$ or $\mathfrak{b}$ factorises essentially uniquely through, respectively, λ^D and λ^E.*

If $\mathfrak{a} \equiv_{\mathcal{K}} \mathfrak{b}$ and $p \in \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$, there exists a morphism $\varphi : p \to q$ of $\mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$ such that $q = \langle g, g' \rangle$ consists of two epimorphisms.

Proof. By induction on the ordinals γ^D and γ^E, we construct a chain $F : \delta \to \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$, two links s and t from $B \circ F$ to D and E, respectively, and two increasing functions $\rho_0 : \gamma^D \to \delta$ and $\theta_0 : \gamma^E \to \delta$ such that

$$\begin{aligned} B(F(\alpha)) &= D(\rho(\alpha))\,, \qquad & s_\alpha &= \mathrm{id}_{D(\rho(\alpha))}\,, \qquad & &\text{for } \alpha \in \operatorname{rng} \rho_0\,, \\ B(F(\alpha)) &= E(\theta(\alpha))\,, \qquad & t_\alpha &= \mathrm{id}_{E(\theta(\alpha))}\,, \qquad & &\text{for } \alpha \in \operatorname{rng} \theta_0\,, \end{aligned}$$

where B is the base projection functor and ρ and θ are the index maps of s and t, respectively.

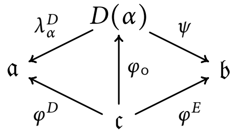
For $\gamma^D, \gamma^E = 0$, we start with $\delta := 1$ and $F(0) := p$. To define s and t, suppose that $p = \langle f, f' \rangle$. By assumption, f and f' factorise essentially uniquely through λ^D and λ^E, respectively. Let $f = \lambda^D_\alpha \circ f_0$ and $f' = \lambda^E_\beta \circ f'_0$ be the corresponding factorisations. We set $s_0 := f_0$ and $t_0 := f'_0$.

For the inductive step, suppose that, for the restrictions $D \restriction \beta^D$ and $E \restriction \beta^E$, we have already defined a chain $F : \delta \to \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$ with $0 < \delta < \kappa$, links s and t from $B \circ F$ to $D \restriction \beta^D$ and $E \restriction \beta^E$, respectively, and increasing functions $\rho_0 : \beta^D \to \delta$ and $\theta_0 : \beta^E \to \delta$.

We will show how to extend these definitions to $D \restriction \beta^D + 1$. (The extension to $E \restriction \beta^E + 1$ works in the same way.) Let μ be a limiting cocone from $B \circ F$ to some object $\mathfrak{c}$. As $\mathcal{K}$ is closed under limits of chains of length $0 < \delta < \kappa$, it follows that $\mathfrak{c} \in \mathcal{K}$. Since $\lambda^D * s$ is a cocone of $B \circ F$, there exists a unique morphism $\varphi^D : \mathfrak{c} \to \mathfrak{a}$ such that $\lambda^D * s = \varphi^D * \mu$. In the same way, we obtain a unique morphism $\varphi^E : \mathfrak{c} \to \mathfrak{b}$ with $\lambda^E * t = \varphi^E * \mu$.

As $\mathfrak{c} \in \mathcal{K}$, there exists an essentially unique factorisation $\varphi^D = \lambda^D_\alpha \circ \varphi_0$, for some morphism $\varphi_0 : \mathfrak{c} \to D(\alpha)$ with $\alpha \geq \beta^D$. Since $\mathfrak{a} \sqsubseteq_{\mathcal{K}} \mathfrak{b}$, we can find a morphism $\psi : \mathfrak{c} \to \mathfrak{b}$ such that

$$\psi \circ \varphi_0 = \varphi^E .$$



As $D(\alpha) \in \mathcal{K}$, there exists an essentially unique factorisation $\psi = \lambda^E_\beta \circ \psi_0$, for some morphism $\psi_0 : D(\alpha) \to E(\beta)$ with $\beta \geq \beta^D$. We set

$$\begin{aligned} F(\delta) &:= \langle \lambda^D_\alpha, \psi \rangle , & F(i, \delta) &:= \varphi_0 \circ \mu_i , \quad \text{for } i < \delta , \\ s_\delta &:= \mathrm{id}_{D(\alpha)} , & \rho_0(\beta^D) &:= \alpha , \\ t_\delta &:= \psi_0 . \end{aligned}$$

Let us show that these morphisms have the desired properties. First, we check that the extension of s is a link from the extension of $B \circ F$ to D. For every $i < \delta$, it follows by choice of φ^D that

$$\lambda^D_\alpha \circ D(\rho(i), \alpha) \circ s_i = \lambda^D_{\rho(i)} \circ s_i = \varphi^D \circ \mu_i = \lambda^D_\alpha \circ \varphi_0 \circ \mu_i .$$

Since $B(F(i)) \in \mathcal{K}$, this morphism has an essentially unique factorisation through λ^D. Hence, the above two factorisations are a.p.-equivalent.

$$D(\varphi(i), \alpha) \circ s_i \mathrel{\wedge\!\!\!\wedge}_D \varphi_0 \circ \mu_i .$$

By Lemma B3.5.3 (d), this implies that

$$s_i \mathrel{\wedge\!\!\!\wedge}_D \varphi_0 \circ \mu_i = s_\delta \circ F(i, \delta) ,$$

as desired.

We also have to check that the extension of t is a link. Let $i < \delta$. Then

$$\begin{aligned} \lambda^E_\beta \circ t_\delta \circ F(i, \delta) &= \lambda^E_\beta \circ \psi_0 \circ \varphi_0 \circ \mu_i \\ &= \psi \circ \varphi_0 \circ \mu_i = \varphi^E \circ \mu_i = \lambda^E_{\theta(i)} \circ t_i . \end{aligned}$$

Since $B(F(i)) \in \mathcal{K}$, this morphism has an essentially unique factorisation through λ^E. Hence, the above two factorisations are a.p.-equivalent.

$$t_\delta \circ F(i, \delta) \mathrel{\wedge\!\!\!\wedge}_E t_i ,$$

as desired.

Having defined $F : \delta \to \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$, we construct the desired partial morphism $q = \langle g, g' \rangle \in \mathfrak{pMor}(\mathfrak{a}, \mathfrak{b})$ as follows. Let λ^F be a limiting cocone from $B \circ F$ to some object $\mathfrak{c} \in \mathcal{C}$. Since $\lambda^D * s$ and $\lambda^E * t$ are cocones of F, there exist unique morphisms $g : \mathfrak{c} \to \mathfrak{a}$ and $g' : \mathfrak{c} \to \mathfrak{b}$ such that $\lambda^D * s = g * \lambda^F$ and $\lambda^E * t = g' * \lambda^F$. We claim that g and g' are epimorphisms. By symmetry, it is sufficient to give a proof for g. Hence, let $h, h' : \mathfrak{a} \to \mathfrak{d}$ be morphisms such that $h \circ g = h' \circ g$. For every $i < \gamma^D$,

it follows that

$$\begin{aligned}
h \circ \lambda^D_i &= h \circ \lambda^D_{\rho(\rho_0(i))} \circ D(i, \rho(\rho_0(i))) \\
&= h \circ \lambda^D_{\rho(\rho_0(i))} \circ s_{\rho_0(i)} \circ D(i, \rho(\rho_0(i))) \\
&= h \circ g \circ \lambda^F_{\rho_0(i)} \circ D(i, \rho(\rho_0(i))) \\
&= h' \circ g \circ \lambda^F_{\rho_0(i)} \circ D(i, \rho(\rho_0(i))) \\
&= h' \circ \lambda^D_{\rho(\rho_0(i))} \circ s_{\rho_0(i)} \circ D(i, \rho(\rho_0(i))) \\
&= h' \circ \lambda^D_{\rho(\rho_0(i))} \circ D(i, \rho(\rho_0(i))) \\
&= h' \circ \lambda^D_i \,.
\end{aligned}$$

Consequently, Lemma B3.4.2 implies that $h = h'$.

Finally, note that $\lambda^F_0 : B(F(0)) \to \mathfrak{c}$ is the desired morphism $p \to q$ since, by choice of g, g', s_0, t_0, we have

$$g \circ \lambda^F_0 = \lambda^D_{\rho(0)} \circ s_0 = f \quad \text{and} \quad g' \circ \lambda^F_0 = \lambda^E_{\theta(0)} \circ t_0 = f' \,. \qquad \square$$

The preceding two results are phrased in a quite general form. Their statements can be simplified significantly if we assume that the category is $\aleph_0$-accessible, all morphisms are monomorphisms, and all epimorphisms are isomorphisms. Since in the applications below we will mainly be working in $\mathfrak{Emb}(\Sigma)$ and similar categories where these assumptions are met, we record here the corresponding simplified versions. We start by proving that, under these assumptions, every object can be written as the colimit of a chain.

Lemma 2.6. *Let $\mathcal{C}$ be a category where every morphism is a monomorphism. For every κ-filtered diagram $D : \mathcal{I} \to \mathcal{C}$ of size λ that has a colimit, there exists a κ-directed diagram $E : \mathfrak{K} \to \mathcal{C}$ of size at most λ with*

$$\varinjlim E = \varinjlim D \quad \textit{and} \quad \operatorname{rng} E^{\mathrm{obj}} = \operatorname{rng} D^{\mathrm{obj}}.$$

Proof. Fix a limiting cocone $\mu \in \operatorname{Cone}(D, \mathfrak{a})$. For the index order $\mathfrak{K}$ of the diagram E, we choose the set $K := \mathcal{I}^{\mathrm{obj}}$ where we define the order by

$$\mathfrak{i} \leq \mathfrak{j} \quad :\text{iff} \quad \mathcal{I}(\mathfrak{i}, \mathfrak{j}) \neq \varnothing \,.$$

Since $\mathcal{I}$ is κ-filtered, this preorder is clearly κ-directed. We define the diagram E by setting

$$E^{\text{obj}}(\mathfrak{i}) := D(\mathfrak{i})$$

$$\text{and} \quad E^{\text{mor}}(\mathfrak{i},\mathfrak{j}) := D(f)\,, \quad \text{for an arbitrary } f \in \mathcal{I}(\mathfrak{i},\mathfrak{j})\,.$$

First, note that E is well-defined in the sense that the value of $E(\mathfrak{i},\mathfrak{j})$ does not depend on the choice of f: if $f, f' \in \mathcal{I}(\mathfrak{i},\mathfrak{j})$, then

$$\mu_{\mathfrak{j}} \circ D(f) = \mu_{\mathfrak{i}} = \mu_{\mathfrak{j}} \circ D(f') \quad \text{implies} \quad D(f) = D(f')\,,$$

as $\mu_{\mathfrak{j}}$ is a monomorphism. Furthermore, it follows immediately from the definition that $\operatorname{rng} E^{\text{obj}} = \operatorname{rng} D^{\text{obj}}$.

Hence, it remains to show that D and E have the same colimit. We will prove below that $\operatorname{Cone}(E,\mathfrak{b}) = \operatorname{Cone}(D,\mathfrak{b})$, for every $\mathfrak{b} \in \mathcal{C}$. Hence, the identity maps provide a natural isomorphism

$$\operatorname{id} : \operatorname{Cone}(D,-) \to \operatorname{Cone}(E,-)$$

and it follows by Lemma B3.4.3 that D and E have the same colimits.

To prove the claim, let $\nu \in \operatorname{Cone}(D,\mathfrak{b})$. For all $\mathfrak{i} \leq \mathfrak{j}$ and $f \in \mathcal{I}(\mathfrak{i},\mathfrak{j})$, it follows that

$$\nu_{\mathfrak{i}} = \nu_{\mathfrak{j}} \circ D(f) = \nu_{\mathfrak{j}} \circ E(\mathfrak{i},\mathfrak{j})\,.$$

Hence, $\nu \in \operatorname{Cone}(E,\mathfrak{b})$. Conversely, let $\nu \in \operatorname{Cone}(E,\mathfrak{b})$. For all $f : \mathfrak{i} \to \mathfrak{j}$ in $\mathcal{I}$, it follows that

$$\nu_{\mathfrak{i}} = \nu_{\mathfrak{j}} \circ E(\mathfrak{i},\mathfrak{j}) = \nu_{\mathfrak{j}} \circ D(f)\,.$$

Hence, $\nu \in \operatorname{Cone}(D,\mathfrak{b})$. □

Corollary 2.7. *Let $\mathcal{C}$ be an $\aleph_0$-accessible category where every morphism is a monomorphism. For every κ^+-presentable object $\mathfrak{a} \in \mathcal{C}$, there exists a chain $D : \kappa \to \mathcal{C}$ such that*

- $\varinjlim D = \mathfrak{a}$,
- *every object $D(\alpha)$ is κ-presentable and,*
- *for each κ-presentable object $\mathfrak{b}$, every morphism $f : \mathfrak{b} \to \mathfrak{a}$ factorises essentially uniquely through every limiting cocone from D to $\mathfrak{a}$.*

Proof. If $\mathfrak{a}$ is κ-presentable, we can take the constant diagram $D : \kappa \to \mathcal{C}$ where $D(\alpha) = \mathfrak{a}$ and $D(\alpha, \beta) = \mathrm{id}_{\mathfrak{a}}$, for all $\alpha \leq \beta < \kappa$. Hence, it remains to consider the case where $\mathfrak{a}$ is κ^+-presentable, but not κ-presentable. Then we can use Theorem B4.4.3 to find an $\aleph_0$-filtered diagram $E : \mathcal{I} \to \mathcal{C}$ of size at most κ with colimit $\mathfrak{a}$ such that every object $E(\mathfrak{i})$ is $\aleph_0$-presentable. We use Lemma 2.6 to construct a $\aleph_0$-directed diagram $F : \mathfrak{K} \to \mathcal{C}$ of size at most κ with $\varinjlim F = \mathfrak{a}$ such that every $F(\mathfrak{i})$ is $\aleph_0$-presentable. By Proposition B3.4.16, there exists a chain $D : \gamma \to \mathcal{C}$ of length $\gamma \leq |K| \leq \kappa$ with colimit $\mathfrak{a}$ such that each object $D(\alpha)$ is a colimit of a directed diagram of size less than $|K|$. In particular, every $D(\alpha)$ is κ-presentable. As $\mathfrak{a}$ is not κ-presentable, it follows by Theorem B4.4.3 that $\gamma = \kappa$.

Finally, let $\mu \in \mathrm{Cone}(D, \mathfrak{a})$ be limiting. If κ is regular, the index order $\langle \kappa, \leq \rangle$ of D is κ-directed and every morphism $f : \mathfrak{b} \to \mathfrak{a}$ from a κ-presentable object $\mathfrak{b}$ to $\mathfrak{a}$ factorises essentially uniquely through μ. Hence, suppose that κ is singular. Then it follows by Lemma B4.1.4 that an object is κ-presentable if, and only if, it is κ^+-presentable. This contradicts our assumption that $\mathfrak{a}$ is κ^+-presentable but not κ-presentable. □

In the following theorem let us state the special cases of Propositions 2.4 and 2.5 that we will need below.

Theorem 2.8. *Let $\mathcal{C}$ be an $\aleph_0$-accessible category where every morphism is a monomorphism and every epimorphism is an isomorphism.*

(a) *If $\mathfrak{a} \in \mathcal{C}$ is κ^+-presentable and $\mathfrak{a} \sqsubseteq^{\kappa}_{\mathrm{pres}} \mathfrak{b}$, then there exists a morphism $f : \mathfrak{a} \to \mathfrak{b}$.*

(b) *Let $\mathfrak{a}, \mathfrak{b} \in \mathcal{C}$ be κ^+-presentable objects with $\mathfrak{a} \equiv^{\kappa}_{\mathrm{pres}} \mathfrak{b}$. For every partial morphism $p = \langle f, f' \rangle \in \mathfrak{pMor}_{\kappa}(\mathfrak{a}, \mathfrak{b})$, there exists an isomorphism $\pi : \mathfrak{a} \to \mathfrak{b}$ with $f' = \pi \circ f$.*

Proof. We start by proving that $\mathcal{C}$ and the class $\mathcal{K}$ of all κ-presentable objects satisfy the conditions of Propositions 2.4 and 2.5. Clearly, being $\aleph_0$-accessible $\mathcal{C}$ has colimits of chains.

To show that $\mathcal{K}$ is closed under colimits of nonempty chains of length less than κ, let $F : \gamma \to \mathcal{K}$ be such a chain. As every object $F(i)$, for $i < \gamma$, is κ-presentable, it follows by Proposition B4.3.7 that the colimit of F is $(\kappa \oplus |\gamma|^+)$-presentable, i.e., κ-presentable.

(a) We can use Corollary 2.7 to express $\mathfrak{a}$ as the colimit of a chain $D : \kappa \to \mathcal{K}$ of the form required by Proposition 2.4. Consequently, we obtain a diagram $F : \kappa \to \mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$ such that $D = B \circ F$. Let λ be a limiting cocone from D to $\mathfrak{a}$ and set $\mu_\alpha := Q(F(\alpha))$, for $\alpha < \kappa$, where Q is the codomain projection functor. Then $\mu := (\mu_\alpha)_{\alpha<\kappa}$ is a cocone from D to $\mathfrak{b}$. As λ is limiting, there exists a morphism $f : \mathfrak{a} \to \mathfrak{b}$ such that $\mu = f * \lambda$.

(b) We can use Corollary 2.7 to express $\mathfrak{a}$ and $\mathfrak{b}$ as colimits of chains $D : \kappa \to \mathcal{K}$ and $E : \kappa \to \mathcal{K}$ of the form required by Proposition 2.5. Therefore, we obtain a morphism $h : p \to q$ of $\mathfrak{pMor}_{\mathcal{K}}(\mathfrak{a}, \mathfrak{b})$ where $q = \langle g, g' \rangle$ consists of two isomorphisms. It follows that $\pi := g' \circ g^{-1}$ is the desired isomorphism between $\mathfrak{a}$ and $\mathfrak{b}$. □

3. *Fraïssé limits*

In this section we will present a method to construct structures with an $\aleph_0$-categorical theory. These structures will be approximated by a directed diagram of finitely generated substructures. Since this construction has further applications, we will present it in the general setting of an accessible category.

Ultrahomogeneous objects

As in the case of κ-saturated structures and atomic ones, we can characterise the maximal objects of the order $\sqsubseteq^{\kappa}_{\mathrm{pres}}$. For the category $\mathfrak{Emb}(\Sigma)$, these structures will have an $\aleph_0$-categorical theory.

Definition 3.1. Let $\mathcal{C}$ be a category. An object $\mathfrak{u} \in \mathcal{C}$ is *κ-ultrahomogeneous* if, for every κ-presentable object $\mathfrak{a}$ and all pairs of morphisms $f, f' : \mathfrak{a} \to \mathfrak{u}$, there exists an automorphism $\pi : \mathfrak{u} \to \mathfrak{u}$ with $f' = \pi \circ f$.

We call an object $\mathfrak{u}$ *ultrahomogeneous* if it is $\|\mathfrak{u}\|$-ultrahomogeneous.

Example. (a) The order $\langle \mathbb{Q}, \leq \rangle$ of the rationals is ultrahomogeneous in $\mathfrak{Emb}(\leq)$.

(b) Let $\langle \omega, p \rangle$ be the structure where $p(0) := 0$ and $p(n+1) := n$. This structure is ultrahomogeneous in $\mathfrak{Emb}(p)$ since no two distinct substructures are isomorphic.

(c) We have shown in Corollary B6.5.31 that algebraically closed fields are $\aleph_0$-ultrahomogeneous.

Exercise 3.1. Find a dense linear order that is not $\aleph_0$-ultrahomogeneous in $\mathfrak{Emb}(\leq)$. Can you find an open one?

One important parameter of an ultrahomogeneous structure is the class of its substructures.

Definition 3.2. Let $\mathcal{C}$ be a category, κ an infinite cardinal, and $\mathfrak{a} \in \mathcal{C}$. We denote by $\mathrm{Sub}_\kappa(\mathfrak{a})$ the class of all κ-presentable objects $\mathfrak{c} \in \mathcal{C}$ such that there exists a morphism $\mathfrak{c} \to \mathfrak{a}$.

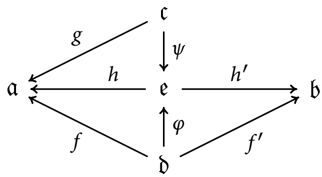
For accessible categories this class is well-behaved.

Lemma 3.3. *Let $\mathcal{C}$ be a κ-accessible category.*

$$\mathfrak{a} \sqsubseteq^\kappa_{\mathrm{pres}} \mathfrak{b} \quad \textit{implies} \quad \mathrm{Sub}_\kappa(\mathfrak{a}) \subseteq \mathrm{Sub}_\kappa(\mathfrak{b})\,.$$

Proof. Let $\mathfrak{c} \in \mathrm{Sub}_\kappa(\mathfrak{a})$ and let $g : \mathfrak{c} \to \mathfrak{a}$ be a corresponding morphism. Since $\mathfrak{a} \sqsubseteq^\kappa_{\mathrm{pres}} \mathfrak{b}$, there exists a partial morphism $\langle f, f' \rangle \in \mathfrak{pMor}_\kappa(\mathfrak{a}, \mathfrak{b})$. According to Proposition B4.4.12, the category $\mathfrak{Sub}_\kappa(\mathfrak{a})$ is κ-filtered. Therefore, there exist an object $h : \mathfrak{e} \to \mathfrak{a}$ of $\mathfrak{Sub}_\kappa(\mathfrak{a})$ and morphisms $\varphi : f \to h$ and $\psi : g \to h$. Since $\mathfrak{a} \sqsubseteq^\kappa_{\mathrm{pres}} \mathfrak{b}$, we can find a morphism $h' : \mathfrak{e} \to \mathfrak{b}$ such that $\langle h, h' \rangle \in \mathfrak{pMor}_\kappa(\mathfrak{a}, \mathfrak{b})$ and

$$\varphi : \langle f, f' \rangle \to \langle h, h' \rangle .$$



We obtain a morphism $h' \circ \psi : \mathfrak{c} \to \mathfrak{b}$ witnessing the fact that $\mathfrak{c} \in \mathrm{Sub}_\kappa(\mathfrak{b})$. ☐

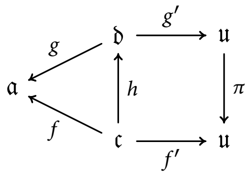
Corollary 3.4. *Let $\mathcal{C}$ be an $\aleph_0$-accessible category where every morphism is a monomorphism, and let $\mathfrak{u}$ be κ-ultrahomogeneous. Then*

$$\mathfrak{a} \sqsubseteq^\kappa_{\mathrm{pres}} \mathfrak{u} \quad \text{iff} \quad \mathrm{Sub}_\kappa(\mathfrak{a}) \subseteq \mathrm{Sub}_\kappa(\mathfrak{u}) , \quad \textit{for all objects } \mathfrak{a} .$$

Proof. ($\Rightarrow$) Since $\aleph_0$-accessible categories are κ-accessible, for all infinite cardinals κ, this direction follows from Lemma 3.3.

($\Leftarrow$) Let $p = \langle f, f' \rangle \in \mathfrak{pMor}_\kappa(\mathfrak{a}, \mathfrak{u})$ be a partial morphism with base $\mathfrak{c}$ and let $h : \mathfrak{c} \to \mathfrak{d}$ and $g : \mathfrak{d} \to \mathfrak{a}$ be morphisms with $g \circ h = f$ where $\mathfrak{d}$ is κ-presentable. Since $\mathfrak{d} \in \mathrm{Sub}_\kappa(\mathfrak{a}) \subseteq \mathrm{Sub}_\kappa(\mathfrak{u})$, there exists some morphism $g' : \mathfrak{d} \to \mathfrak{u}$. As $\mathfrak{u}$ is κ-ultrahomogeneous, we can find an automorphism $\pi : \mathfrak{u} \to \mathfrak{u}$ such that

$$f' = \pi \circ g' \circ h .$$



We obtain a partial morphism $q := \langle g, \pi \circ g' \rangle \in \mathfrak{pMor}_\kappa(\mathfrak{a}, \mathfrak{u})$ such that $h : p \to q$. ☐

The statement of the previous corollary can be used to characterise ultrahomogeneous objects.

Proposition 3.5. *Let $\mathcal{C}$ be an $\aleph_0$-accessible category where every morphism is a monomorphism and every epimorphism an isomorphism. For a κ^+-presentable object $\mathfrak{u} \in \mathcal{C}$, the following statements are equivalent:*

(1) $\mathfrak{u}$ *is κ-ultrahomogeneous.*

(2) $\mathfrak{a} \sqsubseteq^{\kappa}_{\text{pres}} \mathfrak{u}$, *for all* $\mathfrak{a} \in \mathcal{C}$ *with* $\operatorname{Sub}_\kappa(\mathfrak{a}) \subseteq \operatorname{Sub}_\kappa(\mathfrak{u})$.

(3) $\mathfrak{u} \equiv^{\kappa}_{\text{pres}} \mathfrak{u}$

Proof. (1) ⇒ (2) was already proved in Corollary 3.4 and (2) ⇒ (3) is trivial. Hence, it remains to prove (3) ⇒ (1). To show that $\mathfrak{u}$ is κ-ultrahomogeneous, consider morphisms $f, f' : \mathfrak{c} \to \mathfrak{u}$ with κ-presentable domain $\mathfrak{c}$. By assumption, we have $\mathfrak{u} \equiv^{\kappa}_{\text{pres}} \mathfrak{u}$. Consequently, we can use Theorem 2.8 (b) to find an isomorphism $\pi : \mathfrak{u} \to \mathfrak{u}$ such that $f' = \pi \circ f$. □

Corollary 3.6. *Let $\mathcal{C}$ be an $\aleph_0$-accessible category where every morphism is a monomorphism and every epimorphism an isomorphism.*

(a) *Let $\mathfrak{u}, \mathfrak{v}$ be κ^+-presentable κ-ultrahomogeneous objects. Then*

$$\operatorname{Sub}_\kappa(\mathfrak{u}) = \operatorname{Sub}_\kappa(\mathfrak{v}) \quad \textit{implies} \quad \mathfrak{u} \cong \mathfrak{v}\,.$$

(b) *Let $\mathfrak{u}$ be κ-ultrahomogeneous and $\mathfrak{a}$ κ^+-presentable. Then*

$$\operatorname{Sub}_\kappa(\mathfrak{a}) \subseteq \operatorname{Sub}_\kappa(\mathfrak{u}) \quad \textit{implies} \quad \mathfrak{a} \in \operatorname{Sub}_{\kappa^+}(\mathfrak{u})\,.$$

Proof. (a) This follows by Theorem 2.8 (b) and Proposition 3.5.

(b) By Corollary 3.4, $\operatorname{Sub}_\kappa(\mathfrak{a}) \subseteq \operatorname{Sub}_\kappa(\mathfrak{b})$ implies $\mathfrak{a} \sqsubseteq^{\kappa}_{\text{pres}} \mathfrak{u}$. Hence, the claim follows by Theorem 2.8 (a). □

We have claimed above that ultrahomogeneous structures in $\mathfrak{Emb}(\Sigma)$ have an $\aleph_0$-categorical theory. We start by showing that they are existentially closed.

Proposition 3.7. *Let $\mathfrak{U}$ be an $\aleph_0$-ultrahomogeneous structure in $\mathfrak{Emb}(\Sigma)$. Then $\mathfrak{U}$ is existentially closed in the class*

$$\mathcal{C} := \{\, \mathfrak{M} \in \operatorname{Str}[\Sigma] \mid \operatorname{Sub}_{\aleph_0}(\mathfrak{M}) \subseteq \operatorname{Sub}_{\aleph_0}(\mathfrak{U}) \,\}\,.$$

Proof. Suppose that $\mathfrak{U} \subseteq \mathfrak{M}$ for some structure $\mathfrak{M} \in \mathcal{C}$. Let $\varphi(\bar{x}, \bar{y})$ be a quantifier-free formula and $\bar{a} \subseteq U$ parameters such that

$$\mathfrak{M} \vDash \exists \bar{y} \varphi(\bar{a}, \bar{y}) .$$

We have to show that $\mathfrak{U} \vDash \exists \bar{y} \varphi(\bar{a}, \bar{y})$. Fix a tuple $\bar{b} \subseteq M$ with $\mathfrak{M} \vDash \varphi(\bar{a}, \bar{b})$. By Corollary 3.6 (b), there exists an embedding $h : \langle\!\langle \bar{a}\bar{b} \rangle\!\rangle_{\mathfrak{M}} \to \mathfrak{U}$. Since $\mathfrak{U}$ is $\aleph_0$-ultrahomogeneous and

$$\langle\!\langle \bar{a} \rangle\!\rangle_{\mathfrak{U}} \cong \langle\!\langle h(\bar{a}) \rangle\!\rangle_{\mathfrak{U}}$$

we can find an automorphism π of $\mathfrak{U}$ with $\pi(h(\bar{a})) = \bar{a}$. Consequently,

$$\begin{aligned} \mathfrak{M} \vDash \varphi(\bar{a}, \bar{b}) \quad &\text{iff} \quad \langle\!\langle \bar{a}\bar{b} \rangle\!\rangle_{\mathfrak{M}} \vDash \varphi(\bar{a}, \bar{b}) \\ &\text{iff} \quad \mathfrak{U} \vDash \varphi(h(\bar{a}), h(\bar{b})) \\ &\text{iff} \quad \mathfrak{U} \vDash \varphi(\bar{a}, \pi(h(\bar{b}))) . \end{aligned}$$

Hence, $\mathfrak{U} \vDash \exists \bar{y} \varphi(\bar{a}, \bar{y})$. □

With slightly stronger assumptions we obtain $\aleph_0$-categoricity.

Proposition 3.8. *Let Σ be a finite relational signature and let $\mathfrak{U}$ be a countable ultrahomogeneous structure in* $\mathfrak{Emb}(\Sigma)$. *Then* $\mathrm{Th}(\mathfrak{U})$ *is $\aleph_0$-categorical.*

Proof. Note that, for every finite tuple $\bar{s}$ of sorts, there are only finitely many substructures $\langle\!\langle \bar{a} \rangle\!\rangle_{\mathfrak{U}}$ of $\mathfrak{U}$ that are generated by a tuple $\bar{a} \in U^{\bar{s}}$ of sort $\bar{s}$. As $\mathfrak{U}$ is $\aleph_0$-ultrahomogeneous, it follows that any isomorphism between two such substructures extends to an isomorphism of $\mathfrak{U}$. Consequently, the automorphism group of $\mathfrak{U}$ is oligomorphic and it follows by Theorem 1.2 that $\mathrm{Th}(\mathfrak{U})$ is $\aleph_0$-categorical. □

Example. (a) We have seen above that $\langle \mathbb{Q}, \leq \rangle$ is $\aleph_0$-ultrahomogeneous. Consequently, it follows by the proposition that $\mathrm{Th}(\mathbb{Q}, \leq)$ is $\aleph_0$-categorical.

(b) That the restriction on the signature Σ is necessary, is shown by the example $\langle \omega, p \rangle$. We have seen above that this structures is $\aleph_0$-ultrahomogeneous, but its theory is not $\aleph_0$-categorical.

The theorems of Fraïssé

We have seen in Corollary 3.6 (a) that an ultrahomogeneous object $\mathfrak{u}$ is uniquely determined by the class $\mathrm{Sub}_\kappa(\mathfrak{u})$. Therefore it is worthwhile to characterise such classes. In the present section we will provide a characterisation in terms of the following properties.

Definition 3.9. Let $\mathcal{C}$ be a category, κ a cardinal, and $\mathcal{K} \subseteq \mathcal{C}$.

(a) The class $\mathcal{K}$ is *κ-hereditary* if

$$\mathfrak{a} \in \mathcal{K} \quad \text{implies} \quad \mathrm{Sub}_\kappa(\mathfrak{a}) \subseteq \mathcal{K} .$$

We call $\mathcal{K}$ *hereditary* if it is κ-hereditary, for all cardinals κ.

(b) $\mathcal{K}$ has the *κ-joint embedding property* if, for every set $X \subseteq \mathcal{K}$ of size $|X| < \kappa$, there exist an object $\mathfrak{c} \in \mathcal{K}$ and morphisms $\mathfrak{a} \to \mathfrak{c}$, for each $\mathfrak{a} \in X$.

(c) $\mathcal{K}$ has the *κ-amalgamation property* if, for every family of morphisms $f_i : \mathfrak{a} \to \mathfrak{b}_i$, $i < \gamma$, with $\mathfrak{a}, \mathfrak{b}_i \in \mathcal{K}$ and $\gamma < \kappa$, there exist an object $\mathfrak{c} \in \mathcal{K}$ and morphisms $g_i : \mathfrak{b}_i \to \mathfrak{c}$, $i < \gamma$, such that

$$g_i \circ f_i = g_k \circ f_k , \quad \text{for all } i, k < \gamma .$$

Remark. If the subcategory of $\mathcal{C}$ induced by a class $\mathcal{K} \subseteq \mathcal{C}^{\mathrm{obj}}$ is κ-filtered, then Condition (F1) states that $\mathcal{K}$ has the κ-joint embedding property, and Lemma B4.1.2 implies that $\mathcal{K}$ has the κ-amalgamation property.

The converse fails in general. For instance, consider the class $\mathcal{K} \subseteq \mathfrak{Emb}(\Sigma)$ of all finitely generated structures. This class has the $\aleph_0$-joint embedding property and the $\aleph_0$-amalgamation property, but it is not $\aleph_0$-filtered: take finitely generated structures $\mathfrak{A}, \mathfrak{B} \in \mathcal{K}$ such that there are two different embeddings $f, g : \mathfrak{A} \to \mathfrak{B}$. Then $h \circ f \neq h \circ g$, for every embedding h.

Exercise 3.2. For a suitable signature Σ, find a class $\mathcal{K} \subseteq \mathfrak{Emb}(\Sigma)$ with the $\aleph_0$-amalgamation property that does not have the $\aleph_0$-joint embedding property.

Exercise 3.3. Suppose that the class $\mathcal{K}$ is closed under unions of chains of length less than κ. Prove that, if $\mathcal{K}$ has the $\aleph_0$-joint embedding property, it also has the κ-joint embedding property and that, if it has the $\aleph_0$-amalgamation property, it has the κ-amalgamation property.

Before giving a characterisation of classes of the form $\mathrm{Sub}_\kappa(\mathfrak{a})$, we start with a technical remark on such classes for κ-filtered colimits.

Lemma 3.10. *Let* $\mathfrak{a}$ *be the colimit of a* κ*-filtered diagram* $D : \mathcal{I} \to \mathcal{C}$*. Then*

$$\mathrm{Sub}_\kappa(\mathfrak{a}) = \bigcup_{\mathfrak{i}\in\mathcal{I}} \mathrm{Sub}_\kappa(D(\mathfrak{i}))\,.$$

Proof. Let $\lambda \in \mathrm{Cone}(D, \mathfrak{a})$ be a limiting cocone.

($\supseteq$) For every $\mathfrak{b} \in \mathrm{Sub}_\kappa(D(\mathfrak{i}))$, there is some morphism $f : \mathfrak{b} \to D(\mathfrak{i})$. Hence, $\lambda_\mathfrak{i} \circ f$ is a morphism $\mathfrak{b} \to \mathfrak{a}$.

($\subseteq$) Let $\mathfrak{b} \in \mathrm{Sub}_\kappa(\mathfrak{a})$ and let $f : \mathfrak{b} \to \mathfrak{a}$ be the corresponding morphism. Since $\mathfrak{b}$ is κ-presentable, we can find a morphism $f_0 : \mathfrak{b} \to D(\mathfrak{i})$, for some $\mathfrak{i} \in \mathcal{I}$, such that $f = \lambda_\mathfrak{i} \circ f_0$. Hence, $\mathfrak{b} \in \mathrm{Sub}_\kappa(D(\mathfrak{i}))$. □

Let us characterise when a class is of the form $\mathrm{Sub}_\kappa(\mathfrak{a})$, for an arbitrary object $\mathfrak{a}$. We start with an obvious necessary condition.

Proposition 3.11. *Let* $\mathcal{C}$ *be a* κ*-accessible category. For every object* $\mathfrak{a} \in \mathcal{C}$*, the class* $\mathrm{Sub}_\kappa(\mathfrak{a})$ *is* κ*-hereditary and it has the* κ*-joint embedding property.*

Proof. Clearly, if there are morphisms $\mathfrak{b} \to \mathfrak{a}$ and $\mathfrak{c} \to \mathfrak{b}$, there is also a morphism $\mathfrak{c} \to \mathfrak{a}$. Hence, $\mathrm{Sub}_\kappa(\mathfrak{b}) \subseteq \mathrm{Sub}_\kappa(\mathfrak{a})$, for every $\mathfrak{b} \in \mathrm{Sub}_\kappa(\mathfrak{a})$.

Furthermore, we have shown in Proposition B4.4.12 that $\mathfrak{Sub}_\kappa(\mathfrak{a})$ is κ-filtered. This implies that $\mathrm{Sub}_\kappa(\mathfrak{a})$ has the κ-joint embedding property. □

The converse only holds for $\kappa = \aleph_0$ and if $\mathcal{K}$ is small enough.

Theorem 3.12 (Fraïssé). *Let* $\mathcal{C}$ *be an* $\aleph_0$*-accessible category and let* $\mathcal{K} \subseteq \mathcal{C}^{\mathrm{obj}}$ *be a class of* $\aleph_0$*-presentable objects that, up to isomorphism, contains only countably many objects. If* $\mathcal{K}$ *is* $\aleph_0$*-hereditary and if it has the* $\aleph_0$*-joint embedding property, then* $\mathcal{K} = \mathrm{Sub}_{\aleph_0}(\mathfrak{a})$*, for some* $\aleph_1$*-presentable object* $\mathfrak{a} \in \mathcal{C}$*.*

Proof. Fix an enumeration $(\mathfrak{c}_n)_{n<\omega}$ of all objects in $\mathcal{K}$ up to isomorphism. We define a diagram $D : \omega \to \mathcal{K}$ by induction on n. Set $D(0) := \mathfrak{c}_0$. If $D(n)$ is already defined then, by the $\aleph_0$-joint embedding property, we can find an object $D(n+1) \in \mathcal{K}$ with morphisms $\mathfrak{c}_{n+1} \to D(n+1)$ and $f_n : D(n) \to D(n+1)$. Setting

$$D(i,k) := f_{k-1} \circ \cdots \circ f_i\,, \quad \text{for } i < k < \omega\,,$$

we obtain a $\aleph_0$-directed diagram $D : \omega \to \mathcal{K}$. Let $\mathfrak{a}$ be its colimit. According to Proposition B4.3.7, $\mathfrak{a}$ is $\aleph_1$-presentable. Since $\mathcal{K}$ is $\aleph_0$-hereditary,

$$D(n) \in \mathcal{K} \quad \text{implies} \quad \mathrm{Sub}_{\aleph_0}(D(n)) \subseteq \mathcal{K}\,, \quad \text{for every } n < \omega\,.$$

By Lemma 3.10, it follows that $\mathrm{Sub}_{\aleph_0}(\mathfrak{a}) \subseteq \mathcal{K}$. Conversely, we have

$$\mathfrak{c}_n \in \mathrm{Sub}_{\aleph_0}(D(n)) \subseteq \mathrm{Sub}_{\aleph_0}(\mathfrak{a})\,, \quad \text{for every } n < \omega\,.$$

Since $\mathrm{Sub}_{\aleph_0}(\mathfrak{a})$ is closed under isomorphisms, this implies that $\mathcal{K} \subseteq \mathrm{Sub}_{\aleph_0}(\mathfrak{a})$. □

For a given class $\mathcal{K}$ there may be several non-isomorphic objects $\mathfrak{a}$ such that $\mathcal{K} = \mathrm{Sub}_{\aleph_0}(\mathfrak{a})$. For instance, if $\mathcal{K} \subseteq \mathfrak{Emb}(\leq)$ is the class of all finite linear orders then $\mathcal{K} = \mathrm{Sub}_{\aleph_0}(\mathfrak{L})$, for every infinite linear order $\mathfrak{L}$. We are looking for an object $\mathfrak{a}$ with $\mathrm{Sub}_{\aleph_0}(\mathfrak{a}) = \mathcal{K}$ that is in a certain sense the most general one. As we have seen in Corollary 3.6, ultrahomogeneous objects $\mathfrak{u}$ are uniquely determined by $\mathrm{Sub}_\kappa(\mathfrak{u})$. Therefore, we can take ultrahomogeneity as the required additional property.

Definition 3.13. Let $\mathcal{C}$ be a category and $\mathcal{K} \subseteq \mathcal{C}^{\mathrm{obj}}$. An object $\mathfrak{f} \in \mathcal{C}$ is a *Fraïssé limit* of $\mathcal{K}$ if there is some cardinal κ such that $\mathfrak{f}$ is κ^+-presentable, κ-ultrahomogeneous, and $\mathrm{Sub}_\kappa(\mathfrak{f}) = \mathcal{K}$.

Example. $\langle \mathbb{Q}, \leq \rangle$ is the Fraïssé limit of the class of all finite linear orders in $\mathfrak{Emb}(\leq)$.

Before considering their existence, let us prove that Fraïssé limits are unique.

Proposition 3.14. *Let $\mathcal{C}$ be an $\aleph_0$-accessible category where every morphism is a monomorphism and every epimorphism an isomorphism. Up to isomorphism, a class $\mathcal{K} \subseteq \mathcal{C}^{\mathrm{obj}}$ has at most one Fraïssé limit.*

Proof. Suppose that $\mathfrak{f}$ and $\mathfrak{g}$ are Fraïssé limits of $\mathcal{K}$. By definition, there are infinite cardinals κ and λ such that $\mathfrak{f}$ is κ^+-presentable and κ-ultrahomogeneous, $\mathfrak{g}$ is λ^+-presentable and λ-ultrahomogeneous, and

$$\mathrm{Sub}_\kappa(\mathfrak{f}) = \mathcal{K} = \mathrm{Sub}_\lambda(\mathfrak{g}) .$$

By symmetry, we may assume that $\kappa \leq \lambda$. As every object in $\mathrm{Sub}_\lambda(\mathfrak{g}) = \mathrm{Sub}_\kappa(\mathfrak{f})$ is κ-presentable, we have

$$\mathrm{Sub}_\kappa(\mathfrak{g}) = \mathrm{Sub}_\lambda(\mathfrak{g}) = \mathcal{K} = \mathrm{Sub}_\kappa(\mathfrak{f})$$

and it follows by Corollary 3.6 (b) that there exists a morphism $\mathfrak{f} \to \mathfrak{g}$. Consequently,

$$\mathrm{Sub}_\lambda(\mathfrak{f}) \subseteq \mathrm{Sub}_\lambda(\mathfrak{g}) = \mathcal{K} = \mathrm{Sub}_\kappa(\mathfrak{f}) \subseteq \mathrm{Sub}_\lambda(\mathfrak{f}) .$$

Hence, $\mathrm{Sub}_\lambda(\mathfrak{f}) = \mathrm{Sub}_\lambda(\mathfrak{g})$ and, if we can show that $\mathfrak{f}$ is λ-ultrahomogeneous, it will follow by by Corollary 3.6 (a) that $\mathfrak{f} \cong \mathfrak{g}$.

For λ-ultrahomogeneity of $\mathfrak{f}$, consider two morphisms $f, f' : \mathfrak{a} \to \mathfrak{f}$ with λ-presentable domain $\mathfrak{a}$. Then $\mathfrak{a} \in \mathrm{Sub}_\lambda(\mathfrak{f}) = \mathrm{Sub}_\lambda(\mathfrak{g}) = \mathrm{Sub}_\kappa(\mathfrak{g})$ implies that $\mathfrak{a}$ is even κ-presentable. Hence, we can use κ-ultrahomogeneity of $\mathfrak{f}$ to find the desired automorphism $\pi : \mathfrak{f} \to \mathfrak{f}$ with $f' = \pi \circ f$. □

Next, let us describe $\mathrm{Sub}_\kappa(\mathfrak{u})$ for a κ-ultrahomogeneous object $\mathfrak{u}$.

Lemma 3.15. *Let $\mathcal{C}$ be an $\aleph_0$-accessible category where every morphism is a monomorphism. If $\mathfrak{u} \in \mathcal{C}$ is κ-ultrahomogeneous then* $\mathrm{Sub}_\kappa(\mathfrak{u})$ *is κ-hereditary, closed under colimits of nonempty chains of length less than κ, and it has the κ-joint embedding property and the κ-amalgamation property.*

Proof. Note that every $\aleph_0$-accessible category is also κ-accessible. Therefore, it follows by Proposition 3.11 that the class $\mathrm{Sub}_\kappa(\mathfrak{u})$ is κ-hereditary and that it has the κ-joint embedding property. To check the κ-amalgamation property, let $f_i : \mathfrak{a} \to \mathfrak{b}_i$, $i < \gamma$, be a family of $\gamma < \kappa$ morphisms with $\mathfrak{a}, \mathfrak{b}_i \in \mathrm{Sub}_\kappa(\mathfrak{u})$. Fix morphisms $h_i : \mathfrak{b}_i \to \mathfrak{u}$, for $i < \gamma$. Since $\mathfrak{u}$ is κ-ultrahomogeneous, there exist automorphisms $\pi_i \in \mathrm{Aut}(\mathfrak{u})$ such that

$$\pi_i \circ h_i \circ f_i = h_0 \circ f_0 \,, \quad \text{for all } i < \gamma \,.$$

Consequently, $f_i : h_0 \circ f_0 \to \pi_i \circ h_i$ is a morphism of $\mathfrak{Sub}_\kappa(\mathfrak{u})$. We have seen in Proposition B4.4.12 that $\mathfrak{Sub}_\kappa(\mathfrak{u})$ is κ-filtered. Therefore, we can use Lemma B4.1.2 to find an object $g \in \mathfrak{Sub}_\kappa(\mathfrak{u})$ and morphisms

$$\varphi_i : \pi_i \circ h_i \to g \,, \quad \text{for } i < \gamma \,,$$

such that

$$\varphi_i \circ f_i = \varphi_k \circ f_k \,, \quad \text{for all } i, k < \gamma \,.$$

This family witnesses the κ-amalgamation property.

It remains to check that $\mathrm{Sub}_\kappa(\mathfrak{u})$ is closed under colimits of nonempty chains of length less than κ. Let $D : \gamma \to \mathrm{Sub}_\kappa(\mathfrak{u})$ be a chain of length $0 < \gamma < \kappa$. As $\mathcal{C}$ is $\aleph_0$-accessible, D has a colimit $\mathfrak{a}$ which, according to Theorem B4.4.3, is κ-presentable. Furthermore, Lemma 3.10 implies that

$$\mathrm{Sub}_\kappa(\mathfrak{a}) = \bigcup_{\alpha<\kappa} \mathrm{Sub}_\kappa(D(\alpha)) \subseteq \mathrm{Sub}_\kappa(\mathfrak{u}) \,.$$

Hence, it follows by Corollary 3.4 that $\mathfrak{a} \sqsubseteq^\kappa_{\mathrm{pres}} \mathfrak{u}$. Consequently, we can use Lemma 2.3 (b) to find a morphism $\mathfrak{a} \to \mathfrak{u}$. Thus, $\mathfrak{a} \in \mathrm{Sub}_\kappa(\mathfrak{u})$. □

The converse is given by the following theorem, which can be used to construct ultrahomogeneous structures by describing their class of substructures. Again we have to require $\mathcal{K}$ to be small enough.

Theorem 3.16 (Fraïssé). *Let κ be a regular cardinal, let $\mathcal{C}$ be an $\aleph_0$-accessible category where all morphisms are monomorphisms and all epimorphisms are isomorphisms, and let $\mathcal{K} \subseteq \mathcal{C}^{\mathrm{obj}}$ be a κ-hereditary class of κ-presentable objects that is closed under nonempty chains of length less* *than κ and that has the $\aleph_0$-joint embedding property and the $\aleph_0$-amalgamation property, and such that the full subcategory of $\mathcal{C}$ induced by $\mathcal{K}$ has a skeleton $\mathcal{K}_0$ with at most κ morphisms. Then $\mathcal{K}$ has a Fraïssé limit $\mathfrak{f}$.*

Proof. We will construct a diagram $D : \kappa \to \mathcal{K}_0$ satisfying the following condition:

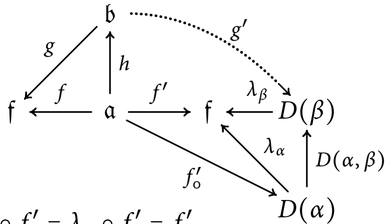
($*$) If $f : \mathfrak{a} \to \mathfrak{b}$ and $g : \mathfrak{a} \to D(\alpha)$ are morphisms with $\mathfrak{a}, \mathfrak{b} \in \mathcal{K}_0$, there is some index $\beta > \alpha$ and a morphism $g' : \mathfrak{b} \to D(\beta)$ such that

$$g' \circ f = D(\alpha, \beta) \circ g . \qquad \begin{array}{ccc} \mathfrak{b} & \xrightarrow{g'} & D(\beta) \\ {\scriptstyle f}\uparrow & & \uparrow{\scriptstyle D(\alpha,\beta)} \\ \mathfrak{a} & \xrightarrow[g]{} & D(\alpha) \end{array}$$

Let $\mathfrak{f}$ be the colimit of this diagram. By Theorem B4.4.3, $\mathfrak{f}$ is κ^+-presentable, and Lemma 3.10 implies that $\mathrm{Sub}_\kappa(\mathfrak{f}) \subseteq \mathcal{K}$. Conversely, if $\mathfrak{a} \in \mathcal{K}$ then, by the $\aleph_0$-joint embedding property, there are an object $\mathfrak{b} \in \mathcal{K}$ and morphisms $h : \mathfrak{a} \to \mathfrak{b}$ and $f : D(0) \to \mathfrak{b}$. By ($*$), we can extend the identity morphism $\mathrm{id} : D(0) \to D(0)$ to a morphism $g' : \mathfrak{b} \to D(\alpha)$, for some $\alpha > 0$. Consequently, $\mathfrak{b} \in \mathrm{Sub}_\kappa(D(\alpha)) \subseteq \mathrm{Sub}_\kappa(\mathfrak{f})$ and $\mathfrak{a} \in \mathrm{Sub}_\kappa(\mathfrak{b}) \subseteq \mathrm{Sub}_\kappa(\mathfrak{f})$. It follows that $\mathcal{K} = \mathrm{Sub}_\kappa(\mathfrak{f})$.

To show that $\mathfrak{f}$ is ultrahomogeneous it is sufficient, by Proposition 3.5, to prove that $\mathfrak{f} \sqsubseteq^\kappa_{\mathrm{pres}} \mathfrak{f}$. Consider morphisms $f : \mathfrak{a} \to \mathfrak{f}$, $f' : \mathfrak{a} \to \mathfrak{f}$, $g : \mathfrak{b} \to \mathfrak{f}$, $h : \mathfrak{a} \to \mathfrak{b}$ such that $f = g \circ h$ and $\mathfrak{a}$ and $\mathfrak{b}$ are κ-presentable. As κ is regular, the order $\langle \kappa, \leq \rangle$ is κ-directed. Since $\mathfrak{a}$ is κ-presentable, there therefore exists an essentially unique factorisation $f' = \lambda_\alpha \circ f'_0$, for some index $\alpha < \kappa$, some morphism $f'_0 : \mathfrak{a} \to D(\alpha)$, and a limiting cocone λ from D to $\mathfrak{f}$. Hence, we can use ($*$) to find an index $\beta > \alpha$ and a morphism $g' : \mathfrak{b} \to D(\beta)$ such that

$$g' \circ h = D(\alpha, \beta) \circ f'_0 .$$



Since

$$\lambda_\beta \circ g' \circ h = \lambda_\beta \circ D(\alpha, \beta) \circ f'_0 = \lambda_\alpha \circ f'_0 = f' ,$$

it follows that $\langle g, \lambda_\beta \circ g' \rangle$ is a partial morphism with

$$h : \langle f, f' \rangle \to \langle g, \lambda_\beta \circ g' \rangle .$$

Consequently, $\mathfrak{f}$ is a Fraïssé limit of $\mathcal{K}$.

It remains to construct a chain $D : \kappa \to \mathcal{K}_0$ satisfying $(*)$. Choose a bijection $\pi : \kappa \times \kappa \to \kappa$ such that $\pi(\alpha, \beta) \geq \alpha$, for all $\alpha, \beta < \kappa$. (For instance, the bijection constructed in the proof of Theorem A4.3.8 has this property.) We construct $D(\alpha)$ by induction on α. We start with an arbitrary object $D(0) \in \mathcal{K}_0$. For the successor step, suppose that $D(\alpha)$ has already been defined. Fix a list of all pairs $\langle f_{\alpha\beta}, g_{\alpha\beta} \rangle$, for $\beta < \kappa$, where $f_{\alpha\beta} : \mathfrak{a}_{\alpha\beta} \to \mathfrak{b}_{\alpha\beta}$ is a morphism in $\mathcal{K}_0$ and $g_{\alpha\beta} : \mathfrak{a}_{\alpha\beta} \to D(\alpha)$ is an arbitrary morphism. Let $\langle \gamma, \beta \rangle := \pi^{-1}(\alpha)$. Note that we have chosen π such that $\gamma \leq \alpha$. By the $\aleph_0$-amalgamation property, we can find a structure $\mathfrak{c} \in \mathcal{K}$ and morphisms $h_{\gamma\beta} : \mathfrak{b}_{\gamma\beta} \to \mathfrak{c}$ and $h'_{\gamma\beta} : D(\alpha) \to \mathfrak{c}$ such that

$$h_{\gamma\beta} \circ f_{\gamma\beta} = h'_{\gamma\beta} \circ D(\gamma, \alpha) \circ g_{\gamma\beta} .$$

We set

$$D(\alpha + 1) := \mathfrak{c} \quad \text{and} \quad D(i, \alpha + 1) := h'_{\gamma\beta} \circ D(i, \alpha) , \quad \text{for } i \leq \alpha .$$

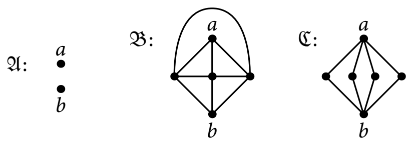
For the limit step, suppose that $D(\alpha)$ is already defined for all $\alpha < \delta$. Let $D(\delta) := \varinjlim (D \restriction \delta)$ and let λ be a corresponding limiting cocone. By assumption $D(\delta) \in \mathcal{K}_0$ and we can set $D(\alpha, \delta) := \lambda_\alpha$, for $\alpha < \delta$.

We claim that the diagram D defined this way satisfies Condition $(*)$. Let $f : \mathfrak{a} \to \mathfrak{b}$ and $g : \mathfrak{a} \to D(\alpha)$ be morphisms with $\mathfrak{a}, \mathfrak{b} \in \mathcal{K}_0$. Then

$\langle f, g\rangle = \langle f_{\alpha\beta}, g_{\alpha\beta}\rangle$, for some ordinal $\beta < \kappa$. Consequently, the morphism $h_{\alpha\beta} : \mathfrak{b}_{\alpha\beta} \to D(\pi(\alpha,\beta)+1)$ chosen in the inductive step above satisfies

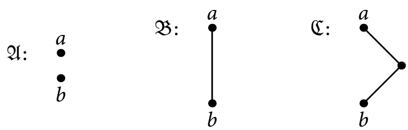
$$\begin{aligned} h_{\alpha\beta} \circ f_{\alpha\beta} &= h'_{\alpha\beta} \circ D(\alpha, \pi(\alpha,\beta)) \circ g_{\alpha\beta} \\ &= D(\alpha, \pi(\alpha,\beta)+1) \circ g_{\alpha\beta} \,. \end{aligned}$$ □

Example. (a) Let $\mathcal{P} \subseteq \mathfrak{Emb}(E)$ be the class of all finite planar graphs. Clearly, $\mathcal{P}$ is hereditary. The class $\mathcal{P}$ does not have a Fraïssé limit since it does not have the $\aleph_0$-amalgamation property. Consider the following graphs:



Let $f : \mathfrak{A} \to \mathfrak{B}$ and $g : \mathfrak{A} \to \mathfrak{C}$ be the embeddings with $a \mapsto a$ and $b \mapsto b$. There is no planar graph $\mathfrak{D}$ such that we can find embeddings $h : \mathfrak{B} \to \mathfrak{D}$ and $k : \mathfrak{C} \to \mathfrak{D}$ with $h \circ f = k \circ g$.

(b) Similarly we can show that the class $\mathcal{F} \subseteq \mathfrak{Emb}(E)$ of all finite acyclic graphs does not have the $\aleph_0$-amalgamation property. The counterexample is given by the graphs:



4. *Zero-one laws*

In this section we study Fraïssé limits by axiomatising their theories.

Definition 4.1. (a) Let $\mathfrak{M}$ be a structure. The *atomic type* of $\bar{a} \subseteq M$ is the set

$$\operatorname{atp}(\bar{a}) := \{\, \varphi \mid \varphi \text{ a literal such that } \mathfrak{M} \vDash \varphi(\bar{a}) \,\}\,.$$

An *atomic n-type* $\mathfrak{p}$ is a set of the form $\mathfrak{p} = \mathrm{atp}(\bar{a})$, for $\bar{a} \in M^n$.

(b) Let $\mathfrak{p}$ be an atomic n-type and $\mathfrak{q}$ an atomic $(n+1)$-type such that $\mathfrak{p} \subseteq \mathfrak{q}$. The *extension axiom* associated with $\mathfrak{p}$ and $\mathfrak{q}$ is the sentence

$$\eta_{\mathfrak{p}\mathfrak{q}} := \forall \bar{x}[\mathfrak{p}(\bar{x}) \to \exists y \mathfrak{q}(\bar{x}, y)]\,.$$

(We write $\mathfrak{p}(\bar{x})$ for the formula $\bigwedge \mathfrak{p}$.)

(c) Let $\mathcal{K}$ be a hereditary class of finitely generated structures. We define

$$\Gamma_{\mathcal{K}} := \{\, \mathrm{atp}(\bar{a}/\mathfrak{M}) \mid \bar{a} \text{ is a finite tuple generating } \mathfrak{M} \in \mathcal{K} \,\}\,,$$

and $\quad T[\mathcal{K}] := \{\, \eta_{\mathfrak{p}\mathfrak{q}} \mid \mathfrak{q} \in \Gamma_{\mathcal{K}} \,\} \cup \{\, \forall \bar{x} \neg \mathfrak{p}(\bar{x}) \mid \mathfrak{p} \notin \Gamma_{\mathcal{K}} \,\}$

The set of all extension axioms over a signature Σ is $T_{\mathrm{ran}}[\Sigma] := T[\mathcal{C}]$, where $\mathcal{C}$ is the class of all finitely generated Σ-structures.

Remark. Note that, in general, $T[\mathcal{K}]$ is an infinitary theory. It is a first-order theory if the signature in question is finite and relational.

Example. An important example of a Fraïssé limit is the *random graph,* also called the *Rado graph.* It can be defined as follows. $\mathfrak{R} := \langle V, E \rangle$ where $V := \mathrm{HF}$ is the set of all hereditary finite sets and the edge relation is

$$E := \{\, \langle a, b \rangle \mid a \in b \text{ or } b \in a \,\}\,.$$

This graph satisfies the following extension axiom: for every pair X, Y of finite disjoint sets of vertices, there exists some vertex $c \in V$ that is adjacent to every vertex in X, but not adjacent to any in Y. For a proof, note that, if $X = \{a_0, \dots, a_{m-1}\}$ and $Y = \{b_0, \dots, b_{n-1}\}$ then we can take $c := \{a_0, \dots, a_{m-1}, x\}$ where the set $x := \{b_0, \dots, b_{n-1}\}$ is needed to ensure that $c \notin b_i$.

Let us investigate the relationship between the theories $T[\mathcal{K}]$ and ultrahomogeneous structures.

Lemma 4.2. *If* $\mathfrak{U}$ *is ultrahomogeneous then* $\mathfrak{U} \vDash T[\mathrm{Sub}_{\aleph_0}(\mathfrak{U})]$.

Lemma 4.3. *If* $\mathfrak{A}, \mathfrak{B} \vDash T[\mathcal{K}]$ *then*

$$\mathfrak{A} \equiv_0 \mathfrak{B} \quad \textit{implies} \quad \mathfrak{A} \cong_0^{\aleph_0} \mathfrak{B}\,.$$

Proof. Since $\mathfrak{A} \equiv_0 \mathfrak{B}$ we have $\mathrm{pIso}_{\aleph_0}(\mathfrak{A}, \mathfrak{B}) \neq \emptyset$. To check the forth condition, let $\bar a \mapsto \bar b \in \mathrm{pIso}_{\aleph_0}(\mathfrak{A}, \mathfrak{B})$ and $c \in A$. Set $\mathfrak{p} := \mathrm{atp}(\bar a)$ and $\mathfrak{q} := \mathrm{atp}(\bar a c)$. Then $\mathfrak{p} \subseteq \mathfrak{q}$ and $\mathfrak{q} \in \Gamma_{\mathcal{K}}$. Hence, $\eta_{\mathfrak{pq}} \in T[\mathcal{K}]$ and $\mathfrak{B} \vDash \eta_{\mathfrak{pq}}$. Since $\mathrm{atp}(\bar b) = \mathfrak{p}$ we can, therefore, find some $d \in B$ with $\mathrm{atp}(\bar b d) = \mathfrak{q}$. Consequently, $\bar a c \mapsto \bar b d \in \mathrm{pIso}_{\aleph_0}(\mathfrak{A}, \mathfrak{B})$. □

Corollary 4.4. *Every model of* $T[\mathcal{K}]$ *is ultrahomogeneous.*

It follows that the theories $T[\mathcal{K}]$ axiomatise Fraïssé limits.

Theorem 4.5. *Let* $\mathcal{K}$ *be a hereditary class of finitely generated structures containing a unique* 0*-generated structure* $\mathfrak{A}_0$*. A structure* $\mathfrak{F}$ *is the Fraïssé limit of* $\mathcal{K}$ *if and only if it is countable,* $\langle\!\langle \emptyset \rangle\!\rangle_{\mathfrak{F}} \cong \mathfrak{A}_0$*, and* $\mathfrak{F} \vDash T[\mathcal{K}]$.

Proof. ($\Rightarrow$) A Fraïssé limit $\mathfrak{F}$ is countable by definition. Furthermore, $\mathrm{Sub}_{\aleph_0}(\mathfrak{F}) \subseteq \mathcal{K}$ implies that $\mathfrak{F} \vDash \forall \bar x \neg \mathfrak{p}(\bar x)$, for all $\mathfrak{p} \notin \Gamma_{\mathcal{K}}$.

Finally, let $\eta_{\mathfrak{pq}} \in T[\mathcal{K}]$. Then $\mathfrak{q} \in \Gamma_{\mathcal{K}}$ and $\mathcal{K} \subseteq \mathrm{Sub}_{\aleph_0}(\mathfrak{F})$ implies that there is some tuple $\bar c \subseteq F$ with $\mathrm{atp}(\bar c) = \mathfrak{q}$. Since $\mathfrak{F}$ is ultrahomogeneous it follows that, for every tuple $\bar a$ with $\mathrm{atp}(\bar a) = \mathfrak{p}$, there is some element $b \in F$ such that $\mathrm{atp}(\bar a b) = \mathrm{atp}(\bar c) = \mathfrak{q}$. Hence, $\mathfrak{F} \vDash \eta_{\mathfrak{pq}}$.

($\Leftarrow$) By assumption, $\mathfrak{F}$ is countable, and we have shown in Corollary 4.4 that it is ultrahomogeneous. Furthermore, $\mathfrak{F} \vDash \forall \bar x \neg \mathfrak{p}(\bar x)$, for $\mathfrak{p} \notin \Gamma_{\mathcal{K}}$ implies that $\mathrm{Sub}_{\aleph_0}(\mathfrak{F}) \subseteq \mathcal{K}$. Hence, it remains to show that $\mathcal{K} \subseteq \mathrm{Sub}_{\aleph_0}(\mathfrak{F})$. Let $\mathfrak{B} \in \mathcal{K}$ be generated by a finite tuple $\bar b = b_0 \dots b_{n-1}$. Note that $\langle\!\langle \emptyset \rangle\!\rangle_{\mathfrak{B}} \cong \mathfrak{A}_0 \cong \langle\!\langle \emptyset \rangle\!\rangle_{\mathfrak{F}} \subseteq \mathfrak{F}$. Since $\mathfrak{F}$ satisfies the needed extension axioms we can, therefore, use induction to find elements $a_0, \dots, a_{n-1} \in F$ such that

$$\langle\!\langle b_0 \dots b_{k-1} \rangle\!\rangle_{\mathfrak{B}} \cong \langle\!\langle a_0, \dots, a_{k-1} \rangle\!\rangle_{\mathfrak{F}}\,, \quad \text{for all } k \leq n\,.$$

Consequently, we have $\mathfrak{B} = \langle\!\langle \bar b \rangle\!\rangle_{\mathfrak{B}} \cong \langle\!\langle \bar a \rangle\!\rangle_{\mathfrak{F}} \subseteq \mathfrak{F}$. □

Proposition 4.6. *$T[\mathcal{K}]$ admits quantifier elimination for* $\mathrm{FO}_{\infty\aleph_0}$.

Proof. This follows immediately from Theorem D1.2.9 and Lemma 4.3. □

Corollary 4.7. *Let $\mathcal{K}$ be a class of Σ-structures where the signature Σ is finite and relational. Then $T[\mathcal{K}]$ admits quantifier elimination for* FO.

Proof. Since $T[\mathcal{K}]$ is a first-order theory, the claim follows by Corollary D1.2.10. □

Corollary 4.8. *Let $\mathcal{K}$ be a class of Σ-structures where Σ is a finite, relational signature without* 0*-ary relations. Then $T[\mathcal{K}]$ is complete.*

Proof. Let $\varphi \in \mathrm{FO}^0[\Sigma]$. There exists a sentence $\psi \in \mathrm{QF}^0[\Sigma]$ such that $T[\mathcal{K}] \vDash \varphi \leftrightarrow \psi$. Since Σ is relational and it contains no 0-ary relations, the only quantifier-free sentences are true and false. If $\psi \equiv$ true then $T[\mathcal{K}] \vDash \varphi$ and if $\psi \equiv$ false then $T[\mathcal{K}] \vDash \neg\varphi$. □

The extension axioms have the surprising property that, asymptotically, they hold with probability 1 in every finite structure. Let us make this claim more precise.

Consider a finite signature Σ. For each finite number $n < \omega$, we count how many Σ-structures with universe $[n]$ satisfy a given sentence. Note that, for every n, there are only finitely many such structures.

Definition 4.9. For $\varphi, \psi \in \mathrm{FO}[\Sigma]$ we define

$$\kappa_n(\varphi) := |\{\, \mathfrak{M} \mid \mathfrak{M} \vDash \varphi,\ M = [n] \,\}|,$$

$$\Pr^n_{\mathfrak{M}}[\, \mathfrak{M} \vDash \varphi \mid \mathfrak{M} \vDash \psi \,] := \frac{\kappa_n(\varphi \wedge \psi)}{\kappa_n(\psi)}.$$

We use the shorthand $\Pr^n_{\mathfrak{M}}[\, \mathfrak{M} \vDash \varphi \,] := \Pr^n_{\mathfrak{M}}[\, \mathfrak{M} \vDash \varphi \mid \mathfrak{M} \vDash \text{true} \,]$.

Lemma 4.10. *Let Σ be a finite, relational signature without* 0*-ary relations. Then*

$$\lim_{n\to\infty} \Pr^n_{\mathfrak{M}}[\, \mathfrak{M} \vDash \eta_{\mathfrak{pq}} \,] = 1, \quad \textit{for every } \eta_{\mathfrak{pq}} \in T_{\mathrm{ran}}[\Sigma].$$

Proof. Suppose that $\mathfrak{p}$ is an m-type and $n > m$. Since Σ is finite there exists some constant $p \in (0,1)$ such that

$$\mathrm{Pr}^n_{\mathfrak{M}}\big[\,\mathfrak{M} \nvDash \mathfrak{q}(0,\dots,m-1,m)\,\big|\,\mathfrak{M} \vDash \mathfrak{p}(0,\dots,m-1)\,\big] = p\,.$$

Hence,

$$\begin{aligned}\mathrm{Pr}^n_{\mathfrak{M}}\big[\,\mathfrak{M} \nvDash \exists x_m \mathfrak{q}(0,\dots,m-1,x_m)\,\big|\,\mathfrak{M} \vDash \mathfrak{p}(0,\dots,m-1)\,\big]&\\ = p^{n-m},&\end{aligned}$$

which implies that $\mathrm{Pr}^n_{\mathfrak{M}}[\,\mathfrak{M} \nvDash \eta_{\mathfrak{pq}}\,] \leq n^m k^{n-m}$. Since $p < 1$ we have

$$\lim_{n\to\infty} n^m k^{n-m} = 0\,,$$

and it follows that

$$\lim_{n\to\infty} \mathrm{Pr}^n_{\mathfrak{M}}[\,\mathfrak{M} \vDash \eta_{\mathfrak{pq}}\,] \geq \lim_{n\to\infty}(1 - n^m k^{n-m}) = 1\,.$$ □

Lemma 4.11. *$T_{\mathrm{ran}}[\Sigma]$ is satisfiable, for every finite relational signature Σ without* 0*-ary relations.*

Proof. For a contradiction suppose that $T_{\mathrm{ran}}[\Sigma]$ is inconsistent. Then there exists a finite inconsistent set $\Phi \subseteq T_{\mathrm{ran}}[\Sigma]$. Suppose that $\Phi = \{\varphi_0,\dots,\varphi_{m-1}\}$. By the preceding lemma, we have

$$\lim_{n\to\infty} \mathrm{Pr}^n_{\mathfrak{M}}[\,\mathfrak{M} \vDash \varphi_i\,] = 1\,, \quad \text{for all } i < m\,.$$

Therefore, there exists some number n such that

$$\mathrm{Pr}^n_{\mathfrak{M}}[\,\mathfrak{M} \vDash \neg\varphi_i\,] < \frac{1}{m}\,.$$

It follows that

$$\begin{aligned}\mathrm{Pr}^n_{\mathfrak{M}}[\,\mathfrak{M} \vDash \textstyle\bigwedge \Phi\,] &= 1 - \mathrm{Pr}^n_{\mathfrak{M}}[\,\mathfrak{M} \vDash \textstyle\bigvee_i \neg\varphi_i\,]\\ &\geq 1 - \sum_i \mathrm{Pr}^n_{\mathfrak{M}}[\,\mathfrak{M} \vDash \neg\varphi_i\,] > 1 - m\cdot\frac{1}{m} = 0\,.\end{aligned}$$

Consequently, Φ has a model of size n. Contradiction. □

Theorem 4.12 (Zero-One Law). *Let Σ be a finite, relational signature without 0-ary relations. For every sentence $\varphi \in \mathrm{FO}[\Sigma]$, we have*

$$\lim_{n\to\infty} \Pr^n_{\mathfrak{M}}[\,\mathfrak{M} \vDash \varphi\,] \in \{0,1\}\,.$$

Proof. If $T_{\mathrm{ran}}[\Sigma] \vDash \varphi$ then there are axioms $\eta_{\mathfrak{p}_0\mathfrak{q}_0}, \dots, \eta_{\mathfrak{p}_k\mathfrak{q}_k} \in T_{\mathrm{ran}}[\Sigma]$ such that $\eta_{\mathfrak{p}_0\mathfrak{q}_0} \wedge \dots \wedge \eta_{\mathfrak{p}_k\mathfrak{q}_k} \vDash \varphi$. Hence, we have

$$\lim_{n\to\infty} \Pr^n_{\mathfrak{M}}[\,\mathfrak{M} \vDash \varphi\,] \geq \lim_{n\to\infty} \Pr^n_{\mathfrak{M}}[\,\mathfrak{M} \vDash \eta_{\mathfrak{p}_0\mathfrak{q}_0} \wedge \dots \wedge \eta_{\mathfrak{p}_k\mathfrak{q}_k}\,] = 1\,.$$

Now suppose that $T_{\mathrm{ran}}[\Sigma] \nvDash \varphi$. Since $T_{\mathrm{ran}}[\Sigma]$ is complete, we have $T_{\mathrm{ran}}[\Sigma] \vDash \neg\varphi$. By the first case, it follows that

$$\lim_{n\to\infty} \Pr^n_{\mathfrak{M}}[\,\mathfrak{M} \vDash \varphi\,] = \lim_{n\to\infty} \big(1 - \Pr^n_{\mathfrak{M}}[\,\mathfrak{M} \vDash \neg\varphi\,]\big) = 1 - 1 = 0\,. \qquad \square$$

Exercise 4.1. Prove that the theorem fails for signatures with 0-ary relations.

Lemma 4.13. *The Zero-One Law fails for signatures with functions.*

Proof. Let $\Sigma = \{f\}$ be a signature consisting just of a unary function symbol f, and define

$$\varphi := \forall x(fx \neq x)\,.$$

We have

$$\Pr^n_{\mathfrak{M}}[\,\mathfrak{M} \vDash \varphi\,] = \frac{(n-1)^n}{n^n} = \big(1 - \tfrac{1}{n}\big)^n$$

which implies that

$$\lim_{n\to\infty} \Pr^n_{\mathfrak{M}}[\,\mathfrak{M} \vDash \varphi\,] = \lim_{n\to\infty} \big(1 - \tfrac{1}{n}\big)^n = \frac{1}{e}\,. \qquad \square$$

Lemma 4.14. *Let Σ be a finite relational signature. There exists no sentence $\varphi \in \mathrm{FO}[\Sigma]$ such that*

$$\mathfrak{M} \vDash \varphi \quad \text{iff} \quad |M| \textit{ is even}\,, \qquad \textit{for all finite } \Sigma\textit{-structures } \mathfrak{M}\,.$$

Proof. $\lim_{n\to\infty} \Pr^n_{\mathfrak{M}}[\,\mathfrak{M} \vDash \varphi\,]$ does not exist in contradiction to the Zero-One Law. □

Remark. For every $n < \omega$, we can extend the Zero-One Law to the logic $\mathrm{FO}^{(n)}_{\infty\aleph_0}$ consisting of all $\mathrm{FO}_{\infty\aleph_0}$-formulae using at most n variables (both free and bound). Note that every FO(PFP)-formula can be translated to such a formula, for some suitable n. Hence, the Zero-One Law also holds for FO(LFP) and FO(PFP).

E5. *Indiscernible sequences*

1. *Ramsey Theory*

In this chapter we introduce some technical tools to study properties of sequences. This machinery is based on combinatorial results concerning colourings of linear orders.

Definition 1.1. (a) For a linear order I and a cardinal ν, we define

$$[I]^\nu := \{\, \bar\imath \in I^\nu \mid \bar\imath \text{ is increasing} \,\}\,.$$

For an unordered set X we abuse notation by defining

$$[X]^\nu := \{\, s \subseteq X \mid |s| = \nu \,\}\,.$$

(This is consistent with our convention of identifying sequences with their ranges.)

(b) Let $c : [A]^\nu \to \lambda$ be a function. A subset $C \subseteq A$ is *homogeneous* with respect to c if we have $c(\bar a) = c(\bar a')$, for all $\bar a, \bar a' \in [C]^\nu$.

(c) Let $\kappa, \lambda, \mu, \nu$ be cardinals. We write $\kappa \to (\mu)^\nu_\lambda$ if, for every set A of size $|A| \geq \kappa$ and each function $c : [A]^\nu \to \lambda$, there exists a homogeneous subset $C \subseteq A$ of size $|C| \geq \mu$.

Example. $6 \to (3)^2_2$ is equivalent to the statement that every undirected graph $\mathfrak{G} = \langle V, E\rangle$ with at least 6 elements contains a triangle or an independent set of size 3.

Exercise 1.1. Prove that $6 \to (3)^2_2$.

Let us start with the simplest case, that of unary colourings.

Theorem 1.2 (Pigeon Hole Principle). $\kappa \to (\kappa)^1_\lambda$, *for all infinite cardinals* κ *and every* $\lambda < \operatorname{cf} \kappa$.

Proof. Let A be a set of size $|A| = \kappa$ and suppose that $c : A \to \lambda$ is a function. We have to show that there is some $\alpha < \lambda$ with $|c^{-1}(\alpha)| = \kappa$. Suppose otherwise. Then $\lambda < \operatorname{cf} \kappa$ implies

$$|A| = \sum_{\alpha<\lambda} |c^{-1}(\alpha)| < \kappa \,.$$

A contradiction. □

The Theorem of Ramsey generalises the Pigeon Hole Principle to colourings of higher arities. We present two versions: one for infinite sets and one for finite sets.

Theorem 1.3 (Ramsey). $\aleph_0 \to (\aleph_0)^n_l$, *for all* $0 < n, l < \aleph_0$.

Proof. Let A be a set of size $|A| = \aleph_0$ and $c : [A]^n \to l$ a function. W.l.o.g. we may assume that $A = \omega$. By induction on n, we construct an infinite subset $C \subseteq \omega$ that is homogeneous with respect to c.

For $n = 1$ the claim follows from the Pigeon Hole Principle. Hence, we may assume that $n > 1$. In a first step, we define an infinite subset $B \subseteq \omega$ such that the value of $c(\bar b)$, for $\bar b \in [B]^n$, only depends on the minimal element b_0. For every $a \in \omega$, we define a function $c'_a : [\omega \setminus \{a\}]^{n-1} \to l$ by $c'_a(\bar b) := c(\bar b \cup \{a\})$. We construct an increasing sequence $a_0 < a_1 < \dots$ of elements and a decreasing sequence $A_0 \supseteq A_1 \supseteq \dots$ of subsets of ω as follows. We start with $a_0 := 0$ and $A_0 := \omega$. If a_i and A_i are already defined then we can use the inductive hypothesis to find an infinite subset $A_{i+1} \subseteq A_i \setminus \{a_0, \dots, a_i\}$. that is homogeneous with respect to c'_{a_i}. Let a_{i+1} be the minimal element of A_{i+1}.

Let $B := \{ a_i \mid i < \omega \}$ and set $k_i := c(a_i a_{i+1} \dots a_{i+n-1})$. Note that, for $i_0 < \dots < i_{n-1}$, we have $a_{i_1}, \dots, a_{i_{n-1}} \in A_{i_0+1}$. Hence, the above construction ensures that

$$\begin{aligned} c(a_{i_0} \dots a_{i_{n-1}}) &= c'_{a_{i_0}}(a_{i_1} \dots a_{i_{n-1}}) \\ &= c'_{a_{i_0}}(a_{i_0+1} \dots a_{i_0+n-1}) = c(a_{i_0} \dots a_{i_0+n-1}) = k_{i_0} \,. \end{aligned}$$

By the Pigeon Hole Principle, there exists an infinite subset $C \subseteq B$ such that $k_i = k_j$, for all $a_i, a_j \in C$. This set C is the desired homogeneous subset of ω. □

Example. Let $\langle P, \leq \rangle$ be an infinite partial order. We can use the Ramsey Theorem to prove that there exists an infinite set $C \subseteq P$ such that C is either linearly ordered or all elements of C are pairwise incomparable.

Let $c : [P]^2 \to 2$ be the function such that

$$c(\{a, b\}) := \begin{cases} 1 & \text{if } a \leq b \text{ or } b \leq a\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

By the theorem there exists an infinite homogeneous set $C \subseteq P$. If we have $c(\{a, b\}) = 1$, for all $a, b \in C$, then C is a chain. Otherwise, all elements of C are pairwise incomparable.

The finite version of the Ramsey Theorem is as follows.

Theorem 1.4 (Ramsey). *For all $l, m, n < \aleph_0$, there exists a finite cardinal $k < \aleph_0$ such that $k \to (m)^n_l$.*

Proof. For a contradiction, suppose that there exists no finite k with $k \to (m)^n_l$. Let F_k be the set of all functions $c : [k]^n \to l$ such that there is no subset $C \subseteq [k]$ of size $|C| \geq m$ that is homogeneous with respect to c. It follows that each set F_k is finite and nonempty. Furthermore, $c \in F_{k+1}$ implies that $c \restriction [k]^n \in F_k$. Hence, if we order the set $T := \bigcup_k F_k$ by inclusion then we obtain a tree $\langle T, \subseteq \rangle$. This tree is infinite and finitely branching. By the Lemma of Kőnig it therefore contains an infinite branch $(c_k)_{k<\omega}$ with $c_k \in F_k$. Set $c := \bigcup_k c_k$. Then c is a function $c : [\aleph_0]^n \to l$. By the infinite version of the Ramsey Theorem, there exists an infinite subset $C \subseteq \aleph_0$ that is homogeneous with respect to c. Fix a subset $Z \subseteq C$ of size $|Z| = m$ and let k be the maximal element of Z. It follows that Z is homogeneous with respect to c_{k+1}. A contradiction. □

Next, we consider the case of infinitely many colours and uncountable homogeneous sets. We start with a counterexample.

Lemma 1.5. $2^{\aleph_0} \nrightarrow (3)^2_{\aleph_0}$

Proof. Let $c : [2^{\aleph_0}]^2 \to \aleph_0$ be the function mapping a pair $\{f, g\}$ of distinct functions $f, g : \aleph_0 \to 2$ to the least number n with $f(n) \neq g(n)$. If $\{f, g, h\}$ were homogeneous with respect to c, we would have $f(n) \neq g(n)$, $f(n) \neq h(n)$, and $g(n) \neq h(h)$, for some n. Since $f(n), g(n), h(n) \in \{0, 1\}$ this is impossible. □

Theorem 1.6 (Erdős-Rado). *For all cardinals $\kappa \geq \aleph_0$ and $n < \aleph_0$,*

$$\beth_n(\kappa)^+ \to (\kappa^+)^{n+1}_\kappa .$$

Proof. We prove the claim by induction on n. By the Pigeon Hole Principle, we have $\kappa^+ \to (\kappa^+)^1_\kappa$. Hence, the claim holds for $n = 1$. For the inductive step, suppose we have already proved the theorem for n. Set $\lambda := \beth_{n+1}(\kappa)$ and $\mu := \beth_n(\kappa)$, and let $c : [\lambda^+]^{n+1} \to \kappa$ be a colouring.

As a first step we define an increasing sequence of ordinals $\beta_i < \lambda^+$, for $i < \kappa^+$, with the following property:

(∗) For every set $S \subseteq \beta_i$ of size $|S| \leq \mu$ and all ordinals $\gamma < \lambda^+$, there exists some ordinal $\eta < \beta_{i+1}$ such that

$$\eta \in S \quad \text{iff} \quad \gamma \in S ,$$

$$\text{and} \quad c(\bar{\alpha}\eta) = c(\bar{\alpha}\gamma) , \quad \text{for all } \bar{\alpha} \in S^n .$$

The ordinals β_i will be used as a measuring stick in the construction below. We define β_i by induction on i. Let $\beta_0 := 0$ and set $\beta_\delta := \sup_{i<\delta} \beta_i$, for limit ordinals δ. For the inductive step, we set

$$\beta_{i+1} := \sup \{ \eta(S, \gamma) \mid \gamma < \lambda^+ ,\ S \subseteq \beta_i \text{ with } |S| \leq \mu \} ,$$

where $\eta(S, \gamma)$ denotes the minimal ordinal η such that

$$\eta \in S \quad \text{iff} \quad \gamma \in S ,$$

$$\text{and} \quad c(\bar{\alpha}\eta) = c(\bar{\alpha}\gamma) , \quad \text{for all } \bar{\alpha} \in S^n .$$

Note that there are at most $|\beta_i|^\mu = \lambda^\mu = (2^\mu)^\mu = \lambda$ subsets of β_i of size $|S| \leq \mu$ and there are at most $\kappa^\mu = 2^\mu = \lambda$ functions $S \to \kappa$. Consequently, the supremum above is taken over a set of at most $\lambda \otimes \lambda = \lambda$ ordinals each of which is less than λ^+. Since λ^+ is regular it follows that the supremum β_{i+1} is less than λ^+.

Having defined the β_i we set $\beta^* := \sup_{i<\mu^+} \beta_i$ and we define ordinals $\alpha_i < \beta_{i+1}$, for $i < \mu^+$, such that $\alpha_i \neq \alpha_k$, for $i \neq k$, and

$$c(\alpha_{k_0}, \dots, \alpha_{k_{n-1}}, \alpha_i) = c(\alpha_{k_0}, \dots, \alpha_{k_{n-1}}, \beta^*),$$

for all $k_0, \dots, k_{n-1} < i$. We can find α_i by induction on i using property $(*)$ with $S = \{ \alpha_k \mid k < i \}$ and $\gamma := \beta^*$.

Define a colouring $c' : [\mu^+]^n \to \kappa$ by

$$c'(\bar{\imath}) := c(\alpha_{i_0} \dots \alpha_{i_{n-1}} \beta^*).$$

By inductive hypothesis, there exists a set $I \subseteq \mu^+$ of size $|I| \geq \kappa^+$ such that

$$c'(\bar{\imath}) = c'(\bar{k}), \quad \text{for all } \bar{\imath}, \bar{k} \in [I]^n.$$

Let $J := \{ \alpha_i \mid i \in I \}$. For $\bar{\gamma}, \bar{\eta} \in [J]^{n+1}$ it follows that

$$\begin{aligned} c(\gamma_0 \dots \gamma_{n-1}\gamma_n) &= c(\gamma_0 \dots \gamma_{n-1}\beta^*) \\ &= c(\eta_0 \dots \eta_{n-1}\beta^*) = c(\eta_0 \dots \eta_{n-1}\eta_n). \end{aligned}$$

Hence, J is the desired homogeneous subset of λ^+. □

2. *Ramsey Theory for trees*

So far, we have considered homogeneous subsets of linear orders. A special property of linear orders is that every subset again induces a linear order. When considering colourings of other structures this is no longer the case. In this section we prove variants of the Pigeon Hole Principle and the Theorem of Ramsey for trees where the homogeneous sets we obtain again induce trees. There are two kinds of tree structures we will be working with: trees of the form $\mathfrak{T}_*(\kappa^{<\alpha})$ are equipped with the tree-order $\leq$ and relations $\prec_p$ for the direction of the immediate successors, while trees $\mathfrak{T}_n(\kappa^{<\alpha})$ also have functions pf to compare the levels of elements.

Definition 2.1. Let κ be a cardinal and α an ordinal.

(a) We denote the tree order on $\kappa^{<\alpha}$ by $\leq$ and $\sqcap$ is the infimum operation with respect to $\leq$. For $\eta, \zeta \in \kappa^{<\alpha}$ and $p \in \kappa$, we further set

$$\eta \prec_p \zeta \quad :\text{iff} \quad \eta p \leq \zeta\,.$$

For $|\eta| \leq |\zeta|$, we denote by $\mathrm{pf}(\eta, \zeta)$ the prefix of ζ of length $|\eta|$. If $|\eta| > |\zeta|$, we set $\mathrm{pf}(\eta, \zeta) := \zeta$.

(b) We define

$$\mathfrak{T}_*(\kappa^{<\alpha}) := \langle \kappa^{<\alpha}, \sqcap, \leq, (\prec_p)_{p\in\kappa} \rangle\,,$$

and $\quad \mathfrak{T}_n(\kappa^{<\alpha}) := \langle \kappa^{<\alpha}, \sqcap, \leq, (\prec_p)_{p\in\kappa}, \mathrm{pf}, (\eta)_{\eta\in\kappa^{<n}} \rangle\,, \quad$ for $n \leq \alpha$.

We denote the substructure of $\mathfrak{T}_n(\kappa^{<\alpha})$ generated by a set $X \subseteq \kappa^{<\alpha}$ by $\langle\!\langle X \rangle\!\rangle_n$.

Remark. (a) Note that the substructure $\langle\!\langle X \rangle\!\rangle_n$ generated by a set $X \subseteq \kappa^{<\alpha}$ has universe

$$\langle\!\langle X \rangle\!\rangle_n = \kappa^{<n} \cup \{\, \mathrm{pf}(\xi \sqcap \eta, \zeta) \mid \xi, \eta, \zeta \in X \,\}\,.$$

Thus, it consists of (i) all elements of $X \cup \kappa^{<n}$, (ii) all elements of the form $\eta \sqcap \zeta$, with $\eta, \zeta \in X$, and (iii) all prefixes of some element of X that have the same length as an element of the form (i) or (ii).

(b) Note that we have

$$|\eta| = |\zeta| \quad \text{iff} \quad \mathrm{pf}(\eta, \zeta) = \zeta \text{ and } \mathrm{pf}(\zeta, \eta) = \eta\,.$$

Hence, every embedding $h : \mathfrak{T}_n(\kappa^{<\alpha}) \to \mathfrak{T}_n(\kappa^{<\alpha})$ has the property that

$$|\eta| = |\zeta| \quad \text{implies} \quad h(|\eta|) = h(|\zeta|)\,, \quad \text{for all } \eta, \zeta \in \kappa^{<\alpha}.$$

Definition 2.2. (a) The set of *levels* of a tuple $\bar{\eta} \in (\kappa^{<\alpha})^d$ is

$$\mathrm{Lvl}(\bar{\eta}) := \{\, |\eta_i \sqcap \eta_j| \mid i, j < d \,\} = \{\, |\zeta| \mid \zeta \in \langle\!\langle \bar{\eta} \rangle\!\rangle_0 \,\}.$$

(b) Let $h : \mathfrak{T}_n(\kappa^{<\alpha}) \to \mathfrak{T}_n(\kappa^{<\alpha})$ be an embedding. The *level embedding function* associated with h is the function $f : \alpha \to \alpha$ such that

$$|h(\eta)| = f(|\eta|), \quad \text{for all } \eta \in \kappa^{<\alpha}.$$

Our first result is a generalisation of a strong version of the Pigeon Hole Principle. We omit the proof, which is quite involved.

Theorem 2.3 (Halpern, Läuchli). *Let $m, d < \omega$ and let C be a finite set. For every function $c : (m^{<\omega})^d \to C$ there exist embeddings*

$$g_i : \mathfrak{T}_0(m^{<\omega}) \to \mathfrak{T}_0(m^{<\omega}), \quad \textit{for } i < d,$$

such that all g_i have the same level embedding function and

$$c(g_0(\eta_0), \dots, g_{d-1}(\eta_{d-1})) = c(g_0(\zeta_0), \dots, g_{d-1}(\zeta_{d-1})),$$

for all tuples $\bar{\eta}, \bar{\zeta} \in (m^{<\omega})^d$ with $|\eta_0| = \dots = |\eta_{d-1}|$ and $|\zeta_0| = \dots = |\zeta_{d-1}|$.

In the remainder of this section we generalise the Theorem of Ramsey to trees. In the version for linear orders we required tuples to have the same colour if they have the same order type. When dealing with other kinds of structures we replace the order type of a tuple by its atomic type.

Definition 2.4. (a) Let $c : A^d \to C$ a function, for $d < \omega$, and let $\approx$ be an equivalence relation on A^d. A subset $X \subseteq A$ is *$\approx$-homogeneous* with respect to c if

$$\bar{\eta} \approx \bar{\zeta} \quad \text{implies} \quad c(\bar{\eta}) = c(\bar{\zeta}), \quad \text{for all } \bar{\eta}, \bar{\zeta} \in X^d.$$

(b) For tuples $\bar{\eta}, \bar{\zeta} \subseteq \kappa^{<\alpha}$, we define

$$\bar{\eta} \approx_* \bar{\zeta} \quad :\text{iff} \quad \mathrm{atp}(\bar{\eta}/\mathfrak{T}_*(\kappa^{<\alpha})) = \mathrm{atp}(\bar{\zeta}/\mathfrak{T}_*(\kappa^{<\alpha})),$$
$$\bar{\eta} \approx_n \bar{\zeta} \quad :\text{iff} \quad \mathrm{atp}(\bar{\eta}/\mathfrak{T}_n(\kappa^{<\alpha})) = \mathrm{atp}(\bar{\zeta}/\mathfrak{T}_n(\kappa^{<\alpha})).$$

Our goal is to prove the following variant of the Theorem of Ramsey for trees.

Theorem 2.5 (Milliken). *Let $m, d < \omega$ and let C be a finite set. For every function $c : (m^{<\omega})^d \to C$ there exists an embedding $g : \mathfrak{T}_0(m^{<\omega}) \to \mathfrak{T}_0(m^{<\omega})$ such that* $\operatorname{rng} g$ *is $\approx_0$-homogeneous with respect to c.*

The proof of the Theorem of Ramsey was by induction on the length of tuples. We prove the Theorem of Milliken by a similar argument where the induction is on the number of levels of a tuple. The next lemma contains the inductive step of this argument. It is based on the following variant of the relation $\approx_n$.

Definition 2.6. Let $k, n < \omega$. For $\bar\eta, \bar\zeta \subseteq m^{<\omega}$, we set

$$\bar\eta \approx_{n,k} \bar\zeta \quad :\text{iff} \quad \bar\eta = \bar\zeta\,, \text{ or}$$
$$\bar\eta \approx_n \bar\zeta \text{ and } |\mathrm{Lvl}(\bar\eta) \smallsetminus [n]|,\ |\mathrm{Lvl}(\bar\zeta) \smallsetminus [n]| \leq k\,,$$

and we denote by $\approx_{\omega,k}$ the transitive closure of the union $\bigcup_{n<\omega} \approx_{n,k}$.

Remark. (a) Note that

$$\bar\eta \approx_{n,0} \bar\zeta \qquad \text{iff} \qquad \bar\eta = \bar\zeta\,,$$

and the fact that $|\mathrm{Lvl}(\bar\eta)| \leq 2|\bar\eta|$ implies that

$$\bar\eta \approx_{n,2|\bar\eta|} \bar\zeta \qquad \text{iff} \qquad \bar\eta \approx_n \bar\zeta\,.$$

(b) A set X is $\approx_{\omega,k}$-homogeneous if, and only if, it is $\approx_{n,k}$-homogeneous, for every $n < \omega$.

Lemma 2.7. *Let $m, d < \omega$, let C be a finite set, and let $c : (m^{<\omega})^d \to C$ be a function such that $m^{<\omega}$ is $\approx_{\omega,k}$-homogeneous with respect to c. For every $n < \omega$, there exists an embedding*

$$g : \mathfrak{T}_{n+1}(m^{<\omega}) \to \mathfrak{T}_{n+1}(m^{<\omega})$$

such that $\operatorname{rng} g$ *is $\approx_{n,k+1}$-homogeneous with respect to c.*

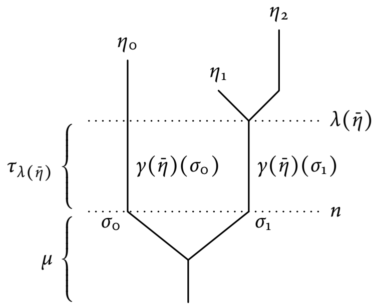


Figure 1.. The definition of μ, λ, τ_l, and γ.

Proof. Given $n < \omega$, set

$$\Gamma := \left\{ \bar\eta \in (m^{<\omega})^d \setminus (m^{<n})^d \;\middle|\; |\mathrm{Lvl}(\bar\eta) \setminus [n]| \leq k + 1 \right\}.$$

For $\bar\eta \in \Gamma$, let

$$\lambda(\bar\eta) := \min(\mathrm{Lvl}(\bar\eta) \setminus [n]).$$

Set $L := m^n$ and let

$$\mu : m^{<\omega} \setminus m^{<n} \to L : \eta \mapsto \eta \restriction n$$

be the function mapping each element to its prefix of length n. For $l \geq n$, let $\tau_l : m^{<\omega} \setminus m^{<l} \to m^{l-n}$ be the function mapping an element $\eta \in m^{<\omega}$ of length $|\eta| \geq l$ to the unique sequence $\sigma \in m^{<\omega}$ such that

$$|\sigma| = l - n \quad \text{and} \quad \mu(\eta)\,\sigma \leq \eta\,.$$

Let H be the set of all functions $h : L \to m^{<\omega}$ such that

$$|h(\rho)| = |h(\sigma)|\,, \quad \text{for all } \rho, \sigma \in L\,.$$

For $h, h' \in H$ and $\bar\eta \in \Gamma$, we set

$$h \sim_{\bar\eta} h' \quad :\text{iff} \quad h(\mu(\eta_i)) = h'(\mu(\eta_i)), \text{ for all } i < d \text{ with } |\eta_i| \geq n.$$

We define a function $\gamma : \Gamma \to H : \bar\eta \mapsto h_{\bar\eta}$ where

$$h_{\bar\eta}(\sigma) := \begin{cases} \tau_{\lambda(\bar\eta)}(\eta_i) & \text{if } \eta_i \in \mu^{-1}(\sigma), \\ \langle 0, \dots, 0 \rangle & \text{otherwise}. \end{cases}$$

Note that, in the first case of the definition of $h_{\bar\eta}(\sigma)$, the value does not depend on the choice of $i < d$ since

$$\eta_i, \eta_j \in \mu^{-1}(\sigma) \quad \text{implies} \quad \tau_{\lambda(\bar\eta)}(\eta_i) = \tau_{\lambda(\bar\eta)}(\eta_j).$$

Finally, we define a function $\beta : H \times \Gamma/{\approx_n} \to C$ by

$$\beta(h, [\bar\eta]_{\approx_n}) := c(\bar a[\bar\zeta]) \quad \text{where } \bar\zeta \in \gamma^{-1}\big[[h]_{\sim_{\bar\eta}}\big] \cap [\bar\eta]_{\approx_n}.$$

To prove that β is well-defined, we have to check that

$$\gamma^{-1}\big[[h]_{\sim_{\bar\eta}}\big] \cap [\bar\eta]_{\approx_n} \neq \emptyset$$

and that the value of β does not depend on the choice of $\bar\zeta$.

For non-emptiness, fix h and $[\bar\eta]_{\approx_n}$. For $i < d$ with $|\eta_i| \geq n$, let $\rho_i \in m^{<\omega}$ be the sequence such that

$$\eta_i = \mu(\eta_i)\,\tau_{\lambda(\bar\eta)}(\eta_i)\,\rho_i.$$

We set

$$\zeta_i := \mu(\eta_i)\,h(\mu(\eta_i))\,\rho_i.$$

For $i < d$ with $|\eta_i| < n$, we set $\zeta_i := \eta_i$. Then $\bar\zeta \approx_n \bar\eta$ and, since we have

$$\lambda(\bar\zeta) = n + |h(\mu(\eta_i))|, \quad \text{for any } i < d \text{ with } |\eta_i| \geq n,$$

it also follows that $\gamma(\bar{\zeta}) \sim_{\bar{\eta}} h$. Hence, $\bar{\zeta} \in \gamma^{-1}[[h]_{\sim_{\bar{\eta}}}] \cap [\bar{\eta}]_{\approx_n}$.

To show that the value of $\beta(h, [\bar{\eta}]_{\approx_n})$ does not depend on the choice of $\bar{\zeta}$, consider two tuples $\bar{\xi}, \bar{\zeta} \in \gamma^{-1}[[h]_{\sim_{\bar{\eta}}}] \cap [\bar{\eta}]_{\approx_n}$. First of all, note that $\bar{\xi} \approx_n \bar{\zeta}$ implies that $\mu(\xi_i) = \mu(\zeta_i)$, for all i with $|\xi_i| \geq n$, since

$$\sigma <_p \xi_i \quad \text{iff} \quad \sigma <_p \zeta_i\,, \quad \text{for all } \sigma \in m^{n-1} \text{ and all } p < m\,.$$

(For $n = 0$, we have $\mu(\xi_i) = \langle\rangle = \mu(\zeta_i)$, for all i.) Consequently, $\gamma(\bar{\xi}) \sim_{\bar{\eta}} h \sim_{\bar{\eta}} \gamma(\bar{\zeta})$ implies that

$$\tau_{\lambda(\bar{\xi})}(\xi_i) = h(\mu(\xi_i)) = h(\mu(\zeta_i)) = \tau_{\lambda(\bar{\zeta})}(\zeta_i)\,,$$

for all $i < d$ with $|\xi_i| \geq n$. In particular, $\lambda(\bar{\xi}) = \lambda(\bar{\zeta}) =: l$ and

$$\xi_i \upharpoonright l = \mu(\xi_i)\,\tau_l(\xi_i) = \mu(\zeta_i)\,\tau_l(\zeta_i) = \zeta_i \upharpoonright l\,.$$

As $\bar{\xi} \approx_n \bar{\zeta}$ it follows that $\bar{\xi} \approx_{l+1} \bar{\zeta}$. Since

$$|\mathrm{Lvl}(\bar{\xi}) \smallsetminus [l+1]| = |\mathrm{Lvl}(\bar{\zeta}) \smallsetminus [l+1]| \leq k\,,$$

we, therefore, have $\bar{\xi} \approx_{l+1,k} \bar{\zeta}$ and, by assumption on c, it follows that $c(\bar{\xi}) = c(\bar{\zeta})$, as desired.

To conclude the proof, consider the function $c_0 : H \to C^{\Gamma/\approx_n}$ mapping a tuple $h \in H$ to the function $[\bar{\eta}]_{\approx_n} \mapsto \beta(h, [\bar{\eta}]_{\approx_n})$, and let $c_1 : (m^{<\omega})^L \to C^{\Gamma/\approx_n}$ be an arbitrary extension of c_0.

Since $C^{\Gamma/\approx_n}$ is a finite set, we can use the Theorem of Halpern and Läuchli to obtain embeddings $g_\sigma : \mathfrak{T}_0(m^{<\omega}) \to \mathfrak{T}_0(m^{<\omega})$, for $\sigma \in L$, such that all g_σ have the same level embedding function and the restriction $c_1 \upharpoonright H \cap \prod_{\sigma\in L} \operatorname{rng} g_\sigma$ is constant. We can define the desired embedding $g : \mathfrak{T}_{n+1}(m^{<\omega}) \to \mathfrak{T}_{n+1}(m^{<\omega})$ by setting

$$g(\eta) := \begin{cases} \eta & \text{if } |\eta| \leq n\,, \\ \sigma g_\sigma(\xi) & \text{if } \eta = \sigma\xi \text{ for } \sigma \in L \text{ and } \xi \in m^{<\omega}\,. \end{cases}$$

It remains to prove that $\operatorname{rng} g$ is $\approx_{n,k+1}$-homogeneous with respect to c. Let $\bar{\eta}, \bar{\zeta} \in \Gamma \cap (\operatorname{rng} g)^d$ be tuples with $\bar{\eta} \approx_n \bar{\zeta}$. To show that $c(\bar{\eta}) = c(\bar{\zeta})$, set $h := \gamma(\bar\eta)$ and $h' := \gamma(\bar\zeta)$. For each $\sigma \in L$, fix some $\xi_\sigma \in \operatorname{rng} g_\sigma$ and set

$$h_0(\sigma) := \begin{cases} h(\sigma) & \text{if } \sigma \leq \eta_i \text{ for some } i\,, \\ \xi_\sigma & \text{otherwise}\,. \end{cases}$$

Then $h_0 \in \prod_{\sigma\in L} \operatorname{rng} g_\sigma$ and $h_0 \sim_{\bar\eta} h$. Similarly, we can find some $h'_0 \in \prod_{\sigma\in L} \operatorname{rng} g_\sigma$ with $h'_0 \sim_{\bar\zeta} h'$. Since $c_0(h_0) = c_0(h'_0)$ and $[\bar\eta]_{\approx_n} = [\bar\zeta]_{\approx_n}$ it follows that

$$\begin{aligned} c(\bar\eta) = \beta(h, [\bar\eta]_{\approx_n}) &= \beta(h_0, [\bar\eta]_{\approx_n}) \\ &= c_0(h_0)([\bar\eta]_{\approx_n}) \\ &= c_0(h'_0)([\bar\eta]_{\approx_n}) \\ &= c_0(h'_0)([\bar\zeta]_{\approx_n}) \\ &= \beta(h'_0, [\bar\zeta]_{\approx_n}) = \beta(h', [\bar\zeta]_{\approx_n}) = c(\bar\zeta)\,. \end{aligned}$$ □

Lemma 2.8. *Let $m, d < \omega$, let C be a finite set, and let $c : (m^{<\omega})^d \to C$ be a function such that $m^{<\omega}$ is $\approx_{\omega,k}$-homogeneous with respect to c. There exists an embedding $g : \mathfrak{T}_0(m^{<\omega}) \to \mathfrak{T}_0(m^{<\omega})$ such that* rng g *is $\approx_{\omega,k+1}$-homogeneous with respect to c.*

Proof. To simplify notation, we write $c \circ g$ for the function mapping a tuple $\bar\eta \in (m^{<\omega})^d$ to the value $c(g(\eta_0), \dots, g(\eta_{d-1}))$. We construct a sequence of embeddings

$$g_n : \mathfrak{T}_n(m^{<\omega}) \to \mathfrak{T}_n(m^{<\omega})\,, \quad \text{for } n < \omega\,,$$

such that, for all $i < n < \omega$, the set $m^{<\omega}$ is $\approx_{i,k+1}$-homogeneous with respect to the function $c_n := c \circ g_0 \circ \dots\dots g_n$.

We start with $g_0 :=$ id. Then $c_0 = c$ trivially satisfies the above condition. For the inductive step, suppose that we have already found functions $g_0, \dots, g_n$ such that, for every $i < n$, $m^{<\omega}$ is $\approx_{i,k+1}$-homogeneous with respect to c_n. We can use Lemma 2.7 to find an embedding g_{n+1} :

$\mathfrak{T}_{n+1}(m^{<\omega}) \to \mathfrak{T}_{n+1}(m^{<\omega})$ such that $m^{<\omega}$ is $\approx_{n,k+1}$-homogeneous with respect to $c_n \circ g_{n+1} = c_{n+1}$. Furthermore, since $m^{<\omega}$ is $\approx_{i,k+1}$-homogeneous with respect to c_n, for all $i < n$, it follows that it is also $\approx_{i,k+1}$-homogeneous with respect to $c_n \circ g_{n+1}$.

Having constructed the sequence $g_0, g_1, \dots$ we obtain the desired embedding $g : \mathfrak{T}_0(m^{<\omega}) \to \mathfrak{T}_0(m^{<\omega})$ as follows. For $\eta \in m^n$, we set $g(\eta) := (g_0 \circ \cdots \circ g_{n+1})(\eta)$. Clearly, g is an embedding. Hence, it remains to prove that rng g is $\approx_{\omega,k+1}$-homogeneous. Fix n and consider two tuples $\bar\eta, \bar\zeta \subseteq m^{<\omega}$ such that

$$\bar\eta \approx_n \bar\zeta \quad \text{and} \quad \big|\mathrm{Lvl}(\bar\eta) \smallsetminus [n]\big|,\ \big|\mathrm{Lvl}(\bar\zeta) \smallsetminus [n]\big| \leq k+1\,.$$

Choose $n < l < \omega$ such that $\bar\eta, \bar\zeta \subseteq m^{<l}$. Then

$$g(\bar\eta) = (g_0 \circ \cdots \circ g_l)(\bar\eta) \quad \text{and} \quad g(\bar\zeta) = (g_0 \circ \cdots \circ g_l)(\bar\zeta)\,.$$

As $\mathrm{rng}(g_0 \circ \cdots \circ g_l)$ is $\approx_{n,k+1}$-homogeneous with respect to c, it follows that $c(g(\bar\eta)) = c(g(\bar\zeta))$. □

Proof of Theorem 2.5. Note that, for every $n < \omega$, the set $m^{<\omega}$ is $\approx_{n,0}$-homogeneous with respect to c. Hence, repeating Lemma 2.8 we obtain embeddings

$$g_k : \mathfrak{T}_0(m^{<\omega}) \to \mathfrak{T}_0(m^{<\omega})\,, \quad \text{for } k \leq 2d\,,$$

such that $\mathrm{rng}(g_0 \circ \cdots \circ g_k)$ is $\approx_{\omega,k}$-homogeneous with respect to c. Setting $g := g_0 \circ \cdots \circ g_{2d}$ it follows that rng g is $\approx_{0,2d}$-homogeneous with respect to c. Since $|\mathrm{Lvl}(\bar\eta)| \leq 2d$, for all $\bar\eta \in (m^{<\omega})^d$, this is the same as saying that rng g is $\approx_0$-homogeneous with respect to c. □

As for the Theorem of Ramsey, the Theorem of Milliken also has a finitary version. The proof follows exactly the same lines as that of Theorem 1.4.

Theorem 2.9. *Let $m, d, k < \omega$ and let C be a finite set. There exists a number $n < \omega$ such that, for every function $c : (m^{<n})^d \to C$, there exists an embedding $g : \mathfrak{T}_0(m^{<k}) \to \mathfrak{T}_0(m^{<n})$ such that* rng g *is $\approx_0$-homogeneous with respect to c.*

Proof. For a contradiction, suppose that there exists no number n as above. For $n < \omega$, let F_n be the set of all functions $c : (m^{<n})^d \to C$ such that there is no embedding $g : \mathfrak{T}_0(m^{<k}) \to \mathfrak{T}_0(m^{<n})$ such that $\operatorname{rng} g$ is $\approx_0$-homogeneous with respect to c. Each set F_n is finite and nonempty. Furthermore, $c \in F_{n+1}$ implies that $c \restriction (m^{<n})^d \in F_n$. Hence, if we order the set $T := \bigcup_n F_n$ by inclusion, we obtain a tree $\langle T, \subseteq \rangle$. This tree is infinite and finitely branching. By the Lemma of Kőnig it therefore contains an infinite branch $(c_n)_{n<\omega}$ where $c_n \in F_n$. Set $c := \bigcup_n c_n$. Then c is a function $c : (m^{<\omega})^d \to C$. By Theorem 2.5, there exists an embedding $g : \mathfrak{T}_0(m^{<\omega}) \to \mathfrak{T}_0(m^{<\omega})$ such that $\operatorname{rng} g$ is $\approx_0$-homogeneous with respect to c. Fix a number $n < \omega$ such that $\operatorname{rng}(g \restriction m^{<k}) \subseteq m^{<n}$. Then $g \restriction m^{<k} : \mathfrak{T}_0(m^{<k}) \to \mathfrak{T}_0(m^{<n})$ is an embedding such that $\operatorname{rng} g$ is $\approx_0$-homogeneous with respect to c_n. A contradiction. □

Note that every $\approx_*$-homogeneous set is also $\approx_0$-homogeneous. Hence, we would obtain a stronger version of the Theorem of Milliken if we could replace the relation $\approx_0$ by $\approx_*$. For the finitary version this is possible.

Theorem 2.10. *Let $m, d, k < \omega$ and let C be a finite set. There exists a number $n < \omega$ such that, for every function $c : (m^{<n})^d \to C$, there exists an embedding $g : \mathfrak{T}_*(m^{<k}) \to \mathfrak{T}_*(m^{<n})$ such that* $\operatorname{rng} g$ *is $\approx_*$-homogeneous with respect to c.*

The proof consists in finding sets where the relations $\approx_*$ and $\approx_0$ coincide. To do so we introduce the following family of embeddings.

Definition 2.11. For $0 < k < \omega$, the *k-th skew embedding*

$$h_k : \mathfrak{T}_*(m^{<k}) \to \mathfrak{T}_*(m^{<l(k)})$$

is defined inductively as follows. We start with $h_1 : \langle\rangle \mapsto \langle\rangle$ and $l(1) = 1$. If h_k and $l(k)$ are already defined, we set

$$h_{k+1}(\langle\rangle) := \langle\rangle \quad \text{and} \quad h_{k+1}(p\eta) := \langle \underbrace{p, \dots, p}_{p+2+pl(k) \text{ times}} \rangle h_k(\eta),$$

for $\eta \in m^{<\omega}$ and $p < m$. Furthermore, $l(k+1) := ml(k) + m + 1$.

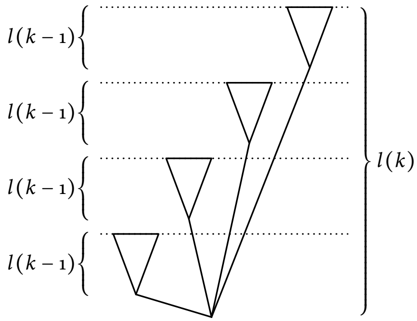


Figure 2.. The k-th skew embedding h_k.

Lemma 2.12. *The k-th skew embedding* $h_k : \mathfrak{T}_*(m^{<k}) \to \mathfrak{T}_*(m^{<l(k)})$ *is an embedding.*

Proof. By an easy induction on $|\eta|$, one can show that

$$\eta \leq \zeta \quad \text{implies} \quad h_k(\eta) \leq h_k(\zeta)\,,$$
$$\text{and} \quad \eta \prec_p \zeta \quad \text{implies} \quad h_k(\eta) \prec_p h_k(\zeta)\,.$$

Similarly, an induction on $|\eta \sqcap \zeta|$ yields

$$h_k(\eta \sqcap \zeta) = h_k(\eta) \sqcap h_k(\zeta)\,. \qquad \square$$

A useful property of a skew embedding is that it upgrades $\approx_*$-equivalence to $\approx_0$-equivalence.

Lemma 2.13. *Let* $\bar{\eta}, \bar{\zeta} \subseteq m^{<k}$. *Then* $\bar{\eta} \approx_* \bar{\zeta}$ *implies* $h_k(\bar{\eta}) \approx_0 h_k(\bar{\zeta})$.

Proof. Let $\bar{\eta}, \bar{\zeta} \in (m^{<k})^d$ with $\bar{\eta} \approx_* \bar{\zeta}$. We start by proving the following

claims:

(a) $h_k(\bar\eta) \approx_* h_k(\bar\zeta)$.

(b) $|h_k(\eta_i)| < |h_k(\eta_j)|$ iff $|h_k(\zeta_i)| < |h_k(\zeta_j)|$, for all $i, j < d$.

(c) $\mathrm{pf}(h_k(\eta_i), h_k(\eta_j)) \prec_p h_k(\eta_j)$
iff $\mathrm{pf}(h_k(\zeta_i), h_k(\zeta_j)) \prec_p h_k(\zeta_j)$, for all $i, j < d$.

(a) Since $h_k : \mathfrak{T}_*(m^{<k}) \to \mathfrak{T}_*(m^{<l(k)})$ is an embedding, it preserves atomic types. Consequently, we have $h_k(\bar\eta) \approx_* \bar\eta \approx_* \bar\zeta \approx_* h_k(\bar\zeta)$.

(b) It follows by induction on $|\eta_i \sqcap \eta_j|$ that

$$|h_k(\eta_i)| < |h_k(\eta_j)| \quad \text{iff} \quad \eta_i <_{\mathrm{lex}} \eta_j .$$

Hence, $\bar\eta \approx_* \bar\zeta$ implies that

$$\begin{aligned} |h_k(\eta_i)| < |h_k(\eta_j)| \quad &\text{iff} \quad \eta_i <_{\mathrm{lex}} \eta_j \\ &\text{iff} \quad \zeta_i <_{\mathrm{lex}} \zeta_j \quad \text{iff} \quad |h_k(\zeta_i)| < |h_k(\zeta_j)| . \end{aligned}$$

(c) By definition of h_k, we have

$$\mathrm{pf}(h_k(\eta_i), h_k(\eta_j)) \prec_p h_k(\eta_j)$$

$$\text{iff} \quad |h_k(\eta_i)| < |h_k(\eta_j)| \text{ and } h_k(\eta_i \sqcap \eta_j) \prec_p h_k(\eta_j) .$$

Therefore, (c) follows from (a) and (b).

To conclude the proof, suppose that $\bar\eta \approx_* \bar\zeta$. W.l.o.g. we may assume that, for all $i, j < d$, there is some $l < d$ such that $\eta_l = \eta_i \sqcap \eta_j$. Then it follows by (a), (b), and (c) that $h_k(\bar\eta) \approx_0 h_k(\bar\zeta)$. □

Proof of Theorem 2.10. Let $h_k : m^{<k} \to m^{<l(k)}$ be the k-th skew embedding. By Theorem 2.9, there exists a number n such that, for every function $c : (m^{<n})^d \to C$, we can find an embedding $g : \mathfrak{T}_0(m^{<l(k)}) \to \mathfrak{T}_0(m^{<n})$ such that $\operatorname{rng} g$ is $\approx_0$-homogeneous with respect to c. We

claim that $g \circ h_k : \mathfrak{T}_*(m^{<k}) \to \mathfrak{T}_*(m^{<n})$ is the desired embedding. For $\bar{\eta}, \bar{\zeta} \in (m^{<k})^d$ it follows by Lemma 2.13 that

$$\begin{aligned} \bar{\eta} \approx_* \bar{\zeta} \quad &\Rightarrow \quad h_k(\bar{\eta}) \approx_\circ h_k(\bar{\zeta}) \\ &\Rightarrow \quad g(h_k(\bar{\eta})) \approx_\circ g(h_k(\bar{\zeta})) \\ &\Rightarrow \quad c(g(h_k(\bar{\eta}))) = c(g(h_k(\bar{\zeta}))) . \end{aligned}$$

Hence, $\mathrm{rng}(g \circ h_k)$ is $\approx_*$-homogeneous with respect to c. □

3. *Indiscernible sequences*

If we apply the Ramsey Theorem to sequences of elements in a structure coloured by their types we obtain subsequences where each tuple has the same type. Such sequences, called *indiscernible,* can be used to investigate the structure of the given model. Let us fix some notation.

Definition 3.1. Let $\langle I, \leq \rangle$ be a linear order and $(\bar{a}^i)_{i \in I}$ a sequence of tuples $\bar{a}^i \in A^\alpha$, for some ordinal α.

(a) For $\bar{\imath} \in I^n$, we set $\bar{a}[\bar{\imath}] := \bar{a}^{i_0} \dots \bar{a}^{i_{n-1}}$.

(b) The *order type* of a tuple $\bar{\imath} \in I^n$ is the atomic type of $\bar{\imath}$ in $\langle I, \leq \rangle$.

Definition 3.2. Suppose that X and Y are disjoint sets of variables and $\Delta \subseteq \mathrm{FO}[\Sigma, X \cup Y]$ a set of formulae. Let $\mathfrak{M}$ be a Σ-structure, $U \subseteq M$, and $(\bar{a}^i)_{i \in I}$ a sequence of tuples in M.

(a) The *Δ-type* of a tuple $\bar{b} \subseteq M$ over U is the set

$$\mathrm{tp}_\Delta(\bar{b}/U) := \big\{ \varphi(\bar{x}; \bar{c}) \bigm| \mathfrak{M} \vDash \varphi(\bar{b}; \bar{c}),\ \bar{c} \subseteq U,\ \varphi(\bar{x}, \bar{y}) \in \Delta, \\ \bar{x} \subseteq X,\ \bar{y} \subseteq Y \big\}$$

(b) We call $(\bar{a}^i)_{i \in I}$ a *Δ-indiscernible sequence* over U, or a *sequence of Δ-indiscernibles,* if

$$\mathrm{tp}_\Delta(\bar{a}[\bar{\imath}]/U) = \mathrm{tp}_\Delta(\bar{a}[\bar{k}]/U) , \quad \text{for all } \bar{\imath}, \bar{k} \in [I]^{<\omega} .$$

For $\Delta = \mathrm{FO}[\Sigma, X \cup Y]$ we drop the Δ and simply speak of *indiscernible sequences.*

(c) The sequence $(\bar{a}^i)_i$ is *totally Δ-indiscernible* over U if

$$\mathrm{tp}_\Delta(\bar{a}[\bar{\imath}]/U) = \mathrm{tp}_\Delta(\bar{a}[\bar{k}]/U)\,,$$

for all finite sequences $\bar{\imath}, \bar{k} \in I^{<\omega}$ of distinct elements with $|\bar{\imath}| = |\bar{k}|$.

Example. (a) If $\langle A, < \rangle$ is an open dense linear order then every strictly increasing sequence $(a^i)_{i \in I}$ in A is indiscernible. Such a sequence is obviously not totally indiscernible.

(b) Let $\mathfrak{K}$ be an algebraically closed field. Every sequence of algebraically independent elements is totally indiscernible. Similarly, if $\mathfrak{V}$ is a vector space then every sequence of linearly independent elements is totally indiscernible.

For finite sets Δ, we can use the Ramsey Theorem to show that every infinite sequence contains a Δ-indiscernible subsequence. For infinite Δ, we need to apply the Compactness Theorem to find Δ-indiscernible sequences.

Lemma 3.3. *Let $(\bar{a}^i)_{i \in I}$ be an infinite sequence. For every finite set Δ of formulae there exists an infinite subset $I_0 \subseteq I$ such that $(\bar{a}^i)_{i \in I_0}$ is Δ-indiscernible.*

Proof. Let n be the maximal number such that Δ contains a formula $\varphi(\bar{x}^0, \dots, \bar{x}^{n-1})$ with n tuples of variables. We define a colouring $c : [I]^n \to \mathscr{P}(\Delta)$ by

$$c(\bar{\imath}) := \left\{ \varphi(\bar{x}^0, \dots, \bar{x}^{n-1}) \in \Delta \;\middle|\; \mathbb{M} \models \varphi(\bar{a}[\bar{\imath}]) \right\}.$$

By the Ramsey Theorem there exists an infinite subset $I_0 \subseteq I$ that is homogeneous with respect to c. By definition of c it follows that $(\bar{a}^i)_{i \in I_0}$ is Δ-indiscernible. ☐

To find Δ-indiscernible sequences, for infinite sets Δ, we apply the Compactness Theorem. Before doing so, let us introduce the average type of a sequence.

Definition 3.4. The *average type* of a sequence $(\bar{a}^i)_i$ over U is the set

$$\mathrm{Av}((\bar{a}^i)_i/U) := \{\, \varphi(\bar{x}^0, \dots, \bar{x}^{n-1}; \bar{c}) \mid \bar{c} \subseteq U \text{ and } \mathbb{M} \vDash \varphi(\bar{a}[\bar{\imath}]; \bar{c}) \text{ for all } \bar{\imath} \in [I]^n \,\}\,.$$

Lemma 3.5. *Let $(\bar{a}^i)_{i \in I}$ be a sequence. Then $\mathrm{Av}((\bar{a}^i)_i/U)$ is a partial type. If $(\bar{a}^i)_i$ is indiscernible over U, it is complete.*

Proposition 3.6. *Let $\mathfrak{M}$ be a Σ-structure and $U \subseteq M$ a set of parameters. For every infinite sequence $(\bar{a}^i)_{i \in I}$ and every linear order J there exists an elementary extension $\mathfrak{N} \succeq \mathfrak{M}$ containing an indiscernible sequence $(\bar{b}^j)_{j \in J}$ over U such that*

$$\mathrm{Av}((\bar{a}^i)_i/U) \subseteq \mathrm{Av}((\bar{b}^j)_j/U)\,.$$

Proof. For every $j \in J$, fix a tuple of new constant symbols $\bar{c}^j$ and set

$$\Phi := \big\{\, \varphi(\bar{c}[\bar{\jmath}]; \bar{d}) \;\big|\; \varphi(\bar{x}; \bar{d}) \in \mathrm{Av}((\bar{a}^i)_i/U),\ \bar{\jmath} \in [J]^{<\omega},\ \bar{d} \subseteq U \,\big\}$$
$$\Psi := \big\{\, \psi(\bar{c}[\bar{\imath}]; \bar{d}) \leftrightarrow \psi(\bar{c}[\bar{\jmath}]; \bar{d}) \;\big|\; \psi \text{ a formula, } \bar{\imath}, \bar{\jmath} \in [J]^{<\omega}, \text{ and } \bar{d} \subseteq U \,\big\}\,.$$

It is sufficient to prove that the set $\Gamma := \mathrm{Th}(\mathfrak{M}_M) \cup \Phi \cup \Psi$ is satisfiable. Consider a finite subset $\Gamma_0 \subseteq \Gamma$. Since $\mathrm{Th}(\mathfrak{M}_M)$ is closed under conjunctions, we may assume that $\Gamma_0 = \{\vartheta(\bar{d})\} \cup \Phi_0 \cup \Psi_0$ for finite sets $\Phi_0 \subseteq \Phi$ and $\Psi_0 \subseteq \Psi$. By Lemma 3.3, there is an infinite subset $I_0 \subseteq I$ such that we have

$$\mathfrak{M} \vDash \psi(\bar{a}[\bar{\imath}]; \bar{d}) \leftrightarrow \psi(\bar{a}[\bar{\jmath}]; \bar{d})\,,$$

for every formula $\psi(\bar{x}; \bar{d}) \leftrightarrow \psi(\bar{y}; \bar{d}) \in \Psi_0$ and all increasing $\bar{\imath}, \bar{\jmath} \subseteq I_0$. For every formula $\varphi(\bar{x}; \bar{d}) \in \Phi_0$, there are only finitely many indices $\bar{\imath} \subseteq I_0$ such that $\mathfrak{M} \nvDash \varphi(\bar{a}[\bar{\imath}]; \bar{d})$. Hence, we can find an infinite subset $I_1 \subseteq I_0$ containing no such tuple $\bar{\imath}$. Let $J_0 \subseteq J$ be the finite set of all indices $j \in J$ such that the constant $\bar{c}^j$ appears in $\Phi_0 \cup \Psi_0$, and fix an embedding $g : J_0 \to I_1$. We can satisfy Γ_0 by interpreting $\bar{c}^j$ by the tuple $\bar{a}^{g(j)}$. □

We can improve the preceding proposition as follows.

Theorem 3.7. *Let $\mathfrak{M}$ be a Σ-structure, $U \subseteq M$ a set of parameters, $\bar{s}$ a sequence of sorts, and λ a cardinal such that $\lambda \geq |S^{\bar{s}^n}(U)|$, for all $n < \omega$. Set $\mu := \beth_{\lambda^+}$.*

For every sequence $(\bar{a}^\alpha)_{\alpha<\mu}$ with $\bar{a}^\alpha \in M^{\bar{s}}$ and for every linear order I, there exists an elementary extension $\mathfrak{N} \succeq \mathfrak{M}$ containing an indiscernible sequence $(\bar{b}^i)_{i \in I}$ over U such that, for every $\bar{\imath} \in [I]^n$, there are indices $\bar{\alpha} \in [\mu]^n$ with

$$\operatorname{tp}(\bar{b}[\bar{\imath}]/U) = \operatorname{tp}(\bar{a}[\bar{\alpha}]/U)\,.$$

Proof. It is sufficient to prove the claim for $I = \omega$. Then the general statement will follow by compactness. We define a sequence of types $(\mathfrak{p}_n)_{n<\omega}$ with $\mathfrak{p}_n \in S^{\bar{s}^n}(U)$ satisfying the following conditions:

(1) $\mathfrak{p}_n(\bar{x}_0, \dots, \bar{x}_{n-1}) \vDash \mathfrak{p}_m(\bar{x}_{i_0}, \dots, \bar{x}_{i_{m-1}})$, for all $i_0 < \dots < i_{m-1} < n$.

(2) For every cardinal $\nu < \mu$, there is some set $I \subseteq \mu$ of size $|I| = \nu$ such that

$$\operatorname{tp}(\bar{a}[\bar{\imath}]/U) = \mathfrak{p}_n\,, \quad \text{for every tuple } \bar{\imath} \in [I]^n\,.$$

Any sequence $(\bar{b}^n)_{n<\omega}$ realising the limit $\mathfrak{p}_\omega := \bigcup_{n<\omega} \mathfrak{p}_n$ has the desired properties.

We start with $\mathfrak{p}_0 := \operatorname{Th}(\mathfrak{M}_U)$. If we have already defined $\mathfrak{p}_n$, we consider the set X of all $\bar{s}^{n+1}$-types over U satisfying condition (1). If there is some type $\mathfrak{q} \in X$ that also satisfies condition (2), we are done. Suppose there is no such type. Then we can choose, for every $\mathfrak{q} \in X$, a cardinal $\nu_\mathfrak{q} < \mu$ such that no subset $I \subseteq \mu$ of size $\nu_\mathfrak{q}$ satisfies the above condition. Since $|X| \leq \lambda < \lambda^+ = \operatorname{cf} \mu$ it follows that

$$\nu_* := \lambda \oplus \sup\{\, \nu_\mathfrak{q} \mid \mathfrak{q} \in X \,\} < \mu\,.$$

By choice of ν_* there exists, for every $\mathfrak{q} \in X$ and all $I \subseteq \mu$ of size $|I| = \nu_*$, some increasing tuple $\bar{\imath} \in I^{n+1}$ such that $\operatorname{tp}(\bar{a}[\bar{\imath}]/U) \neq \mathfrak{q}$. Since $\nu_* < \mu = \beth_{\lambda^+}$ there is some ordinal $\alpha < \lambda^+$ with $\nu_* < \beth_\alpha$. Let $\rho := \beth_{\alpha+n+1}$. Then

$$\beth_n(\nu_*)^+ \leq \rho < \mu\,.$$

By choice of $\mathfrak{p}_n$ there is some set $I \subseteq \mu$ of size $|I| = \rho$ such that

$$\mathrm{tp}(\bar{a}[\bar{\imath}]/U) = \mathfrak{p}_n\,, \quad \text{for every } \bar{\imath} \in [I]^n.$$

Since $|S^{\bar{s}^n}(U)| \leq \lambda \leq \nu_*$ we can use the Theorem of Erdős and Rado to find a subset $I_0 \subseteq I$ of size $|I_0| = \nu_*^+$ such that the types

$$\mathrm{tp}(\bar{a}[\bar{\imath}]/U)\,, \quad \text{for } \bar{\imath} \in [I_0]^{n+1}\,,$$

are all equal. This contradicts the choice of ν_*. □

There is a close relationship between automorphisms and indiscernible sequences. The next observation follows immediately from the definitions of an indiscernible sequence and a strongly κ-homogeneous structure.

Lemma 3.8. *Let $\mathfrak{M}$ be strongly κ-homogeneous and let $(\bar{a}^i)_{i\in I}$ be a sequence of indiscernible over U. Suppose that $|U| \oplus |I| \oplus |\bar{a}^i| < \kappa$. For every partial automorphism $\pi \in \mathrm{pIso}(I, I)$ of the index set I (considered as a linear order), there exists an automorphism $h \in \mathrm{Aut}\,\mathfrak{M}$ such that*

$$h \restriction U = \mathrm{id}_U \quad \textit{and} \quad h(\bar{a}^i) = \bar{a}^{\pi(i)}\,, \quad \textit{for all } i \in I\,.$$

In a sufficient saturated structure, we can extend every indiscernible sequence to a longer one.

Lemma 3.9. *Let $\mathfrak{M}$ be κ-saturated. If $(\bar{a}^i)_{i\in I}$ is indiscernible over U and $g : I \to J$ is an embedding with $|J| \oplus |U| \oplus |\bar{a}^i| < \kappa$ then there exists an indiscernible sequence $(\bar{b}^j)_{j\in J}$ such that $\bar{a}^i = \bar{b}^{g(i)}$, for $i \in I$.*

Proof. We can use Proposition 3.6 to find an elementary extension $\mathfrak{N} \succeq \mathfrak{M}$ containing an indiscernible sequence $(\bar{c}^j)_{j\in J}$ with $\mathrm{Av}((\bar{c}^j)_j/U) = \mathrm{Av}((\bar{a}^i)_i/U)$. This implies that

$$\mathrm{tp}(\textstyle\bigcup_i \bar{c}^{g(i)}/U) = \mathrm{tp}(\bigcup_i \bar{a}^i/U)\,.$$

W.l.o.g. we may assume that $\mathfrak{N}$ is strongly κ-homogeneous. Therefore, there exists an automorphism π of $\mathfrak{N}_U$ mapping $\bar{c}^{g(i)}$ to $\bar{a}^i$. Since $\mathfrak{M}$ is κ-saturated it contains a sequence $(\bar{b}^j)_{j\in J}$ such that

$$\operatorname{tp}(\bigcup_j \bar{b}^j/U \cup \bigcup_i \bar{a}^i) = \operatorname{tp}(\bigcup_j \pi(\bar{c}^j)/U \cup \bigcup_i \bar{a}^i)\,.$$

It follows that $(\bar{b}^j)_j$ is the desired sequence of indiscernibles. □

Corollary 3.10. *If $(\bar{a}^i)_{i\in I}$ is indiscernible over U and $g : I \to J$ an embedding, then there exists an elementary extension $\mathfrak{N}$ containing an indiscernible sequence $(\bar{b}^i)_{i\in J}$ such that $\bar{b}^{g(i)} = \bar{a}^i$, for $i \in I$.*

Let us record the following consequence of Theorem 3.7.

Lemma 3.11. *Let $(\bar{a}_i)_{i\in I}$ be an indiscernible sequences over U. For every set $C \subseteq \mathbb{M}$, there exists a set $C' \equiv_U C$ such that $(\bar{a}_i)_{i\in I}$ is indiscernible over $U \cup C'$.*

Proof. Let $\kappa := |T| \oplus |U \cup C|$ and $\lambda := \beth_{(2^\kappa)^+}$. By Corollary 3.10, there exists an indiscernible sequence $(\bar{b}_\alpha)_{\alpha<\kappa}$ over U with

$$\operatorname{Av}((\bar{b}_\alpha)_\alpha/U) = \operatorname{Av}((\bar{a}_i)_i/U)\,.$$

Furthermore, with the help of Theorem 3.7 we can find an indiscernible sequence $(\bar{c}_n)_{n<\omega}$ over $U \cup C$ such that, for every $n < \omega$, there are indices $\alpha_0 < \dots < \alpha_{n-1}$ with

$$\bar{c}_0 \dots \bar{c}_{n-1} \equiv_{U\cup C} \bar{b}_{\alpha_0} \dots \bar{b}_{\alpha_{n-1}}\,.$$

By Lemma 3.9, we can extend $(\bar{c}_n)_{n<\omega}$ to an indiscernible sequence $(\bar{c}_i)_{i\in\omega+I}$ over $U \cup C$. Since

$$\operatorname{Av}((\bar{c}_i)_i/U) = \operatorname{Av}((\bar{a}_i)_i/U)\,,$$

there exists an automorphism $\pi \in \operatorname{Aut} \mathbb{M}_U$ such that $\pi(\bar{c}_{\omega+i}) = \bar{a}_i$, for all $i \in I$. Then $\pi[C] \equiv_U C$ and $(\bar{a}_i)_{i\in I}$ is indiscernible over $U \cup \pi[C]$. □

The following technical lemma can be used to simplify proofs of indiscernibility. It states that, if some formula is a witness for the failure of indiscernibility, we can detect this fact already by varying a single element of the sequence.

Lemma 3.12. *Let $\alpha = (\bar{a}_i)_{i \in I}$ be a sequence and $\varphi(\bar{x})$ a formula such that*

$$\mathbb{M} \vDash \varphi(\bar{a}[\bar{\imath}]) \wedge \neg\varphi(\bar{a}[\bar{\jmath}]) , \quad \text{for some } \bar{\imath}, \bar{\jmath} \in [I]^n .$$

Then there are indices $\bar{u} < s < t < \bar{v}$ in I such that

$$\mathbb{M} \vDash \varphi(\bar{a}[\bar{u}s\bar{v}]) \leftrightarrow \neg\varphi(\bar{a}[\bar{u}t\bar{v}]) .$$

Proof. We define a sequence $\bar{k}^0, \dots, \bar{k}^{2n} \in [I]^n$ by setting

$$k^l_m := \begin{cases} \min\{i_m, j_m\} & \text{if } l \leq n \text{ and } m < l , \\ i_m & \text{if } l \leq n \text{ and } m \geq l , \\ \min\{i_m, j_m\} & \text{if } l > n \text{ and } m < 2n - l , \\ j_m & \text{if } l > n \text{ and } m \geq 2n - l . \end{cases}$$

Then every $\bar{k}_l$ belongs to $[I]^n$, $\bar{k}_0 = \bar{\imath}$, $\bar{k}_{2n} = \bar{\jmath}$, and, for each $l < 2n$, the tuples $\bar{k}_l$ and $\bar{k}_{l+1}$ differ in at most one component. Let $l < 2n$ be the maximal index such that $\mathbb{M} \vDash \varphi(\bar{a}[\bar{k}_l])$. Then $\mathbb{M} \vDash \neg\varphi(\bar{a}[\bar{k}_{l+1}])$ and it follows by definition of $\bar{k}_l$ that $\bar{k}_l = \bar{u}s\bar{v}$ and $\bar{k}_{l+1} = \bar{u}t\bar{v}$ for indices $\bar{u} < s < \bar{v}$ and $\bar{u} < t < \bar{v}$. Interchanging $\bar{k}_l$ and $\bar{k}_{l+1}$ if necessary, we may assume that $s < t$. □

Recall that stable theories do not have the order property. This implies that in a model of a stable theory every indiscernible sequence is totally indiscernible.

Theorem 3.13. *A theory T is stable if, and only if, every infinite indiscernible sequence in a model of T is totally indiscernible.*

Proof. ($\Leftarrow$) Suppose that there is a formula $\varphi(\bar{x}, \bar{y})$ with the order property and let $(\bar{a}^n)_{n<\omega}$ and $(\bar{b}^n)_{n<\omega}$ be sequences such that

$$\mathbb{M} \vDash \varphi(\bar{a}^i, \bar{b}^k) \quad \text{iff} \quad i \leq k\,.$$

By Proposition 3.6, there exists an indiscernible sequence $(\bar{c}^n \bar{d}^n)_{n<\omega}$ with $\mathrm{Av}((\bar{a}^n \bar{b}^n)_n) \subseteq \mathrm{Av}((\bar{c}^n \bar{d}^n)_n)$. Setting $\psi(\bar{x}\bar{y}, \bar{x}'\bar{y}') := \varphi(\bar{x}, \bar{y}')$ it follows that

$$\mathbb{M} \vDash \psi(\bar{c}^i \bar{d}^i, \bar{c}^k \bar{d}^k) \quad \text{iff} \quad i \leq k\,.$$

Hence, $(\bar{c}^n \bar{d}^n)_n$ is not totally indiscernible.

($\Rightarrow$) Suppose that $(\bar{a}^i)_{i \in I}$ is an infinite indiscernible sequence over U that is not totally indiscernible. By Corollary 3.10, we may assume that the ordering I is dense. There are a formula φ and two tuples of indices $\bar{\imath}, \bar{k} \in I^n$ such that both $\bar{\imath}$ and $\bar{k}$ consist of distinct elements and we have

$$\mathbb{M} \vDash \varphi(\bar{a}[\bar{\imath}]) \wedge \neg\varphi(\bar{a}[\bar{k}])\,.$$

Set $\bar{l}^r := i_0 \dots i_{r-1} k_r \dots k_{n-1}$ and let r be the maximal number such that

$$\mathbb{M} \vDash \neg\varphi(\bar{a}[\bar{l}^r])\,.$$

Note that r is well-defined since $\bar{l}^0 = \bar{k}$ implies $\mathbb{M} \vDash \neg\varphi(\bar{a}[\bar{l}^0])$. Replacing $\bar{\imath}$ by $\bar{l}^{r+1}$ and $\bar{k}$ by $\bar{l}^r$, we may assume that $\bar{\imath}$ and $\bar{k}$ differ in exactly one component. Hence, suppose that

$$\bar{\imath} = s\bar{u}\bar{v}\bar{w} \quad \text{and} \quad \bar{k} = t\bar{u}\bar{v}\bar{w}\,, \quad \text{where } \bar{u} < s < \bar{v} < t < \bar{w}\,.$$

(Reversing the order of I, if necessary, we may assume that $s < t$.)

By indiscernibility, we know that the tuple $\bar{v}$ is not empty. We claim that we may assume that $\bar{v}$ is a singleton. If $\bar{v} = v_0 \dots v_{n-1}$ with $n > 1$ then, choosing some index $v_0 < v' < v_{n-1}$, we may replace either s or t by v', depending on whether or not the formula $\varphi(\bar{a}[v'\bar{u}\bar{v}\bar{w}])$ holds. Hence, the claim follows by induction. Thus, we have arrived at the situation that

$$\bar{\imath} = sv\bar{u}\bar{w} \quad \text{and} \quad \bar{k} = vt\bar{u}\bar{w}\,, \quad \text{where } \bar{u} < s < v < t < \bar{w}\,.$$

By indiscernibility, it follows that

$$\mathbb{M} \vDash \varphi(\bar{a}[st\bar{u}\bar{w}]) \wedge \neg\varphi(\bar{a}[ts\bar{u}\bar{w}]), \quad \text{for all} \quad \bar{u} < s < t < \bar{w}.$$

Fix an infinite increasing sequence of indices k_n, $n < \omega$, with

$$\bar{u} < k_0 < k_1 < \dots < \bar{w},$$

set $\bar{b}^i := \bar{a}^{k_i}$, and define

$$\psi(\bar{x}, \bar{y}) := \bar{x} = \bar{y} \vee [\varphi(\bar{x}, \bar{y}, \bar{a}[\bar{u}\bar{w}]) \wedge \neg\varphi(\bar{y}, \bar{x}, \bar{a}[\bar{u}\bar{w}])].$$

Then we have

$$\mathbb{M} \vDash \psi(\bar{b}^i, \bar{b}^k) \quad \text{iff} \quad i \leq k.$$

Hence, T is unstable. □

When considering the automorphism group of a structure, an indiscernible sequence looks like a linear order while a totally indiscernible sequence looks like a set. We can generalise the definition of an indiscernible sequence to include automorphism groups of other structures.

Definition 3.14. Let L be an algebraic logic, $\mathfrak{J}$ a Γ-structure, $\mathfrak{M}$ a Σ-structure, and $U \subseteq M$.

(a) A *U-indiscernible system* over $\mathfrak{J}$ (w.r.t. L) is an injective function $\bar{a} : I \to M^\alpha$, for some ordinal α, such that, for every partial isomorphism $\bar{\imath} \mapsto \bar{k} \in \mathrm{pIso}_{\aleph_0}(\mathfrak{J}, \mathfrak{J})$, we have

$$\mathrm{tp}_L(\bar{a}[\bar{\imath}]/U) = \mathrm{tp}_L(\bar{a}[\bar{k}]/U).$$

(b) The *average type* of a U-indiscernible system $\bar{a}$ over $\mathfrak{J}$ is the function $\mathrm{Av}_L(\bar{a})$ with

$$\mathrm{Av}_L(\bar{a}/U) : \mathrm{atp}(\bar{\imath}/\mathfrak{J}) \mapsto \mathrm{tp}_L(\bar{a}[\bar{\imath}]/U), \quad \text{for } \bar{\imath} \in I^{<\omega}.$$

For $L = \mathrm{FO}$, we drop the index and just write $\mathrm{Av}(\bar{a}/U)$.

(c) Let $\mathfrak{I}$ and $\mathfrak{K}$ be two index structures and $\bar{a} : I \to M^\alpha$, $\bar{b} : K \to M^\alpha$ arbitrary families of α-tuples. We say that $\bar{a}$ is *inspired* by $\bar{b}$ over U if, for every finite set of formulae Δ and every finite tuple $\bar{\imath} \in I^{<\omega}$, there is a finite tuple $\bar{k} \in K^{<\omega}$ such that

$$\operatorname{atp}(\bar{\imath}/\mathfrak{I}) = \operatorname{atp}(\bar{k}/\mathfrak{K}) \quad \text{and} \quad \operatorname{tp}_\Delta(\bar{a}[\bar{\imath}]/U) = \operatorname{tp}_\Delta(\bar{b}[\bar{k}]/U)\,.$$

Remark. (a) Using the terminology of the previous definition we can restate Proposition 3.6 as: for every infinite sequence $(\bar{a}^i)_{i\in I}$, every linear order J, and every set U of parameters, there exists an indiscernible sequence $(\bar{b}^i)_{i\in J}$ over U inspired by $(\bar{a}^i)_{i\in I}$.

(b) Note that, for indiscernible systems $\bar{a}$ and $\bar{b}$ over U, $\bar{a}$ is inspired by $\bar{b}$ over U if, and only if, $\operatorname{Av}(\bar{a}/U) = \operatorname{Av}(\bar{b}/U)$.

In the same way as in Proposition 3.6 we can use the Compactness Theorem to show that we can extend every indiscernible system.

Lemma 3.15. *Let $\mathfrak{M}$ be a structure containing a U-indiscernible system $\bar{a}$ over $\mathfrak{I}$. If $\mathfrak{H}$ is a structure with $\operatorname{Sub}_{\aleph_0}(\mathfrak{H}) \subseteq \operatorname{Sub}_{\aleph_0}(\mathfrak{I})$ then there exists an elementary extension $\mathfrak{N} \succeq \mathfrak{M}$ containing a U-indiscernible system $\bar{b}$ over $\mathfrak{H}$ with $\operatorname{Av}(\bar{b}/U) = \operatorname{Av}(\bar{a}/U)$.*

In general, it is hard to prove the existence of indiscernible systems over structures that are not linear orders. For trees we can use the Theorem of Milliken to show that such systems always exist. Recall the trees $\mathfrak{T}_*(\kappa^{<\alpha})$ introduced in Section 2.

Definition 3.16. Let κ be a cardinal and α an ordinal. A family $(\bar{a}_\eta)_{\eta\in\kappa^{<\alpha}}$ is called *tree-indiscernible* over a set U if it is a U-indiscernible system over $\mathfrak{T}_*(\kappa^{<\alpha})$.

Theorem 3.17 (Džamonja, Shelah, B. Kim, H.-J. Kim)**.** *Let $m < \omega$. For every family $\bar{a} = (\bar{a}_\eta)_{\eta\in m^{<\omega}}$ and every set U, there exists a family of tree-indiscernibles $(\bar{b}_\eta)_{\eta\in m^{<\omega}}$ over U inspired by $\bar{a}$.*

Proof. Fix variable symbols $\bar{x}_\eta$, for each $\eta \in m^{<\omega}$, and define

$$\Psi_{\bar{\eta}} := \big\{ \varphi(\bar{x}[\bar{\zeta}]) \bigm| \varphi \text{ a formula over } U\,,\ \bar{\zeta} \approx_* \bar{\eta}\,,\ \text{and } \mathbb{M} \vDash \varphi(\bar{a}[\bar{\xi}]) \text{ for all } \bar{\xi} \approx_* \bar{\eta} \big\}\,,$$

$$\Xi := \big\{ \varphi(\bar{x}[\bar{\eta}]) \leftrightarrow \varphi(\bar{x}[\bar{\zeta}]) \bigm| \varphi \text{ a formula over } U\,, \bar{\eta} \approx_* \bar{\zeta} \big\}\,,$$

and $\quad \Phi := \Xi \cup \bigcup_{\bar{\eta} \subseteq m^{<\omega}} \Psi_{\bar{\eta}}\,.$

We claim that Φ is satisfiable. Let $\Phi_0 \subseteq \Phi$ be finite. There exists a finite set Δ of formulae such that every formula in Φ_0 is of the form

$$\varphi(\bar{x}[\bar{\eta}]) \leftrightarrow \varphi(\bar{x}[\bar{\zeta}]) \quad \text{or} \quad \varphi(\bar{x}[\bar{\zeta}])\,,$$

for some $\varphi(\bar{x}_0, \dots, \bar{x}_{n-1}) \in \Delta$. Let d be the number of variables appearing in Δ and let $c : (m^{<\omega})^d \to S(\Delta)$ be the function mapping each tuple $\bar{\eta} \in (m^{<\omega})^d$ to the type $\mathrm{tp}_\Delta(\bar{a}[\bar{\eta}])$.

Let $k < \omega$ be some number such that Φ_0 only contains variables $\bar{x}_\eta$ with $\eta \in m^{<k}$. We can use Theorem 2.10 to find an embedding $g : \mathfrak{T}_*(m^{<k}) \to \mathfrak{T}_*(m^{<\omega})$ such that $\mathrm{rng}\, g$ is $\approx_*$-homogeneous with respect to c. It follows that the family $(\bar{a}_{g(\eta)})_{\eta \in m^{<k}}$ satisfies Φ_0.

By the Compactness Theorem we conclude that Φ is satisfiable. Let $\bar{b} = (\bar{b}_\eta)_{\eta \in m^{<\omega}}$ be a family realising Φ. Then $\bar{b}$ is tree-indiscernible over U since it satisfies Ξ. Hence, it remains to show that $\bar{b}$ is inspired by $\bar{a}$.

For a contradiction, suppose otherwise. Then there exist a finite tuple $\bar{\eta} \subseteq m^{<\omega}$ and a finite set of formulae Δ over U such that

$$\mathrm{tp}_\Delta(\bar{b}[\bar{\eta}]) \neq \mathrm{tp}_\Delta(\bar{a}[\bar{\zeta}])\,, \quad \text{for all } \bar{\zeta} \approx_* \bar{\eta}\,.$$

W.l.o.g. we may assume that Δ is closed under negation. Set

$$\vartheta(\bar{x}) := \bigwedge \mathrm{tp}_\Delta(\bar{b}[\bar{\eta}])\,.$$

Then

$$\mathbb{M} \vDash \neg\vartheta(\bar{a}[\bar{\zeta}])\,, \quad \text{for all } \bar{\zeta} \approx_* \bar{\eta}\,.$$

Consequently, $\neg\vartheta(\bar{x}[\bar{\eta}]) \in \Psi_{\bar{\eta}}$. Since $\bar{b}$ satisfies $\Psi_{\bar{\eta}}$ it therefore follows that $\mathbb{M} \vDash \neg\vartheta(\bar{b}[\bar{\eta}])$. A contradiction. ∎

4. The independence and strict order properties

In this section we use indiscernible sequences to study concepts related to the order property. Recall that

$$[\![\varphi(\bar{a}, \bar{b}^i)]\!]_{i \in I} := \{\, i \in I \mid \mathbb{M} \vDash \varphi(\bar{a}, \bar{b}^i) \,\}\,.$$

Definition 4.1. Let T be a theory. A formula $\varphi(\bar{x}, \bar{y})$ has the *independence property* (with respect to T) if there exists a model $\mathfrak{M} \vDash T$ containing two sequences $(\bar{a}^w)_{w \in \wp(\omega)}$ and $(\bar{b}^n)_{n < \omega}$ such that

$$\mathfrak{M} \vDash \varphi(\bar{a}^w, \bar{b}^n) \quad \text{iff} \quad n \in w\,.$$

If some formula has the independence property with respect to T, we also say that T has the *independence property.*

Proposition 4.2. *Let T be a first-order theory and $\varphi(\bar{x}, \bar{y})$ a formula. The following statements are equivalent:*

(1) *φ has the independence property.*

(2) *For every finite number $m < \omega$, there exist sequences $(\bar{a}^w)_{w \in \wp[m]}$ and $(\bar{b}^n)_{n < m}$ such that*

$$\mathbb{M} \vDash \varphi(\bar{a}^w, \bar{b}^n) \quad \text{iff} \quad n \in w\,.$$

(3) *There exist a sequence $(\bar{a}^w)_{w \in \wp(\omega)}$ and an indiscernible sequence $(\bar{b}^n)_{n < \omega}$ such that*

$$\mathfrak{M} \vDash \varphi(\bar{a}^w, \bar{b}^n) \quad \text{iff} \quad n \in w\,.$$

(4) *There exist a tuple $\bar{a}$ and an indiscernible sequence $(\bar{b}^n)_{n < \omega}$ such that*

$$[\![\varphi(\bar{a}, \bar{b}^n)]\!]_{n < \omega} = \{\, 2n \mid n < \omega \,\}\,.$$

(5) *There exist a tuple $\bar{a}$ and an indiscernible sequence $(\bar{b}^i)_{i \in I}$ such that $[\![\varphi(\bar{a}, \bar{b}^i)]\!]_{i \in I}$ is not a finite union of segments.*

Proof. The implications (3) ⇒ (4) ⇒ (5) are trivial and (2) ⇒ (1) follows by compactness.

For (1) ⇒ (3), let $(\bar{a}^w)_{w \in \wp(\omega)}$ and $(\bar{b}^n)_{n<\omega}$ be sequences such that

$$\mathfrak{M} \vDash \varphi(\bar{a}^w, \bar{b}^n) \quad \text{iff} \quad n \in w\,.$$

By Proposition 3.6, there exists an indiscernible sequence $(\bar{d}^n)_{n<\omega}$ with the same average type as $(\bar{b}^n)_{n<\omega}$. By compactness, we can find a sequence $(\bar{c}^w)_{w \in \wp(\omega)}$ such that

$$\mathfrak{M} \vDash \varphi(\bar{c}^w, \bar{d}^n) \quad \text{iff} \quad n \in w\,.$$

It remains to prove (5) ⇒ (2). Fix $m < \omega$ and let $\bar{a}$ and $(\bar{b}^i)_{i \in I}$ be such that $[\![\varphi(\bar{a}, \bar{b}^i)]\!]_{i \in I}$ is not a finite union of segments. We can find a strictly increasing sequence $i_0 < \dots < i_{2m-1}$ of indices in I such that

$$\mathbb{M} \vDash \varphi(\bar{a}, \bar{b}^{i_k}) \quad \text{iff} \quad k \text{ is odd}\,.$$

Set $\bar{d}^k := \bar{b}^{i_k}$ and let

$$\chi_w(k) := \begin{cases} 0 & \text{if } k \notin w\,, \\ 1 & \text{if } k \in w\,, \end{cases}$$

be the characteristic function of w. Note that the sequence $(\bar{d}^k)_{k<2m}$ is also indiscernible. For each $w \subseteq [m]$, we can therefore find an automorphism π_w of $\mathbb{M}$ such that

$$\pi_w(\bar{d}^k) = \bar{d}^{2n+\chi_w(k)}\,, \quad \text{for } k < m\,.$$

Setting $\bar{c}^w := \pi_w^{-1}(\bar{a})$ it follows that

$$\begin{aligned} \mathbb{M} \vDash \varphi(\bar{c}^w, \bar{d}^k) \quad &\text{iff} \quad \mathbb{M} \vDash \varphi(\pi_w(\bar{c}^w), \pi_w(\bar{d}^k)) \\ &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}, \bar{d}^{2n+\chi_w(k)}) \\ &\text{iff} \quad \chi_w(k) = 1 \\ &\text{iff} \quad k \in w\,. \end{aligned}$$

□

We can generalise Condition (4) above as follows.

Corollary 4.3. *Let $\varphi(\bar{x}; \bar{y}_0, \dots, \bar{y}_{n-1})$ be a formula. If there exist a tuple $\bar{c}$ and an indiscernible sequence $(\bar{a}_i)_{i \in I}$ such that the order I has no last element,*

$$\begin{aligned} &\mathbb{M} \vDash \varphi(\bar{c}; \bar{a}[\bar{\imath}])\,, && \text{for arbitrarily large } \bar{\imath} \in [I]^n, \\ \text{and} \quad &\mathbb{M} \vDash \neg\varphi(\bar{c}; \bar{a}[\bar{\imath}])\,, && \text{for arbitrarily large } \bar{\imath} \in [I]^n, \end{aligned}$$

then φ has the independence property.

Proof. By assumption we can inductively choose tuples $\bar{k}_0 < \bar{k}_1 < \dots$ in $[I]^n$ such that

$$\mathbb{M} \vDash \varphi(\bar{c}; \bar{a}[\bar{k}_i]) \quad \text{iff} \quad i \text{ is even.}$$

Since the sequence $(\bar{a}[\bar{k}_i])_{i<\omega}$ is indiscernible, the claim follows by Proposition 4.2 (4). □

Lemma 4.4. *Let T be a first-order theory. If $\varphi(\bar{x}, \bar{y})$ has the independence property then so does $\varphi(\bar{y}, \bar{x})$.*

Proof. We apply the characterisation in Proposition 4.2 (2). Let $m < \omega$. Since $\varphi(\bar{x}, \bar{y})$ has the independence property there are tuples $\bar{a}^w$ and $\bar{b}^n$ for $w \subseteq \mathscr{P}(2^m)$ and $n < 2^m$ such that

$$\mathbb{M} \vDash \varphi(\bar{a}^w, \bar{b}^n) \quad \text{iff} \quad n \in w\,.$$

We identify each number $k < 2^m$ with the function $k : [m] \to [2]$ such that $k = \sum_{i<m} k(i)2^i$. For $i < m$ and $s \subseteq [m]$, we define

$$\bar{c}^s := \bar{b}^{n_s} \quad \text{and} \quad \bar{d}^i := \bar{a}^{w_i},$$

where

$$n_s := \textstyle\sum_{i \in s} 2^i \quad \text{and} \quad w_i := \{\, k < 2^m \mid k(i) = 1 \,\}\,.$$

It follows that

$$\begin{aligned} \mathbb{M} \vDash \varphi(\bar{d}^i, \bar{c}^s) \quad &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}^{w_i}, \bar{b}^{n_s}) \\ &\text{iff} \quad n_s \in w_i \\ &\text{iff} \quad i \in s\,. \end{aligned}$$

□

Lemma 4.5. *Let T be a first-order theory and $\varphi(\bar{x}, \bar{y})$ a formula with the independence property. There exist formulae $\psi(x, \bar{y})$ and $\vartheta(\bar{x}, y)$ with, respectively, a single variable x and a single variable y that have the independence property.*

Proof. We construction ψ using Proposition 4.2 (3). Let $\bar{a}$ and $(\bar{b}^n)_{n<\omega}$ be tuples such that $[\![\varphi(\bar{a}, \bar{b}^n)]\!]_{n<\omega} = \{ 2n \mid n < \omega \}$. Suppose that $\bar{a} = a_0\bar{a}'$. We define a new sequence $\bar{c}^n := \bar{b}^n\bar{a}'$ and the formula $\psi(x, \bar{y}\bar{z}) := \varphi(x\bar{z}, \bar{y})$. It follows that $[\![\psi(a, \bar{c}^n)]\!]_{n<\omega} = \{ 2n \mid n < \omega \}$. Hence, ψ has the independence property.

To find $\vartheta(\bar{x}, y)$ it is sufficient to note that, according to Lemma 4.4, the formula $\varphi(\bar{y}, \bar{x})$ also has the independence property. Hence, we can apply the first part of the lemma. □

The independence property is closely related to the order property which characterises unstable theories.

Lemma 4.6. *Every formula with the independence property has the order property.*

Proof. Suppose that φ is a formula with the independence property and let $(\bar{a}^w)_{w\subseteq\wp(\omega)}$ and $(\bar{b}^n)_{n<\omega}$ be sequences such that

$$\mathbb{M} \vDash \varphi(\bar{a}^w, \bar{b}^n) \quad \text{iff} \quad n \in w\,.$$

Setting $w_n := \omega \setminus [n]$ and $\bar{c}^n := \bar{a}^{w_n}$ it follows that

$$\mathbb{M} \vDash \varphi(\bar{c}^n, \bar{b}^k) \quad \text{iff} \quad n \leq k\,.$$

Hence, φ has the order property. □

Lemma 4.7. *No o-minimal theory has the independence property.*

Proof. Let T be a theory with the independence property. Then there exist a model $\mathfrak{M}$ of T, a formula $\varphi(x, \bar{y})$, parameters $\bar{c} \subseteq M$, and an indiscernible sequence $(a_n)_{n<\omega}$ such that

$$\mathfrak{M} \vDash \varphi(a_n, \bar{c}) \quad \text{iff} \quad n \equiv 0 \pmod 2 .$$

Since $(a_n)_n$ is indiscernible we either have $a_0 < a_1 < \dots$ or $a_0 > a_1 > \dots$. In both cases it follows that the set $\varphi(x, \bar{c})^{\mathfrak{M}}$ is not a finite union of intervals. Hence, T is not o-minimal. □

Lemma 4.8. *Let $\varphi(\bar{x}, \bar{y})$ be a formula without the independence property. Suppose that there exists a tuple $\bar{c}$ and a sequence $(\bar{a}^i)_{i\in I}$ such that the sets $[\![\varphi(\bar{c}, \bar{a}^i)]\!]_i$ and $[\![\neg\varphi(\bar{c}, \bar{a}^i)]\!]_i$ are both infinite. Then there exists a formula $\chi(\bar{y}, \bar{y}'; \bar{d})$ with parameters $\bar{d}$ such that*

$$\mathbb{M} \vDash \chi(\bar{a}^i, \bar{a}^k; \bar{d}) \quad \text{iff} \quad i \leq k .$$

Proof. Let J be an open dense linear order with $I \subseteq J$ such that J contains infinitely many elements above I and below I. By Lemma 3.9, we can extend $(\bar{a}^i)_{i\in I}$ to an indiscernible sequence $(\bar{a}^i)_{i\in J}$. Replacing φ by $\neg\varphi$ if necessary, we may assume that $[\![\varphi(\bar{c}, \bar{a}^i)]\!]_i$ contains a final segment of J. By Proposition 4.2 (2), there exists a number m such that, for all indices $\bar{s} \in [I]^m$,

$$\mathbb{M} \vDash \neg\exists\bar{x} \bigwedge_{i<m-1} [\varphi(\bar{x}, \bar{a}^{s_i}) \leftrightarrow \neg\varphi(\bar{x}, \bar{a}^{s_{i+1}})] .$$

Consequently, there exists a number $0 < n \leq m$, a set $w \subseteq [n]$, and indices $\bar{s} \in [I]^n$ such that there is no $\bar{c}'$ with

$$\begin{aligned} &\downarrow s_0 \cup \{ s_i \mid i \notin w \} \subseteq [\![\neg\varphi(\bar{c}', \bar{a}^i)]\!]_i \\ \text{and} \quad &\uparrow s_{n-1} \cup \{ s_i \mid i \in w \} \subseteq [\![\varphi(\bar{c}', \bar{a}^i)]\!]_i . \end{aligned}$$

We choose n and w such that $\langle n, w\rangle$ is minimal with respect to the lexicographic order (treating $w \subseteq [n]$ as a word in $[2]^n$). By minimality of n, it follows that $0 \in w$ and $n-1 \notin w$. Hence, there is some index $k < n$ with $[k] \subseteq w$ and $k \notin w$.

By compactness, there are finite sets $J_- \subseteq \downarrow s_0$ and $J_+ \subseteq \uparrow s_{n-1}$ such that there is no $\bar{c}'$ with

$$\begin{aligned} & J_- \cup \{ s_i \mid i \notin w \} \subseteq [\![\neg\varphi(\bar{c}', \bar{a}^i)]\!]_i \\ \text{and} \quad & J_+ \cup \{ s_i \mid i \in w \} \subseteq [\![\varphi(\bar{c}', \bar{a}^i)]\!]_i \,. \end{aligned}$$

By indiscernibility, we may assume that

$$J_- \cup \{ s_i \mid i < k \} < I < J_+ \cup \{ s_i \mid i \geq k \} \,.$$

Let $w_+ := w \setminus \{k-1\}$ and $w_- := [n] \setminus (w \cup \{k\})$. We define

$$\psi(\bar{x}) := \bigwedge_{i \in J_- \cup w_-} \neg\varphi(\bar{x}, \bar{a}^i) \wedge \bigwedge_{i \in J_+ \cup w_+} \varphi(\bar{x}, \bar{a}^i) \,.$$

Then

$$\mathbb{M} \vDash \neg\exists\bar{x}\big[\psi(\bar{x}) \wedge \varphi(\bar{x}, \bar{a}^{s_{k-1}}) \wedge \neg\varphi(\bar{x}, \bar{a}^{s_k})\big] \,.$$

Hence,

$$\mathbb{M} \vDash \forall\bar{x}\big[\psi(\bar{x}) \wedge \varphi(\bar{x}, \bar{a}^{s_{k-1}}) \rightarrow \varphi(\bar{x}, \bar{a}^{s_k})\big] \,.$$

Moreover, $(w \setminus \{k-1\}) \cup \{k\} <_{\text{lex}} w$ implies, by choice of w, that

$$\mathbb{M} \vDash \exists\bar{x}\big[\psi(\bar{x}) \wedge \neg\varphi(\bar{x}, \bar{a}^{s_{k-1}}) \wedge \varphi(\bar{x}, \bar{a}^{s_k})\big] \,.$$

Consequently, it follows by indiscernibility that, for all $i, l \in [s_{k-1}, s_k]$,

$$\mathbb{M} \vDash \forall\bar{x}[\psi(\bar{x}) \wedge \varphi(\bar{x}, \bar{a}^i) \rightarrow \varphi(\bar{x}, \bar{a}^l)] \quad \text{iff} \quad i \leq l \,.$$

In particular, this holds for all $i, l \in I$. ☐

Lemma 4.7 shows that there are unstable theories without the independence property. Such theories can be characterised as follows.

Definition 4.9. Let T be a theory. A formula $\varphi(\bar{x}, \bar{y})$ has the *strict order property* (with respect to T) if there exists a model $\mathfrak{M} \vDash T$ containing a sequence $(\bar{a}^n)_{n<\omega}$ such that

$$\mathfrak{M} \vDash \exists \bar{x}[\neg\varphi(\bar{x}, \bar{a}^i) \wedge \varphi(\bar{x}, \bar{a}^k)] \quad \text{iff} \quad i < k\,.$$

If some formula has the strict order property with respect to T then we also say that T has the *strict order property.*

Lemma 4.10. *A theory T has the strict order property if and only if there exists a formula $\varphi(\bar{x}, \bar{y})$ such that $\varphi^{\mathbb{M}}$ is a preorder with infinite chains.*

Proof. ($\Leftarrow$) Suppose that $\varphi(\bar{x}, \bar{y})$ defines a preorder with an infinite chain $(\bar{a}^i)_{i\in I}$. By compactness, there exists an infinite ascending $\varphi^{\mathbb{M}}$-chain $(\bar{b}^n)_{n<\omega}$. It follows that

$$\mathbb{M} \vDash \exists \bar{x}[\neg\varphi(\bar{x}, \bar{b}^i) \wedge \varphi(\bar{x}, \bar{b}^k)] \quad \text{iff} \quad i < k\,.$$

($\Rightarrow$) Suppose that there exists a formula $\psi(\bar{x}, \bar{y})$ with the strict order property and let $(\bar{a}^n)_{n<\omega}$ be a sequence with

$$\mathbb{M} \vDash \exists \bar{x}[\neg\psi(\bar{x}, \bar{a}^i) \wedge \psi(\bar{x}, \bar{a}^k)] \quad \text{iff} \quad i < k\,.$$

We set

$$\varphi(\bar{y}, \bar{y}') := \forall \bar{x}[\psi(\bar{x}, \bar{y}) \rightarrow \psi(\bar{x}, \bar{y}')]\,.$$

Clearly, $\varphi^{\mathbb{M}}$ is reflexive and transitive. Furthermore, we have

$$\mathbb{M} \vDash \varphi(\bar{a}^i, \bar{a}^k) \quad \text{iff} \quad i \geq k\,.$$

Hence, $(\bar{a}^n)_{n<\omega}$ is an infinite descending $\varphi^{\mathbb{M}}$-chain. ☐

Proposition 4.11. *A first-order theory T is unstable if, and only if, it has the independence property or the strict order property.*

Proof. ($\Leftarrow$) If there is a formula φ with the independence property then, according to Lemma 4.6, φ has also the order property and T is unstable.

Similarly, suppose that there exists a formula φ with the strict order property and let $(\bar{a}^n)_{n<\omega}$ be a sequence with

$$\mathbb{M} \vDash \exists \bar{x}[\neg\varphi(\bar{x}, \bar{a}^i) \wedge \varphi(\bar{x}, \bar{a}^k)] \quad \text{iff} \quad i < k\,.$$

Setting

$$\psi(\bar{x}, \bar{y}) :\!= \bar{x} = \bar{y} \vee \exists \bar{z}[\neg\varphi(\bar{z}, \bar{x}) \wedge \varphi(\bar{z}, \bar{y})]$$

it follows that

$$\mathbb{M} \vDash \psi(\bar{a}^i, \bar{a}^k) \quad \text{iff} \quad i \leq k\,.$$

Hence, ψ has the order property and T is unstable.

($\Rightarrow$) Let $\varphi(\bar{x}, \bar{y})$ be a formula with the order property and suppose that $(\bar{a}^n)_{n<\omega}$ and $(\bar{b}^n)_{n<\omega}$ are indiscernible sequences such that

$$\mathbb{M} \vDash \psi(\bar{a}^i, \bar{b}^k) \quad \text{iff} \quad i \leq k\,.$$

By compactness, there are indiscernible sequences $(\bar{a}^i)_{i\in\mathbb{Z}}$ and $(\bar{b}^i)_{i\in\mathbb{Z}}$ such that

$$\mathbb{M} \vDash \psi(\bar{a}^i, \bar{b}^k) \quad \text{iff} \quad i \leq k\,.$$

If φ has the independence property we are done. Hence, suppose otherwise. Since $[\![\psi(\bar{a}^\circ, \bar{b}^i)]\!]_i$ and $[\![\neg\psi(\bar{a}^\circ, \bar{b}^i)]\!]_i$ are both infinite, we can use Lemma 4.8 to construct a formula $\chi(\bar{y}, \bar{y}; \bar{d})$ such that

$$\mathbb{M} \vDash \chi(\bar{b}^i, \bar{b}^k; \bar{d}) \quad \text{iff} \quad i \leq k\,.$$

It follows that

$$\mathbb{M} \vDash \exists \bar{x}[\neg\chi(\bar{x}, \bar{b}^i; \bar{d}) \wedge \chi(\bar{x}, \bar{b}^k; \bar{d})] \quad \text{iff} \quad i < k\,.$$

Consequently, the sequence $(\bar{b}^i \bar{d})_{i<\omega}$ witnesses that $\chi(\bar{x}, \bar{y}; \bar{z})$ has the strict order property. □

Proposition 4.12. *Let $\varphi(\bar{x}, \bar{y})$ be a formula over a set U. The following statements are equivalent:*

(1) *$\varphi(\bar{x}, \bar{y})$ has the order property.*

(2) *There exist an indiscernible sequence $(\bar{a}^i)_{i \in I}$ over U and a tuple $\bar{c}$ such that both the set $[\![\varphi(\bar{a}^i, \bar{c})]\!]_{i \in I}$ and its complement are infinite.*

(3) *There exists an indiscernible sequence $(\bar{a}^i)_{i \in I}$ such that, for every number $m < \omega$, there exists a tuple $\bar{c}$ such that*

$$\big|[\![\varphi(\bar{a}^i; \bar{c})]\!]_{i \in I}\big| > m \quad \textit{and} \quad \big|[\![\neg\varphi(\bar{a}^i; \bar{c})]\!]_{i \in I}\big| > m\,.$$

Proof. (1) ⇒ (3) By Proposition 3.6 and compactness, it is sufficient to find, for every $m < \omega$, a tuple $\bar{c}$ and a sequence $(\bar{a}^i)_{i < \omega}$ such that

$$\big|[\![\varphi(\bar{a}^i, \bar{c})]\!]_{i \in I}\big| \geq m \quad \text{and} \quad \big|[\![\neg\varphi(\bar{a}^i, \bar{c})]\!]_{i \in I}\big| \geq m\,.$$

Since φ has the order property there are sequences $(\bar{c}^n)_{n < \omega}$ and $(\bar{d}^n)_{n < \omega}$ such that

$$\mathbb{M} \vDash \varphi(\bar{c}^i, \bar{d}^k) \quad \text{iff} \quad i \leq k\,.$$

Given $m < \omega$ we consider the tuple $\bar{c} := \bar{d}^m$ and the sequence $\bar{a}^i := \bar{c}^i$, $i < \omega$. Then

$$[\![\varphi(\bar{a}^i, \bar{c})]\!]_{i \in I} = [\![\varphi(\bar{c}^i, \bar{d}^m)]\!]_{i \in I} = \{m, m+1, \dots\}$$

and $$[\![\neg\varphi(\bar{a}^i, \bar{c})]\!]_{i \in I} = [\![\neg\varphi(\bar{c}^i, \bar{d}^m)]\!]_{i \in I} = \{0, \dots, m-1\}$$

contain both at least m elements.

(2) ⇒ (1) Let $\bar{c}$ and $(\bar{a}^i)_{i \in I}$ be given. According to Proposition 4.2, if neither

$$I_0 := [\![\neg\varphi(\bar{a}^i, \bar{c})]\!]_{i \in I} \quad \text{nor} \quad I_1 := [\![\varphi(\bar{a}^i, \bar{c})]\!]_{i \in I}$$

can be written as a finite union of segments then φ has the independence property. By Lemma 4.6, this implies that φ has the order property.

Hence, it remains to consider the case that both I_0 and I_1 are finite unions of segments. Since these sets are both infinite it follows that each contains at least one infinite segment. By taking a suitable subsequence of $(\bar{a}^i)_{i \in I}$ we may assume that both sets consist of a single infinite segment. Reversing the sequence $(\bar{a}^i)_{i \in I}$ if necessary, we may further assume that $I_0 < I_1$.

By compactness it is sufficient to find, for every $m < \omega$, sequences $(\bar{c}^i)_{i<m}$ and $(\bar{d}^i)_{i<m}$ such that

$$\mathbb{M} \vDash \varphi(\bar{c}^i, \bar{d}^k) \quad \text{iff} \quad i \leq k\,.$$

Given $m < \omega$ we pick indices $k_0 < \cdots < k_{m-1}$ in I_0 and $k_m < \cdots < k_{2m-1}$ in I_1. For $i < m$, let π_i be an automorphism with $\pi_i(\bar{a}^{k_j}) = \bar{a}^{k_{j-i}}$ and define

$$\bar{c}^i := \bar{a}^{k_{m-i}} \quad \text{and} \quad \bar{d}^i := \pi_i(\bar{c})\,.$$

For $i, l < m$, it then follows that

$$\begin{aligned} \mathbb{M} \vDash \varphi(\bar{c}^i, \bar{d}^l) \quad &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}^{k_{m-i}}, \pi_l(\bar{c})) \\ &\text{iff} \quad \mathbb{M} \vDash \varphi(\pi_l(\bar{a}^{k_{m-i+l}}), \pi_l(\bar{c})) \\ &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}^{k_{m-i+l}}, \bar{c}) \\ &\text{iff} \quad m - i + l \geq m \\ &\text{iff} \quad i \leq l\,. \end{aligned}$$

(3) ⇒ (2) By Corollary 3.10, we may assume that the order I is dense. Set

$$\Phi := \mathrm{Av}((\bar{a}^i)_i/U) \cup \{\, \varphi(\bar{x}^n; \bar{y}) \leftrightarrow \neg\varphi(\bar{x}^{n+1}; \bar{y}) \mid n < \omega \,\}\,.$$

If Φ is satisfiable, there exists an indiscernible sequence $(\bar{b}^n)_{n<\omega}$ over U and a tuple $\bar{c}$ such that

$$[\![\varphi(\bar{b}^n; \bar{c})]\!]_{n<\omega} = \{\, 2n \mid n < \omega \,\}$$

and $\quad [\![\neg\varphi(\bar{b}^n;\bar{c})]\!]_{n<\omega} = \{\, 2n+1 \mid n<\omega \,\}$.

In particular, both sets are infinite.

Hence, it remains to prove that Φ is satisfiable. Consider a finite subset $\Phi_0 \subseteq \Phi$. Let $n < \omega$ be the maximal number such that Φ_0 contains a formula of the form

$$\varphi(\bar{x}^n;\bar{y}) \leftrightarrow \neg\varphi(\bar{x}^{n+1};\bar{y})\,.$$

By (3), there exists a tuple $\bar{c}$ such that

$$\left|[\![\varphi(\bar{a}^i;\bar{c})]\!]_{i\in I}\right| > n \quad\text{and}\quad \left|[\![\neg\varphi(\bar{a}^i;\bar{c})]\!]_{i\in I}\right| > n\,.$$

If both sets are infinite, we are done. Hence, suppose that one of them is finite. Choose indices $k_0 < \dots < k_{n-1}$ in the finite set. As the other set is dense and cofinite, it contains indices $l_0 < \dots < l_{n-1}$ such that

$$k_0 < l_0 < k_1 < l_1 < \dots < k_{n-1} < l_{n-1}\,.$$

Let K be this set of indices. Then $(\bar{a}^i)_{i\in K}$ and $\bar{c}$ satisfy Φ_0. □

Corollary 4.13. *A first-order theory T is stable if, and only if, for every formula $\varphi(\bar{x})$ with parameters and all indiscernible sequences $(\bar{a}^i)_{i\in I}$ at least one of the sets $[\![\varphi(\bar{a}^i)]\!]_{i\in I}$ and $[\![\neg\varphi(\bar{a}^i)]\!]_{i\in I}$ is finite.*

Corollary 4.14. *Let T be a stable theory and $(\bar{a}^i)_{i\in I}$ an indiscernible sequence over U. For every set $C \subseteq \mathbb{M}$, the set*

$$\mathrm{Av}_1((\bar{a}^i)_i/C) := \{\, \varphi(\bar{x}) \mid \varphi \text{ a formula over } C \text{ such that } [\![\varphi(\bar{a}^i)]\!]_{i\in I} \text{ is cofinite} \,\}$$

forms a complete type over C.

Proof. By the preceding corollary, we have

$$\begin{aligned}
\varphi(\bar{x}) \in \mathrm{Av}_1((\bar{a}^i)_i/C) \quad &\text{iff} \quad [\![\varphi(\bar{a}^i)]\!]_{i\in I} \text{ is cofinite} \\
&\text{iff} \quad [\![\neg\varphi(\bar{a}^i)]\!]_{i\in I} \text{ is finite} \\
&\text{iff} \quad [\![\neg\varphi(\bar{a}^i)]\!]_{i\in I} \text{ is not cofinite} \\
&\text{iff} \quad \neg\varphi(\bar{x}) \notin \mathrm{Av}_1((\bar{a}^i)_i/C)\,.
\end{aligned}$$

Hence, it remains to prove that $\mathrm{Av}_1((\bar{a}_i)_i/C)$ is consistent with T. Let $\varphi_0, \dots, \varphi_n \in \mathrm{Av}_1((\bar{a}_i)_i/C)$. Then

$$[\![\varphi_0(\bar{a}_i)]\!]_{i \in I}, \dots, [\![\varphi_n(\bar{a}_i)]\!]_{i \in I} \text{ are cofinite.}$$

Hence, so is

$$[\![\varphi_0(\bar{a}_i) \wedge \dots \wedge \varphi_n(\bar{a}_i)]\!]_{i \in I} = [\![\varphi_0(\bar{a}_i)]\!]_{i \in I} \cap \dots \cap [\![\varphi_n(\bar{a}_i)]\!]_{i \in I} .$$

Fixing some index i in this set, it follows that

$$\mathbb{M} \vDash \varphi_0(\bar{a}_i) \wedge \dots \wedge \varphi_n(\bar{a}_i) .$$

Consequently, every finite subset of $\mathrm{Av}_1((\bar{a}_i)_i/C)$ is satisfiable. □

E6. Functors and embeddings

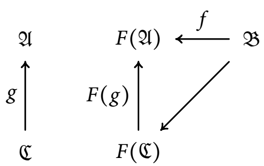
1. Local functors

In this section we consider functors preserving back-and-forth equivalence. Recall that $\mathrm{Sub}_\kappa(\mathfrak{M})$ denotes the class of all κ-generated substructures of $\mathfrak{M}$, and that a class $\mathcal{K}$ is κ-hereditary if $\mathfrak{M} \in \mathcal{K}$ implies $\mathrm{Sub}_\kappa(\mathfrak{M}) \subseteq \mathcal{K}$.

Definition 1.1. Let $\mathcal{K}$ be a class of Σ-structures. We denote the subcategory of $\mathfrak{Emb}(\Sigma)$ induced by $\mathcal{K}$ by $\mathfrak{Emb}(\mathcal{K})$.

Below we will show that functors preserving direct limits also preserve ∞-equivalence. We start by giving an alternative characterisation of such functors.

Definition 1.2. A functor $F : \mathfrak{Emb}(\mathcal{C}) \to \mathfrak{Emb}(\mathcal{K})$ is *κ-local* if, for every embedding $f : \mathfrak{B} \to F(\mathfrak{A})$ where $\mathfrak{B} \in \mathcal{K}$ is κ-generated and $\mathfrak{A} \in \mathcal{C}$, there exists an embedding $g : \mathfrak{C} \to \mathfrak{A}$ where $\mathfrak{C} \in \mathcal{C}$ is κ-generated such that the map f factors through $F(g)$.



Example. The following operations are $\aleph_0$-local functors.

(a) The function mapping a ring $\mathfrak{R}$ to the polynomial ring $\mathfrak{R}[x]$.

(b) The function mapping an integral domain $\mathfrak{R}$ to its quotient field.

(c) The function mapping a set X to the free group generated by X.

(d) The function mapping a structure $\mathfrak{M}$ to the structure $\mathrm{HF}(\mathfrak{M})$ consisting of all hereditary finite sets with elements from $\mathfrak{M}$.

Lemma 1.3. *If* $F : \mathfrak{Emb}(\mathcal{C}) \to \mathfrak{Emb}(\mathcal{D})$ *and* $G : \mathfrak{Emb}(\mathcal{D}) \to \mathfrak{Emb}(\mathcal{K})$ *are* κ*-local then so is* $G \circ F$.

Exercise 1.1. Prove the preceding lemma.

As a further, more involved example we show that quantifier-free interpretations are $\aleph_0$-local functors. While every interpretation is local in an intuitive sense we need the restriction to quantifier-free formulae to prove that the interpretation is a functor.

Lemma 1.4. *Every* $\mathrm{QF}_{\infty\aleph_0}$*-interpretation* $\mathcal{I} : \mathfrak{Emb}(\mathcal{K}) \to \mathfrak{Emb}(\Sigma)$ *is an* $\aleph_0$*-local functor.*

Proof. First, we show that quantifier-free interpretations are functors. Suppose that

$$\mathcal{I} = \langle \alpha, (\delta_s)_{s\in S}, (\varepsilon_s)_{s\in S}, (\varphi_\xi)_{\xi\in\Sigma} \rangle$$

is quantifier-free and let $h : \mathfrak{A} \to \mathfrak{B}$ be an embedding. For $\bar{a} \in \delta_s^{\mathfrak{A}}$, we denote by $[\bar{a}]_s$ the element encoded by $\bar{a}$. We define $\mathcal{I}(h)$ by

$$\mathcal{I}(h)[\bar{a}]_s := [h(\bar{a})]_s \,.$$

Since embeddings preserve quantifier-free formulae it follows that this mapping is a well-defined embedding $\mathcal{I}(h) : \mathcal{I}(\mathfrak{A}) \to \mathcal{I}(\mathfrak{B})$. Obviously, we have $\mathcal{I}(f \circ g) = \mathcal{I}(f) \circ \mathcal{I}(g)$. Consequently, $\mathcal{I}$ is a functor.

To show that it is $\aleph_0$-local let $X \subseteq \mathcal{I}(\mathfrak{A})$ be finite. For each equivalence class $[\bar{a}]_s \in X$, fix a representative $\bar{a}$ and let A_0 be the set of these representatives. Then A_0 is finite and we have $X \subseteq \mathcal{I}(\langle\!\langle A_0 \rangle\!\rangle_{\mathfrak{A}})$. Note that $\mathcal{I}(\langle\!\langle A_0 \rangle\!\rangle_{\mathfrak{A}})$ is defined since $\mathcal{I}$ is quantifier-free. □

Local functors can be characterised in purely category-theoretical terms as those functors that preserve direct limits.

Theorem 1.5. *Let $F : \mathfrak{Emb}(\mathcal{C}) \to \mathfrak{Emb}(\mathcal{K})$ be a functor where the classes $\mathcal{C}$ and $\mathcal{K}$ are κ-hereditary. The functor F is κ-local if and only if it preserves κ-filtered colimits.*

Proof. ($\Leftarrow$) Let $f : \mathfrak{B} \to F(\mathfrak{A})$ be an embedding where $\mathfrak{B} \in \mathcal{K}$ is κ-generated. According to Lemma ?? we can write $\mathfrak{A} = \varinjlim D$ where $D : \mathcal{I} \to \mathrm{Sub}_\kappa(\mathfrak{A})$ is the canonical κ-filtered diagram of all κ-generated substructures. The corresponding cocone μ from D to $\mathfrak{A}$ consists of all inclusion maps $\mu_{\mathfrak{i}} : D(\mathfrak{i}) \to \mathfrak{A}$. Since F preserves κ-direct limits we have $F(\mathfrak{A}) = \varinjlim(F \circ D)$ and the corresponding cone is $F[\mu]$.

To find the desired embedding $g : \mathfrak{C} \to \mathfrak{A}$ we fix a set $X \subseteq B$ of size $|X| < \kappa$ generating $\mathfrak{B}$. For each $x \in X$, we choose an index $\mathfrak{i}_x \in \mathcal{I}$ such that $f(x) \in \operatorname{rng} F(\mu_{\mathfrak{i}_x})$. Since I is κ-filtered there is some index $\mathfrak{k} \in I$ and morphisms $h_x : \mathfrak{i}_x \to \mathfrak{k}$, for all x. Hence, we have

$$f[X] \subseteq \operatorname{rng} F(\mu_{\mathfrak{k}}) ,$$

which, by Lemma B1.2.8, implies that

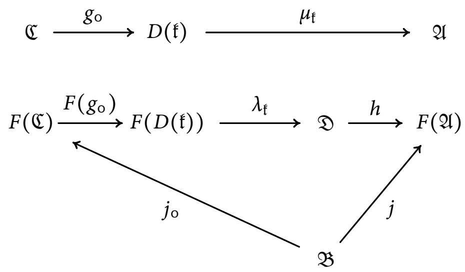
$$\begin{aligned} \operatorname{rng} f = f[\langle\!\langle X \rangle\!\rangle_{\mathfrak{B}}] &= \langle\!\langle f[X] \rangle\!\rangle_{F(\mathfrak{A})} \\ &\subseteq \langle\!\langle \operatorname{rng} F(\mu_{\mathfrak{k}}) \rangle\!\rangle_{F(\mathfrak{A})} = \operatorname{rng} F(\mu_{\mathfrak{k}}) . \end{aligned}$$

Since f and $F(\mu_{\mathfrak{k}})$ are injective and $\operatorname{rng} f \subseteq \operatorname{rng} F(\mu_{\mathfrak{k}})$ we can define a function $g : B \to F(D(\mathfrak{k}))$ by $g := F(\mu_{\mathfrak{k}})^{-1} \circ f$. Since f and $F(\mu_{\mathfrak{k}})$ preserve all quantifier-free formulae so does g. Hence, g is an embedding. Furthermore, we have $F(\mu_{\mathfrak{k}}) \circ g = f$.

($\Rightarrow$) Let $D : \mathcal{I} \to \mathfrak{Emb}(\mathcal{C})$ be a κ-filtered diagram with $\mathfrak{A} := \varinjlim D$, and suppose that μ is a limiting cocone from D to $\mathfrak{A}$. We claim that $\varinjlim(F \circ D) = F(\mathfrak{A})$. Let $\mathfrak{D} := \varinjlim(F \circ D)$ and let λ be a limiting cocone from $F \circ D$ to $\mathfrak{D}$. Since $F[\mu]$ is a cocone from $F \circ D$ to $F(\mathfrak{A})$ it follows that there exists an embedding $h : \mathfrak{D} \to F(\mathfrak{A})$ with $h * \lambda = F[\mu]$.

We only have to show that h is surjective. Fix $c \in F(\mathfrak{A})$. There exists some substructure $\mathfrak{B} \in \mathrm{Sub}_\kappa(F(\mathfrak{A}))$ with $c \in B$. Let $j : \mathfrak{B} \to F(\mathfrak{A})$ be the inclusion map. Since F is κ-local we can find a κ-generated structure $\mathfrak{C} \in \mathcal{C}$ and an embedding $g : \mathfrak{C} \to \mathfrak{A}$ such that $j = F(g) \circ j_0$, for some

$j_\circ : \mathfrak{B} \to F(\mathfrak{C})$. In the same way as above we can show that there is some index $\mathfrak{k} \in \mathcal{I}$ and an embedding $g_\circ : \mathfrak{C} \to D(\mathfrak{k})$ with $g = \mu_\mathfrak{k} \circ g_\circ$.



Since $a = j(a) = (h \circ \lambda_\mathfrak{k} \circ F(g_\circ) \circ j_\circ)(a)$ it follows that $a \in \operatorname{rng} h$. □

Let us show that local functors preserve back-and-forth equivalences.

Definition 1.6. Suppose that $F : \mathfrak{Emb}(\mathcal{C}) \to \mathfrak{Emb}(\mathcal{K})$ is a functor where the classes $\mathcal{C}$ and $\mathcal{K}$ are κ-hereditary. Let $p = \bar{a} \mapsto \bar{b} \in \operatorname{pIso}(\mathfrak{A}, \mathfrak{B})$ be a partial isomorphism between $\mathfrak{A}, \mathfrak{B} \in \mathcal{C}$ and let $\pi : \mathfrak{A}_\circ \to \mathfrak{B}_\circ$ be the unique isomorphism extending p, where $\mathfrak{A}_\circ := \langle\!\langle \bar{a} \rangle\!\rangle_\mathfrak{A}$ and $\mathfrak{B}_\circ := \langle\!\langle \bar{b} \rangle\!\rangle_\mathfrak{B}$ are the structures induced by, respectively, the domain and range of p. Let $i : \mathfrak{A}_\circ \to \mathfrak{A}$ and $j : \mathfrak{B}_\circ \to \mathfrak{B}$ be the corresponding inclusion maps and suppose that $F(\pi) = \bar{a}' \mapsto \bar{b}'$. We define

$$p^F := F(i)(\bar{a}') \mapsto F(j)(\bar{b}') .$$

Proposition 1.7. *Let $F : \mathfrak{Emb}(\mathcal{C}) \to \mathfrak{Emb}(\mathcal{K})$ be an $\aleph_\circ$-local functor where the classes $\mathcal{C}$ and $\mathcal{K}$ are $\aleph_\circ$-hereditary.*

$$p \in I^{\aleph_\circ}_{\omega\alpha}(\mathfrak{A}, \mathfrak{B}) \quad \textit{implies} \quad p^F \in I_\alpha(F(\mathfrak{A}), F(\mathfrak{B})) .$$

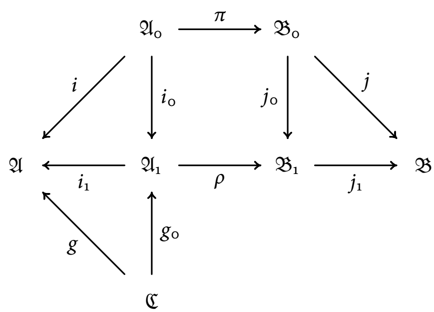
Proof. The claim follows by induction on α. Let $p := \bar{a} \mapsto \bar{b} \in I^{\aleph_\circ}_{\omega\alpha}(\mathfrak{A}, \mathfrak{B})$, set $\mathfrak{A}_\circ := \langle\!\langle \bar{a} \rangle\!\rangle_\mathfrak{A}$ and $\mathfrak{B}_\circ := \langle\!\langle \bar{b} \rangle\!\rangle_\mathfrak{B}$, and let $\pi : \mathfrak{A}_\circ \to \mathfrak{B}_\circ$ be the isomorphism extending p. Let $i : \mathfrak{A}_\circ \to \mathfrak{A}$ and $j : \mathfrak{B}_\circ \to \mathfrak{B}$ be the corresponding inclusion maps and suppose that $F(\pi) = \bar{a}' \mapsto \bar{b}'$.

For $\alpha = 0$, we have to check that $F(p)$ is a partial isomorphism. Since $F(i)$, $F(j)$, and $F(\pi)$ are embeddings it follows, for every quantifier-free formula $\varphi(\bar{x})$, that

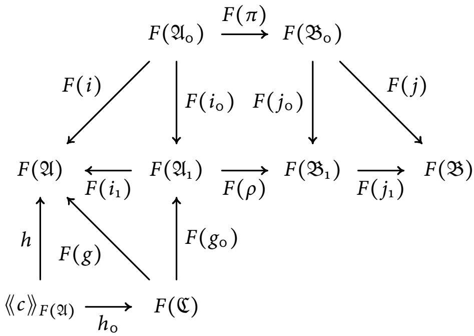
$$\begin{aligned} F(\mathfrak{A}) \vDash \varphi(F(i)(\bar{a}')) \quad &\text{iff} \quad F(\mathfrak{A}_0) \vDash \varphi(\bar{a}') \\ &\text{iff} \quad F(\mathfrak{B}_0) \vDash \varphi(\bar{b}') \\ &\text{iff} \quad F(\mathfrak{B}) \vDash \varphi(F(j)(\bar{b}'))\,. \end{aligned}$$

If α is a limit ordinal then the claim follows immediately by inductive hypothesis. Hence, suppose that $\alpha = \beta + 1$. By symmetry, we only need to check the forth property. Fix $c \in F(\mathfrak{A})$. Since F is $\aleph_0$-local there exist a finitely generated structure $\mathfrak{C}$ and an embedding $g : \mathfrak{C} \to \mathfrak{A}$ such that the inclusion $h : \langle\!\langle c \rangle\!\rangle_{F(\mathfrak{A})} \to F(\mathfrak{A})$ factors through $F(g)$, i.e., $h = F(g) \circ h_0$. Choose a finite tuple $\bar{e}_0$ of generators of $\mathfrak{C}$ and set $\bar{e} := g(\bar{e}_0)$ and $\mathfrak{A}_1 := \langle\!\langle \bar{a}\bar{e} \rangle\!\rangle_{\mathfrak{A}}$. Since $p = \bar{a} \mapsto \bar{b} \in \mathrm{pIso}^{\aleph_0}_{\omega(\beta+1)}(\mathfrak{A}, \mathfrak{B})$ we can find some $\bar{f} \subseteq B$ with $q := \bar{a}\bar{e} \mapsto \bar{b}\bar{f} \in \mathrm{pIso}^{\aleph_0}_{\omega\beta}(\mathfrak{A}, \mathfrak{B})$. Set $\mathfrak{B}_1 := \langle\!\langle \bar{b}\bar{f} \rangle\!\rangle_{\mathfrak{B}}$ and let $\rho : \mathfrak{A}_1 \to \mathfrak{B}_1$ be the unique isomorphism extending q. We claim that q^F is an extension of p^F with $c \in \mathrm{dom}\, q^F$.

Let i_0, i_1, j_0, j_1, g_0 be the inclusion maps as depicted in the following diagram



Applying F to this diagram we obtain



First, let us show that $c \in \operatorname{dom} q^F$. We have

$$c = h(c) = (F(i_1) \circ F(g_0) \circ h_0)(c)$$

which implies that $c \in \operatorname{rng} F(i_1) = \operatorname{dom} q^F$.

It remains to prove that $p^F \subseteq q^F$. Let $x \in \operatorname{dom} p^F$. Then $x = F(i)(a'_l)$, for some l. Setting $w := F(i_0)(a'_l)$ we have

$$F(i_1)(w) = (F(i_1) \circ F(i_0))(a'_l) = F(i)(a'_l) = x\,.$$

It follows that

$$\begin{aligned} q^F(x) &= (F(j_1) \circ F(\rho))(w) \\ &= (F(j_1) \circ F(\rho) \circ F(i_0))(a'_l) \\ &= (F(j_1) \circ F(j_0) \circ F(\pi))(a'_l) \\ &= (F(j) \circ F(\pi))(a'_l) = p^F(x)\,. \end{aligned}$$

□

Corollary 1.8. *Let $F : \mathfrak{Emb}(\mathcal{C}) \to \mathfrak{Emb}(\mathcal{K})$ be an $\aleph_0$-local functor where the classes $\mathcal{C}$ and $\mathcal{K}$ are $\aleph_0$-hereditary. For all $\mathfrak{A}, \mathfrak{B}$, we have*

$$\mathfrak{A} \cong_{\omega\alpha} \mathfrak{B} \quad \textit{implies} \quad F(\mathfrak{A}) \cong_{\alpha} F(\mathfrak{B})\,.$$

In particular,

$$\mathfrak{A} \cong_{\infty} \mathfrak{B} \quad \textit{implies} \quad F(\mathfrak{A}) \cong_{\infty} F(\mathfrak{B})\,.$$

We conclude this section by showing that local functors are compatible with universal theories.

Definition 1.9. Let $F : \mathfrak{Emb}(\mathcal{C}) \to \mathfrak{Emb}(\mathcal{K})$ be a functor and L a logic. The *L-theory* of F is the set

$$\mathrm{Th}_L(F) := \{ \varphi \in L \mid F(\mathfrak{A}) \vDash \varphi \text{ for all } \mathfrak{A} \in \mathcal{C} \}.$$

Lemma 1.10. *Let $F : \mathfrak{Emb}(\mathcal{C}) \to \mathfrak{Emb}(\mathcal{K})$ be an $\aleph_0$-local functor where the classes $\mathcal{C}$ and $\mathcal{K}$ are $\aleph_0$-hereditary. If $\mathfrak{U} \in \mathcal{C}$ is $\aleph_0$-universal then*

$$\mathrm{Th}_{\forall_{\infty\aleph_0}}(F(\mathfrak{U})) = \mathrm{Th}_{\forall_{\infty\aleph_0}}(F).$$

Proof. ($\supseteq$) follows immediately from the definitions.

($\subseteq$) We prove by induction on $\psi(\bar{x}) \in \forall_{\infty\aleph_0}$ that

$$F(\mathfrak{U}) \vDash \psi(\bar{c}), \quad \text{for all } \bar{c} \subseteq F(\mathfrak{U}),$$

implies that

$$F(\mathfrak{A}) \vDash \psi(\bar{a}), \quad \text{for all } \mathfrak{A} \in \mathcal{C} \text{ and every } \bar{a} \subseteq F(\mathfrak{A}).$$

First, suppose that ψ is quantifier-free. Let $\mathfrak{A} \in \mathcal{C}$ and $\bar{a} \subseteq F(\mathfrak{A})$. We have to show that $F(\mathfrak{A}) \vDash \psi(\bar{a})$. Since F is $\aleph_0$-local we can find a finitely generated substructure $\mathfrak{A}_0 \subseteq \mathfrak{A}$ with $\bar{a} \subseteq F(\mathfrak{A}_0)$. Since $\mathfrak{U}$ is $\aleph_0$-universal there exists an embedding $f : \mathfrak{A}_0 \to \mathfrak{U}$. We set $\bar{b} := F(f)(\bar{a})$. By assumption $F(\mathfrak{U}) \vDash \psi(\bar{b})$. Since ψ is quantifier-free and $F(f)$ is an embedding it follows that $F(\mathfrak{A}_0) \vDash \psi(\bar{a})$. Hence, $F(\mathfrak{A}) \vDash \psi(\bar{a})$.

For the inductive step, we have to distinguish three cases. Either

$$\psi(\bar{x}) = \bigwedge \Psi, \quad \text{or} \quad \psi(\bar{x}) = \bigvee \Psi, \quad \text{or} \quad \psi(\bar{x}) = \forall y \vartheta(\bar{x}, y).$$

In each of these cases the claim follows directly from the inductive hypothesis. □

2. Word constructions

Local functors can be characterised in terms of a certain family of comorphisms called *word constructions.* Instead of defining these operations as a single, complex construction we will introduce several simple operations which, when combined with first-order interpretations, yield the required expressive power.

We start with the main ingredient in a word construction, the so-called *term-algebra* operation.

Definition 2.1. Let Γ be a functional S-sorted signature and Σ a relational one that is S_0-sorted for some $S_0 \subseteq S$. The *Γ-term algebra* $\mathcal{T}[\Gamma, \mathfrak{A}]$ over a Σ-structure $\mathfrak{A}$ is the $T[\Gamma, S_0]$-sorted structure whose universe $T[\Gamma, A]$ consists of all Γ-terms over A. Every element $t(\bar{a}) \in T[\Gamma, A]$ has sort $t(\bar{s})$, where $\bar{s}$ are the sorts of $\bar{a}$. For each relation symbol $R \in \Sigma$, we have the relation

$$R^{\mathcal{T}[\Gamma,\mathfrak{A}]} = R^{\mathfrak{A}},$$

and, for each n-ary function symbol $f \in \Gamma$, we have an n-ary function defined by

$$f^{\mathcal{T}[\Gamma,\mathfrak{A}]}(t_0, \dots, t_{n-1}) := f t_0 \dots t_{n-1} .$$

Example. Let us give two simple examples showing the versatility of the term algebra operation in conjunction with a first-order interpretation.

(a) First, we interpret the product $\mathfrak{A} \times \mathfrak{A}$ in the structure $\mathcal{T}[\{f\}, \mathfrak{A}]$ where f is a binary function symbol. When we encode a pair $\langle a, b\rangle \in A \times A$ by the term $f(a, b)$, we can define the universe by the formula

$$\delta(x) := \text{“}x = f(a, b) \text{ for some } a, b \in A.\text{”}$$

Then we define each relation R by

$$\varphi_R(\bar{x}) := \text{“}x_i = f(a_i, b_i) \text{ for some } a_i, b_i \in A \text{ such that } \bar{a}, \bar{b} \in R.\text{”}$$

(b) Similarly, we can interpret the disjoint union $\mathfrak{A} \cup \mathfrak{A}$ in the structure $\mathcal{T}[\{f\}, \mathfrak{A}]$ where f is a unary function symbol. The universe is the set

$$A \cup \{ f(a) \mid a \in A \}$$

which is obviously definable in $\mathcal{T}[\{f\}, \mathfrak{A}]$. We can define the relations R by

$$\varphi_R(\bar{x}) := \text{“Either } \bar{x} = \bar{a} \text{ or } \bar{x} = f(\bar{a}), \text{ for some } \bar{a} \in R.\text{”}$$

Lemma 2.2. *Let Σ a relational signature and Γ a functional one. The Γ-term-algebra operation*

$$\mathcal{T}[\Gamma, -] : \mathfrak{Emb}(\Sigma) \to \mathfrak{Emb}(\Sigma \cup \Gamma)$$

is an $\aleph_0$-local functor.

Proof. First, let us show that it is a functor. Let $h : \mathfrak{A} \to \mathfrak{B}$ be an embedding of Σ-structures. We obtain an embedding

$$\mathcal{T}[\Gamma, h] : \mathcal{T}[\Gamma, \mathfrak{A}] \to \mathcal{T}[\Gamma, \mathfrak{B}]$$

by setting

$$\mathcal{T}[\Gamma, h](t(\bar{a})) := t(h(\bar{a}))\,.$$

To prove that $\mathcal{T}[\Gamma, -]$ is $\aleph_0$-local suppose that $X \subseteq T[\Gamma, A]$ is finite. Then we have $X \subseteq T[\Gamma, A_0] = \langle\!\langle A_0 \rangle\!\rangle_{\mathcal{T}[\Gamma, \mathfrak{A}]}$ where the set

$$A_0 := \bigcup \{ \bar{a} \mid t(\bar{a}) \in X \}$$

is finite. ☐

It follows from the results of the previous section that $\mathcal{T}[\Gamma, -]$ preserves ∞-equivalence. The next lemma gives a more precise statement.

Lemma 2.3. *Suppose that Σ is a relational signature, Γ a functional one, and κ an infinite cardinal. For each* $\mathrm{FO}_{\kappa\aleph_0}$*-formula $\varphi(x_0, \dots, x_{n-1})$ and all terms $t_i(\bar{x}^i) \in T^{<\omega}[\Gamma]$, for $i < n$, we can construct an* $\mathrm{FO}_{\kappa\aleph_0}$*-formula $\varphi_{t_0 \dots t_{n-1}}(\bar{x}^0, \dots, \bar{x}^{n-1})$ such that*

$$\mathcal{T}[\Gamma, \mathfrak{A}] \vDash \varphi(t_0(\bar{a}_0), \dots, t_{n-1}(\bar{a}_{n-1}))$$

iff $$\mathfrak{A} \vDash \varphi_{t_0 \dots t_{n-1}}(\bar{a}_0, \dots, \bar{a}_{n-1}) .$$

Proof. W.l.o.g. we may assume that φ is term reduced. We construct $\varphi_{\bar{t}}$ inductively. First, suppose that φ is an atomic formula. If $\varphi = R\bar{x}$ with $R \in \Sigma$ then we can set

$$(R\bar{x})_{\bar{t}} := \begin{cases} Rx^0 \dots x^{n-1} & \text{if } t_i = x \text{ for all } i \,, \\ \text{false} & \text{otherwise} . \end{cases}$$

For $\varphi = x = y$ we set

$$(x = y)_{st} := \begin{cases} \bigwedge_i x_i = y_i & \text{if } s = t \,, \\ \text{false} & \text{otherwise} . \end{cases}$$

Finally, if $\varphi = f\bar{x} = y$ then we define

$$(f\bar{x} = y)_{\bar{s}t} := \begin{cases} \bigwedge_{i,j} x^i_j = y^i_j & \text{if } f\bar{s} = t \,, \\ \text{false} & \text{otherwise} , \end{cases}$$

where $s_i = s_i(\bar{x}^i)$ and $t = t(\bar{y}^0, \dots, \bar{y}^{n-1})$. Boolean operations are unchanged:

$$(\neg\varphi)_{\bar{t}} := \neg\varphi_{\bar{t}} \quad \text{and} \quad (\bigwedge \Phi)_{\bar{t}} := \bigwedge \{ \varphi_{\bar{t}} \mid \varphi \in \Phi \} .$$

For a quantifier over a variable y of sort $s \in T[\Gamma, S_0]$, we have

$$(\exists y \varphi(\bar{x}, y))_{\bar{t}} := \exists \bar{y} \varphi_{\bar{t}s}(\bar{x}^0, \dots, \bar{x}^{n-1}, \bar{y}) .$$ ☐

The term-algebra operation creates structures with many sorts. To reduce the number of sorts we employ a second operation that merges several sorts into a single one. Recall that with every morphism $\langle S, \Sigma\rangle \to \langle T, \Gamma\rangle$ of $\mathfrak{Sig}$ we have associated a reduct mapping $\mathfrak{Str}[\Gamma] \to \mathfrak{Str}[\Sigma]$. For relational signatures we can also define a mapping $\mathfrak{Str}[\Sigma] \to \mathfrak{Str}[\Gamma]$ in the other direction.

Definition 2.4. Let $\alpha = \langle \chi, \mu\rangle : \langle S, \Sigma\rangle \to \langle T, \Gamma\rangle$ be a morphism of $\mathfrak{Sig}$ where the signatures Σ and Γ are relational. The *inverse α-reduct* of a Σ-structure $\mathfrak{A}$ is the Γ-structure $\mathfrak{A}^\alpha$ where the domain of sort $t \in T$ is

$$A^\alpha_t := \dot{\bigcup} \{ A_s \mid s \in \chi^{-1}(t) \},$$

and, for each relation symbol $R \in \Gamma$, we have

$$R^{\mathfrak{A}^\alpha} := \dot{\bigcup} \{ Q^{\mathfrak{A}} \mid Q \in \mu^{-1}(R) \}.$$

Remark. We have defined inverse reducts only for relational signatures in order to avoid the complications arising from the fact that we require functions to be total. For instance, if $\mathfrak{V} = \langle V, K, +, \cdot\rangle$ is a $\{v, s\}$-sorted vector space and α maps both sorts to the same value, then we get problems defining $\mathfrak{V}^\alpha$ since the operation $+$ is a function $V \times V \to V$ and not a function $(V \cup K) \times (V \cup K) \to V \cup K$.

Lemma 2.5. *Let α be a morphism of $\mathfrak{Sig}$. The operation $\mathfrak{A} \mapsto \mathfrak{A}^\alpha$ is an $\aleph_0$-local functor.*

Proof. Clearly the operation is $\aleph_0$-local: for every finite subset $X \subseteq \mathfrak{A}^\alpha$ we have $X \subseteq (\langle\!\langle X \rangle\!\rangle_{\mathfrak{A}})^\alpha$. It remains to show that it is a functor. Let $h : \mathfrak{A} \to \mathfrak{B}$ be an embedding. We define

$$h^\alpha : \mathfrak{A}^\alpha \to \mathfrak{B}^\alpha \quad \text{by setting} \quad h^\alpha(a) := h(a).$$

To show that this function is an embedding suppose that $\bar a \in R^{\mathfrak{A}^\alpha}$. Then there is some relation $Q \in \alpha^{-1}(R)$ with $\bar a \in Q^{\mathfrak{A}}$. Hence, $h(\bar a) \in Q^{\mathfrak{B}} \subseteq R^{\mathfrak{B}^\alpha}$. □

It follows that inverse reducts preserve $\mathrm{FO}_{\infty\aleph_0}$-equivalence. The next lemma states that they also preserve $\mathrm{FO}_{\kappa\aleph_0}$-equivalence for sufficiently large cardinals κ.

Lemma 2.6. *Let $\alpha = \langle \chi, \mu \rangle : \langle S, \Sigma \rangle \to \langle T, \Gamma \rangle$ be a morphism of $\mathfrak{Sig}$ where the signatures Σ and Γ are relational, and let κ be an infinite cardinal such that*

$$|\chi^{-1}(t)| < \kappa \quad \text{and} \quad |\mu^{-1}(R)| < \kappa\,, \quad \text{for all} \quad t \in T \text{ and } R \in \Gamma\,.$$

For every formula $\varphi(\bar{x}) \in \mathrm{FO}_{\kappa\aleph_0}[\Sigma]$ where x_i is of sort t_i and for all sorts $s_i \in \chi^{-1}(t_i)$, there exists a formula $\varphi^\alpha_{\bar{s}}(\bar{x}) \in \mathrm{FO}_{\kappa\aleph_0}[\Gamma]$ such that

$$\mathfrak{A}^\alpha \vDash \varphi(\bar{a}) \quad \text{iff} \quad \mathfrak{A} \vDash \varphi^\alpha_{\bar{s}}(\bar{a})\,,$$

for every Σ-structure $\mathfrak{A}$ and all $a_i \in A_{s_i}$.

Proof. We construct $\varphi^\alpha_{\bar{s}}$ by induction on φ. For atomic formulae we set

$$(x_0 = x_1)^\alpha_{\bar{s}} := x_0 = x_1 \quad \text{and} \quad (R\bar{x})^\alpha_{\bar{s}} := \bigvee \{\, Q \in \mu^{-1}(R) \mid Q\bar{x} \,\}$$

(where we consider x_i now to be of sort s_i). Boolean operations remain unchanged:

$$(\neg\varphi)^\alpha_{\bar{s}} := \neg\varphi^\alpha_{\bar{s}} \quad \text{and} \quad (\bigwedge \Phi)^\alpha_{\bar{s}} := \bigwedge \{\, \varphi^\alpha_{\bar{s}} \mid \varphi \in \Phi \,\}\,.$$

A quantifier with a variable y of sort $t \in T$ is replaced by a disjunction over all sorts $r \in \chi^{-1}(t)$

$$(\exists y\varphi)^\alpha_{\bar{s}} := \bigvee \{\, \exists y \varphi^\alpha_{\bar{s}r} \mid r \in \chi^{-1}(t) \,\}\,. \qquad \square$$

We obtain an alternative characterisation of $\aleph_0$-local functors by combining these two operations with quantifier-free interpretations.

Definition 2.7. (a) Let Σ be a signature and let Σ_{rel} be the signature obtained from Σ by replacing every function symbol f of type $\bar{s} \to t$ by a relation symbol R_f of type $\bar{s}t$. The *relational variant* of a Σ-structure $\mathfrak{M}$ is the Σ_{rel}-structure $\mathcal{R}(\mathfrak{M})$ obtained from $\mathfrak{M}$ by replacing every function f by its graph.

(b) A *κ-word construction* is an operation of the form

$$F = \mathcal{I} \circ \mathcal{S} \circ \mathcal{R} \circ \mathcal{T} \circ \mathcal{R},$$

where $\mathcal{I}$ is a $\mathrm{QF}_{\kappa\aleph_0}$-interpretation, $\mathcal{R}$ is the operation defined in (a), $\mathcal{S}$ is an inverse reduct, and $\mathcal{T}$ is a Γ-term-algebra operation where $|\Gamma| < \kappa$.

Remark. Note that $\mathcal{R}$ is a quantifier-free first-order interpretation.

Theorem 2.8. *Let $\mathcal{C}$ be an $\aleph_0$-hereditary class of Σ-structures and $\mathcal{K}$ a class of Γ-structures. Suppose that κ is a cardinal such that*

$$\kappa > 2^{|\Sigma| \oplus \aleph_0} \quad \text{and} \quad \kappa > |F(\mathfrak{C})|, \quad \text{for all finitely generated } \mathfrak{C} \in \mathcal{C}.$$

A mapping $F : \mathfrak{Emb}(\mathcal{C}) \to \mathfrak{Emb}(\mathcal{K})$ is an $\aleph_0$-local functor if and only if it is an κ-word construction.

Proof. ($\Leftarrow$) We have already seen that all operations a word construction is built up from are $\aleph_0$-local functors. Since $\aleph_0$-local functors are closed under composition the claim follows.

($\Rightarrow$) We have to express F as composition

$$F = \mathcal{I} \circ \mathcal{S} \circ \mathcal{R} \circ \mathcal{T} \circ \mathcal{R}.$$

To define $\mathcal{T}$ we use Theorem 1.5 which tells us that F preserves direct limits. Let $\mathcal{D} : I \to \mathrm{Sub}_{\aleph_0}(\mathfrak{A})$ be the canonical diagram with limit $\varinjlim \mathcal{D} = \mathfrak{A}$. We are looking for an operation mapping $\mathfrak{A}$ to $\varinjlim(F \circ \mathcal{D})$.

Fix an enumeration $(\mathfrak{C}_\alpha)_{\alpha<\lambda}$ of $\bigcup_{\mathfrak{A}\in\mathcal{K}} \mathrm{Sub}_{\aleph_0}(\mathfrak{A})$. Note that each structure $\mathfrak{C}_\alpha$ has at most $|\Sigma| \oplus \aleph_0$ elements. Hence, there are at most $2^{|\Sigma|\oplus\aleph_0}$ of them and we have $\lambda \leq 2^{|\Sigma|\oplus\aleph_0} < \kappa$.

For each $\alpha < \lambda$, we choose a finite tuple $\bar{c}_\alpha \subseteq C_\alpha$ generating $\mathfrak{C}_\alpha$. Set

$$\Xi := \{ f^\alpha_b \mid \alpha < \lambda,\ b \in F(\mathfrak{C}_\alpha) \},$$

where f_b^α is a new function symbol of arity $|\bar{c}_\alpha|$. Note that $|\Xi| < \kappa$ since $\lambda < \kappa$ and $|F(\mathfrak{C}_\alpha)| < \kappa$, for all α. For $\mathcal{T}$ we choose the Ξ-term-algebra operation $\mathfrak{A} \mapsto \mathcal{T}[\Xi, \mathfrak{A}]$. The inverse reduct $\mathcal{S}$ maps each element to the correct sort.

The main work is done by the interpretation $\mathcal{I}$. It creates the structures $F(\mathfrak{C}_\alpha)$ and pastes them together. The domain formula $\delta(x)$ states that x is a term of the form $f_b^\alpha(\bar{a})$, for some $\alpha < \lambda$ and $b \in F(\mathfrak{C}_\alpha)$, such that the substructure generated by $\bar{a}$ is isomorphic to $\mathfrak{C}_\alpha$. Each relation $R \in \Gamma$ can be defined by a formula $\varphi_R(\bar{x})$ stating that $x_i = f_{b_i}^{\alpha_i}(\bar{a})$ and the tuple $\bar{b}$ is in the relation $R^{F(\mathfrak{C}_\alpha)}$. The functions in Γ are defined in the same way. Two elements $f_b^\alpha(\bar{a})$ and $f_{b'}^{\alpha'}(\bar{a}')$ are defined to be equal iff we have $i(b) = i'(b')$ where $i : \mathfrak{C}_\alpha \to \langle\!\langle \bar{c}_\alpha \bar{c}_{\alpha'} \rangle\!\rangle_{\mathfrak{A}}$ and $i' : \mathfrak{C}_{\alpha'} \to \langle\!\langle \bar{c}_\alpha \bar{c}_{\alpha'} \rangle\!\rangle_{\mathfrak{A}}$ are the canonical inclusion maps. Since $\lambda < \kappa$ and every $\mathfrak{C}_\alpha$ has less than κ elements, it follows that each of the above statements can be expressed in $\mathrm{FO}_{\kappa\aleph_0}$. □

Corollary 2.9. *Let $F : \mathfrak{Emb}(\mathcal{C}) \to \mathfrak{Emb}(\mathcal{K})$ be $\aleph_0$-local and let Σ be the signature of $\mathcal{C}$. If κ is a cardinal such that*

$$\kappa > 2^{|\Sigma| \oplus \aleph_0} \quad \textit{and} \quad \kappa > |F(\mathfrak{C})|\,, \quad \textit{for all finitely generated } \mathfrak{C} \in \mathcal{C}\,,$$

then $\mathfrak{A} \equiv_{\mathrm{FO}_{\kappa\aleph_0}} \mathfrak{B}$ implies $F(\mathfrak{A}) \equiv_{\mathrm{FO}_{\kappa\aleph_0}} F(\mathfrak{B})$.

Remark. We have characterised $\aleph_0$-local functors in terms of word operations and we have shown that they preserve $\mathrm{FO}_{\infty\aleph_0}$-equivalence. These results can be generalised to κ-local functors for arbitrary cardinals κ. To do so we have to allow term algebras with operations of infinite arity less than κ. It follows that these operations preserve equivalence for the logic $\mathrm{FO}_{\infty\kappa}$ which extends $\mathrm{FO}_{\infty\aleph_0}$ by quantifiers $\exists\{\, x_i \mid i < \alpha \,\}$ and $\forall\{\, x_i \mid i < \alpha \,\}$ over sets of $\alpha < \kappa$ variables. We can give a back-and-forth characterisation of this logic if we replace the usual back-and-forth property by the requirement that, for every tuple $\bar{c}$ with $|\bar{c}| < \kappa$, we can find a corresponding tuple $\bar{d}$ in the other structure.

As an application of word constructions we consider varieties. With each variety $\mathcal{V}$ we can associated a so-called *replica functor* that maps a given structure to its closest approximation in $\mathcal{V}$.

Definition 2.10. Let $\Sigma \subseteq \Sigma_+$ be signatures, $P \in \Sigma_+ \setminus \Sigma$ a unary predicate, and $\mathcal{V}$ a quasivariety of Σ_+-structures.

The *replica functor* $R_{\mathcal{V}} : \mathfrak{Hom}(\Sigma) \to \mathfrak{Hom}(\mathcal{V})$ of $\mathcal{V}$ maps an arbitrary Σ-structure $\mathfrak{A}$ to the free model of the $\mathcal{V}$-presentation $\langle A; \Phi_{\mathfrak{A}} \rangle$ where

$$\Phi_{\mathfrak{A}} := \{ Pa \mid a \in A \} \cup \{ \varphi(\bar{a}) \mid \varphi \text{ atomic}, \bar{a} \subseteq A,\ \mathfrak{A} \vDash \varphi(\bar{a}) \}.$$

Remark. Note that replica functors differ from the functors considered so far since, in general, they do not preserve embeddings. Hence, they are functors $\mathfrak{Hom}(\Sigma) \to \mathfrak{Hom}(\Sigma_+)$, and not $\mathfrak{Emb}(\Sigma) \to \mathfrak{Emb}(\Sigma_+)$.

Lemma 2.11. *The replica functor $R_{\mathcal{V}} : \mathfrak{Hom}(\Sigma) \to \mathfrak{Hom}(\mathcal{V})$ is a functor.*

Proof. Let $h : \mathfrak{A} \to \mathfrak{B}$ be a homomorphism. By definition, the structure $R_{\mathcal{V}}(\mathfrak{A})$ is the free model of $\langle A; \Phi_{\mathfrak{A}} \rangle$. Let $\bar{a}$ be an enumeration of A and set $\bar{b} := h(\bar{a})$. Since homomorphisms preserve atomic formulae it follows that

$$\langle R_{\mathcal{V}}(\mathfrak{B}), \bar{b} \rangle \vDash \Phi_{\mathfrak{A}},$$

that is, $R_{\mathcal{V}}(\mathfrak{B})$ is a model of $\langle A; \Phi_{\mathfrak{A}} \rangle$. Since $R_{\mathcal{V}}(\mathfrak{A})$ is the free model of this presentation there exists a unique homomorphism $g : R_{\mathcal{V}}(\mathfrak{A}) \to R_{\mathcal{V}}(\mathfrak{B})$ with $g \restriction A = h$. It is straightforward to check that we obtain a functor if we define $R_{\mathcal{V}}(h) := g$. ☐

Proposition 2.12. *Each replica functor is a word construction.*

Proof. Since the structure $R_{\mathcal{V}}(\mathfrak{A})$ is generated by the set A there exists a homomorphism $\mathfrak{T}[\Sigma_+, A] \to R_{\mathcal{V}}(\mathfrak{A})$ such that $h \restriction A = \mathrm{id}_A$. We define a quantifier-free interpretation $\mathcal{I}$ such that

$$R_{\mathcal{V}} = \mathcal{I} \circ \mathcal{S} \circ \mathcal{R} \circ \mathcal{T} \circ \mathcal{R},$$

where $\mathcal{T}(\mathfrak{A}) := \mathcal{T}[\Sigma_+, \mathfrak{A}]$ and $\mathcal{S}$ is the inverse reduct that maps every sort $t \in T[\Sigma_+, S_0]$ of $\mathcal{T}[\Sigma_+, \mathfrak{A}]$ to the sort s such that $t \in T_s[\Sigma_+, S_0]$.

According to Lemma D2.4.2, we have

$$R_{\mathcal{V}}(\mathfrak{A}) \vDash \psi(\bar{a}) \quad \text{iff} \quad \mathrm{Th}(\mathcal{V}) \vDash \bigwedge \Phi_{\mathfrak{A}} \to \psi(\bar{a}),$$

for every atomic formula $\psi(\bar{x}) \in \mathrm{FO}^{<\omega}[\Sigma_+]$ and all $\bar{a} \subseteq A$.

Note that, by the interpolation theorem, we have

$$\mathrm{Th}(\mathcal{V}) \vDash \bigwedge \Phi_{\mathfrak{A}} \to \psi(\bar{a}) \quad \text{iff} \quad \mathrm{Th}(\mathcal{V}) \vDash \bigwedge \Phi_{\langle\!\langle \bar{a} \rangle\!\rangle_{\mathfrak{A}}} \to \psi(\bar{a}).$$

For each atomic formula $\psi(\bar{x})$, we define

$$D_\psi := \{ \langle\!\langle \bar{a} \rangle\!\rangle_{\mathfrak{A}} \mid \mathrm{Th}(\mathcal{V}) \vDash \bigwedge \Phi_{\mathfrak{A}} \to \psi(\bar{a}) \}.$$

Let $\eta_\psi(\bar{x})$ be the $\mathrm{FO}_{\infty\aleph_0}$-formula expressing that

$$\langle\!\langle \bar{x} \rangle\!\rangle_{\mathfrak{A}} \cong \mathfrak{C}, \quad \text{for some } \mathfrak{C} \in D_\psi.$$

It follows that

$$R_{\mathcal{V}}(\mathfrak{A}) \vDash \psi(\bar{a}) \quad \text{iff} \quad \langle\!\langle \bar{a} \rangle\!\rangle_{\mathfrak{A}} \in D_\psi \quad \text{iff} \quad \mathfrak{A} \vDash \eta_\psi.$$

Consequently, we can define the desired interpretation

$$\mathcal{I} = \langle \alpha, (\delta_s)_{s \in S}, (\varepsilon_s)_{s \in S}, (\varphi_\xi)_{\xi \in \Sigma_+} \rangle$$

by setting

$$\begin{aligned} \alpha &:= \text{true}, \\ \delta_s(x) &:= \text{true}, \\ \varepsilon_s(x, y) &:= \text{“}x = s(\bar{a}) \text{ and } y = t(\bar{b}) \text{ and } \mathfrak{A} \vDash \eta_{s(\bar{x})=t(\bar{y})}(\bar{a}, \bar{b})\text{”}, \\ \varphi_\xi(\bar{x}) &:= \text{“}x_i = t_i(\bar{a}_i) \text{ and } \mathfrak{A} \vDash \eta_{R\bar{t}}(\bar{a}_0, \dots, \bar{a}_{n-1})\text{”}. \end{aligned}$$ □

3. Ehrenfeucht-Mostowski models

If a functor F is $\aleph_0$-local then with every element c of $F(\mathfrak{A})$ we can associate some finitely generated substructure $\mathfrak{A}_0 \subseteq \mathfrak{A}$ such that c is contained in $F(\mathfrak{A}_0)$. We can think of the generators of $\mathfrak{A}_0$ as a code for c. In general, c can have several such codes and the connection between c and its codes is rather loose. In order to obtain a tighter relationship and a canonical way to encode elements of $F(\mathfrak{A})$, we add a function $s : A \to F(\mathfrak{A})$ assigning to every element a of $\mathfrak{A}$ some element of $F(\mathfrak{A})$ encoded by a.

Definition 3.1. Let $\mathcal{K}$ be a class of Γ-structures and Σ a signature. A functor $F : \mathfrak{Emb}(\mathcal{K}) \to \mathfrak{Emb}(\Sigma)$ is *strongly local* if there exists a family of injective functions $s_{\mathfrak{I}} : I \to F(\mathfrak{I})$, for $\mathfrak{I} \in \mathcal{K}$, such that

- $F(\mathfrak{I})$ is generated by $\operatorname{rng} s_{\mathfrak{I}}$ and
- $F(h) \circ s_{\mathfrak{I}} = s_{\mathfrak{K}} \circ h$, for every embedding $h : \mathfrak{I} \to \mathfrak{K}$.

We call $s_{\mathfrak{I}}$ the *spine* of $F(\mathfrak{I})$.

Remark. Translated into category-theoretical terms the second of the above conditions on $s_{\mathfrak{I}}$ simply means that $(s_{\mathfrak{I}})_{\mathfrak{I}}$ is a natural transformation between the functors U and $V \circ F$, where

$$U : \mathfrak{Emb}(\mathcal{K}) \to \mathfrak{Set} \quad \text{and} \quad V : \mathfrak{Emb}(\Sigma) \to \mathfrak{Set}$$

are the forgetful functors mapping a structure to its universe.

Every strongly local functor is $\aleph_0$-local. For the proof we need a technical lemma.

Lemma 3.2. *Let $F : \mathfrak{Emb}(\mathcal{K}) \to \mathfrak{Emb}(\Sigma)$ be a strongly local functor and $h : \mathfrak{I} \to \mathfrak{K}$ an embedding in $\mathcal{K}$. Then*

$$F(h) : F(\mathfrak{I}) \cong \langle\!\langle s_{\mathfrak{K}}[\operatorname{rng} h] \rangle\!\rangle_{F(\mathfrak{K})} .$$

Proof. It is sufficient to show that $\operatorname{rng} F(h) = \langle\!\langle s_{\mathfrak{K}}[\operatorname{rng} h] \rangle\!\rangle_{F(\mathfrak{K})}$. Note that $F(h) \circ s_{\mathfrak{I}} = s_{\mathfrak{K}} \circ h$ implies

$$F(h)[\operatorname{rng} s_{\mathfrak{I}}] = s_{\mathfrak{K}}[\operatorname{rng} h] .$$

Therefore, $\langle\!\langle \operatorname{rng} s_{\mathfrak{I}} \rangle\!\rangle_{F(\mathfrak{I})} = F(I)$ implies

$$\begin{aligned} \operatorname{rng} F(h) &= F(h)\big[\langle\!\langle \operatorname{rng} s_{\mathfrak{I}} \rangle\!\rangle_{F(\mathfrak{I})}\big] \\ &= \big\langle\!\big\langle F(h)[\operatorname{rng} s_{\mathfrak{I}}] \big\rangle\!\big\rangle_{F(\mathfrak{K})} = \langle\!\langle s_{\mathfrak{K}}[\operatorname{rng} h] \rangle\!\rangle_{F(\mathfrak{K})} . \end{aligned} \qquad \square$$

Proposition 3.3. *Let $F : \mathfrak{Emb}(\mathcal{K}) \to \mathfrak{Emb}(\Sigma)$ be a strongly local functor where $\mathcal{K}$ is $\aleph_0$-hereditary. Then F is $\aleph_0$-local.*

Proof. Fix $\mathfrak{I} \in \mathcal{K}$ and suppose that $X \subseteq F(\mathfrak{I})$ is finite. Then there is a finite subset $Z \subseteq \operatorname{rng} s_{\mathfrak{I}}$ such that $X \subseteq \langle\!\langle Z \rangle\!\rangle_{F(\mathfrak{I})}$. Set

$$\mathfrak{I}_0 := \langle\!\langle s_{\mathfrak{I}}^{-1}[Z] \rangle\!\rangle_{\mathfrak{I}} .$$

Note that $\mathfrak{I}_0 \in \mathcal{K}$ since $\mathcal{K}$ is $\aleph_0$-hereditary. By Lemma 3.2, it follows that

$$X \subseteq \langle\!\langle Z \rangle\!\rangle_{F(\mathfrak{I})} = \langle\!\langle \operatorname{rng} s_{\mathfrak{I}_0} \rangle\!\rangle_{F(\mathfrak{I})} \cong F(\mathfrak{I}_0) . \qquad \square$$

By Corollary 2.9 it follows that strongly local functors preserve $\mathrm{FO}_{\kappa\aleph_0}$-equivalence.

Corollary 3.4. *Let $F : \mathfrak{Emb}(\mathcal{K}) \to \mathfrak{Emb}(\Sigma)$ be a strongly local functor where $\mathcal{K}$ is an $\aleph_0$-hereditary class of Γ-structures. For every cardinal $\kappa \geq 2^{|\Gamma| \oplus \aleph_0}$ and all $\mathfrak{I}, \mathfrak{K} \in \mathcal{K}$,*

$$\mathfrak{I} \equiv_{\mathrm{FO}_{\kappa\aleph_0}} \mathfrak{K} \quad \text{implies} \quad F(\mathfrak{I}) \equiv_{\mathrm{FO}_{\kappa\aleph_0}} F(\mathfrak{K}) .$$

Strongly local functors also preserve QF-equivalence.

Lemma 3.5. *Suppose that $F : \mathfrak{Emb}(\mathcal{K}) \to \mathfrak{Emb}(\Sigma)$ is a strongly local functor where the class $\mathcal{K}$ is $\aleph_0$-hereditary.*

Let $\mathfrak{I}, \mathfrak{K} \in \mathcal{K}$ be structures and $\bar{a} \subseteq I$ and $\bar{b} \subseteq K$ finite tuples. Then

$$\langle \mathfrak{I}, \bar{a} \rangle \equiv_0 \langle \mathfrak{K}, \bar{b} \rangle \quad \text{implies} \quad \big\langle F(\mathfrak{I}), s_{\mathfrak{I}}(\bar{a}) \big\rangle \equiv_0 \big\langle F(\mathfrak{K}), s_{\mathfrak{K}}(\bar{b}) \big\rangle .$$

Proof. Set $\mathfrak{L} := \langle\!\langle \bar{a} \rangle\!\rangle_{\mathfrak{I}}$ and let $s_{\mathfrak{L}}$ be the spine of $\mathfrak{L}$. Since $\mathcal{K}$ is $\aleph_0$-hereditary we have $\mathfrak{L} \in \mathcal{K}$. Since $\langle \mathfrak{I}, \bar{a} \rangle \equiv_0 \langle \mathfrak{K}, \bar{b} \rangle$, there are embeddings $f : \mathfrak{L} \to \mathfrak{I}$ and $g : \mathfrak{L} \to \mathfrak{K}$ with $f(\bar{a}) = \bar{a}$ and $g(\bar{a}) = \bar{b}$. Note that

$$(F(f) \circ s_{\mathfrak{L}})(\bar{a}) = (s_{\mathfrak{I}} \circ f)(\bar{a}) = s_{\mathfrak{I}}(\bar{a}) ,$$

$$\text{and} \quad (F(g) \circ s_{\mathfrak{L}})(\bar{a}) = (s_{\mathfrak{K}} \circ g)(\bar{a}) = s_{\mathfrak{K}}(\bar{b}) .$$

Since embeddings preserve every quantifier-free formula φ, it follows that

$$\begin{aligned} F(\mathfrak{J}) \vDash \varphi(s_{\mathfrak{J}}(\bar{a})) \quad &\text{iff} \quad F(\mathfrak{L}) \vDash \varphi(s_{\mathfrak{L}}(\bar{a})) \\ &\text{iff} \quad F(\mathfrak{K}) \vDash \varphi(s_{\mathfrak{K}}(\bar{b})) . \end{aligned}$$

□

Corollary 3.6. *Let $F : \mathfrak{Emb}(\mathcal{K}) \to \mathfrak{Emb}(\Sigma)$ be a strongly local functor where the class $\mathcal{K}$ is $\aleph_0$-hereditary. For every $\mathfrak{J} \in \mathcal{K}$, the spine $s_{\mathfrak{J}}$ of $F(\mathfrak{J})$ is a* QF*-indiscernible system over $\mathfrak{J}$.*

Next we study the first-order theory of structures in the range of a strongly local functor.

Proposition 3.7. *Let $F : \mathfrak{Emb}(\mathcal{K}) \to \mathfrak{Emb}(\Sigma)$ be a strongly local functor and $\mathfrak{U} \in \mathcal{K}$ an $\aleph_0$-universal structure. If $\mathrm{Th}(F(\mathfrak{U}))$ is a Skolem theory then $\mathrm{Th}(F)$ is complete. In particular,*

$$F(\mathfrak{J}) \equiv F(\mathfrak{K}) , \quad \text{for all } \mathfrak{J}, \mathfrak{K} \in \mathcal{K} .$$

Furthermore, each spine $s_{\mathfrak{J}}$ is an indiscernible system over $\mathfrak{J}$.

Proof. A Skolem theory is $\forall$-axiomatisable and admits quantifier elimination. Let $\Phi \subseteq \forall$ be an axiom system for $\mathrm{Th}(F(\mathfrak{U}))$. By Lemma 1.10. we have $\Phi \subseteq \mathrm{Th}(F)$. Hence,

$$\mathrm{Th}(F(\mathfrak{U})) = \Phi^{\vDash} \subseteq \mathrm{Th}(F) \subseteq \mathrm{Th}(F(\mathfrak{U}))$$

implies that $F(\mathfrak{J}) \equiv F(\mathfrak{U})$, for all $\mathfrak{J} \in \mathcal{K}$.

To show that every spine $s_{\mathfrak{J}}$ is indiscernible, fix $\mathfrak{J} \in \mathcal{K}$ and let $\bar{c}, \bar{d} \subseteq I$ be tuples with $\mathrm{atp}(\bar{c}) = \mathrm{atp}(\bar{d})$. For every formula $\varphi(\bar{x})$, there exists a quantifier-free formula $\psi(\bar{x})$ with $F(\mathfrak{J}) \vDash \varphi \leftrightarrow \psi$. By Lemma 3.5, it follows that

$$\begin{aligned} F(\mathfrak{J}) \vDash \varphi(s_{\mathfrak{J}}[\bar{c}]) \quad &\text{iff} \quad F(\mathfrak{J}) \vDash \psi(s_{\mathfrak{J}}[\bar{c}]) \\ &\text{iff} \quad F(\mathfrak{J}) \vDash \psi(s_{\mathfrak{J}}[\bar{d}]) \\ &\text{iff} \quad F(\mathfrak{J}) \vDash \varphi(s_{\mathfrak{J}}[\bar{d}]) . \end{aligned}$$

□

Existence and uniqueness of strongly local functors is proved in the following proposition.

Proposition 3.8. *Let $\mathfrak{A}$ be a Σ-structure, $\mathfrak{U}$ a Γ-structure, and set*

$$\mathcal{K} := \{ \mathfrak{J} \mid \mathrm{Sub}_{\aleph_0}(\mathfrak{J}) \subseteq \mathrm{Sub}_{\aleph_0}(\mathfrak{U}) \} .$$

Suppose that $\mathfrak{A}$ is generated by a QF-*indiscernible system $a : U \to A$ over $\mathfrak{U}$. Up to natural isomorphism there exists a unique strongly local functor $F : \mathfrak{Emb}(\mathcal{K}) \to \mathfrak{Emb}(\Sigma)$ such that*

$$F(\mathfrak{U}) \cong \mathfrak{A} \quad \textit{and} \quad \mathrm{Av}_{\mathrm{QF}}(s_{\mathfrak{U}}) = \mathrm{Av}_{\mathrm{QF}}(a) .$$

Proof. For each $\mathfrak{J} \in \mathcal{K}$, we define a set $\Phi(\mathfrak{J}) \subseteq \mathrm{QF}^{\circ}[\Sigma_I]$ by

$$\Phi(\mathfrak{J}) := \{ \varphi(\bar{c}) \mid \bar{c} \subseteq I \text{ and } \varphi(\bar{x}) \in \mathrm{Av}_{\mathrm{QF}}(a)(\mathrm{atp}(\bar{c}/\mathfrak{J})) \} .$$

We claim that $\Phi(\mathfrak{J})$ is $=$-closed. Since every type $\mathfrak{q}$ contains the equation $t(\bar{x}) = t(\bar{x})$, we have

$$t(\bar{c}) = t(\bar{c}) \in \Phi(\mathfrak{J}) , \quad \text{for every term } t(\bar{c}) \in T[\Sigma_I, \varnothing] .$$

Furthermore, if $\Phi(\mathfrak{J})$ contains the formulae $\varphi(t(\bar{c}), \bar{c})$ and $t(\bar{c}) = t'(\bar{c})$ then

$$\varphi(t(\bar{x}), \bar{x}),\ t(\bar{x}) = t'(\bar{x}) \in \mathrm{Av}_{\mathrm{QF}}(a)(\mathrm{atp}(\bar{c}/\mathfrak{J}))$$

implies

$$\varphi(t'(\bar{x}), \bar{x}) \in \mathrm{Av}_{\mathrm{QF}}(a)(\mathrm{atp}(\bar{c}/\mathfrak{J})) .$$

Consequently, $\varphi(t'(\bar{c}), \bar{c}) \in \Phi(\mathfrak{J})$. Hence, we can use Lemma C2.4.4 to construct a Herbrand model $\mathfrak{H}(\mathfrak{J})$ of $\Phi(\mathfrak{J})$ such that

$$\Phi(\mathfrak{J}) = \{ \varphi \in \mathrm{QF}^{\circ}[\Sigma_I] \mid \mathfrak{H}(\mathfrak{J}) \vDash \varphi \} .$$

We define the desired strongly local functor by setting

$$F(\mathfrak{J}) := \mathfrak{H}(\mathfrak{J})|_{\Sigma} \quad \text{and} \quad s_{\mathfrak{J}}(c) := c^{\mathfrak{H}(\mathfrak{J})} , \quad \text{for } c \in I .$$

First, note that the mapping $s_{\mathfrak{J}}$ is injective since we have $x_0 \neq x_1 \in \mathrm{tp}(a[vv'])$, for all elements $v \neq v'$ of $\mathfrak{U}$. Furthermore, if $h : \mathfrak{J} \to \mathfrak{K}$ is an embedding, $\bar{c} \subseteq I$, and $\varphi(\bar{x})$ quantifier-free, then

$$\begin{aligned} F(\mathfrak{J}) \vDash \varphi(s_{\mathfrak{J}}(\bar{c})) \quad &\text{iff} \quad \varphi(\bar{x}) \in \mathrm{Av}_{\mathrm{QF}}(a)(\mathrm{atp}(\bar{c}/\mathfrak{J})) \\ &\text{iff} \quad \varphi(\bar{x}) \in \mathrm{Av}_{\mathrm{QF}}(a)(\mathrm{atp}(h(\bar{c})/\mathfrak{K})) \\ &\text{iff} \quad F(\mathfrak{K}) \vDash \varphi(s_{\mathfrak{K}}(h(\bar{c}))) \,. \end{aligned}$$

By the Diagram Lemma it follows that the function

$$F(h) : t^{F(\mathfrak{J})}(s_{\mathfrak{J}}(\bar{c})) \mapsto t^{F(\mathfrak{K})}(s_{\mathfrak{K}}(h(\bar{c})))$$

is an embedding $F(h) : F(\mathfrak{J}) \to F(\mathfrak{K})$. Consequently, F is a functor. By construction, it further follows that it is strongly local, that $F(\mathfrak{U}) \cong \mathfrak{A}$, and that $\mathrm{Av}_{\mathrm{QF}}(s_{\mathfrak{U}}) = \mathrm{Av}_{\mathrm{QF}}(a)$. Hence, it remains to check uniqueness.

Suppose that G is another strongly local functor such that $G(\mathfrak{U}) \cong \mathfrak{A}$ and $\mathrm{Av}_{\mathrm{QF}}(s'_{\mathfrak{U}}) = \mathrm{Av}_{\mathrm{QF}}(a)$, where $s'_{\mathfrak{U}}$ is the spine of $G(\mathfrak{U})$. For every $\mathfrak{J} \in \mathcal{K}$, each finite tuple $\bar{c} \subseteq I$, and all quantifier-free formulae $\varphi(\bar{x})$, it follows that

$$\begin{aligned} G(\mathfrak{J}) \vDash \varphi(s'_{\mathfrak{J}}(\bar{c})) \quad &\text{iff} \quad G(\mathfrak{U}) \vDash \varphi((s'_{\mathfrak{U}} \circ g)(\bar{c})) \\ &\text{iff} \quad \varphi(\bar{x}) \in \mathrm{Av}_{\mathrm{QF}}(s'_{\mathfrak{U}})(\mathrm{atp}(g(\bar{c})/\mathfrak{U})) \\ &\text{iff} \quad \varphi(\bar{x}) \in \mathrm{Av}_{\mathrm{QF}}(a)(\mathrm{atp}(g(\bar{c})/\mathfrak{U})) \\ &\text{iff} \quad \varphi(\bar{x}) \in \mathrm{Av}_{\mathrm{QF}}(a)(\mathrm{atp}(\bar{c}/\mathfrak{J})) \\ &\text{iff} \quad F(\mathfrak{J}) \vDash \varphi(s_{\mathfrak{J}}(\bar{c})) \,, \end{aligned}$$

where $g : \langle\!\langle \bar{c} \rangle\!\rangle_{\mathfrak{J}} \to \mathfrak{U}$ is an arbitrary embedding and $s'_{\mathfrak{J}}$ and $s'_{\mathfrak{U}}$ are the spines of $G(\mathfrak{J})$ and $G(\mathfrak{U})$, respectively. Since $F(\mathfrak{J})$ and $G(\mathfrak{J})$ are generated by, respectively, $\mathrm{rng}\, s_{\mathfrak{J}}$ and $\mathrm{rng}\, s'_{\mathfrak{J}}$ it follows that we obtain an isomorphism $\pi : F(\mathfrak{J}) \to G(\mathfrak{J})$ by setting

$$\pi\big(t^{F(\mathfrak{J})}(s_{\mathfrak{J}}(\bar{c}))\big) := t^{G(\mathfrak{J})}(s'_{\mathfrak{J}}(\bar{c})) \,,$$

for all terms $t(\bar{x})$ and all $\bar{c} \subseteq I$. □

Of particular importance are strongly local functors $F : \mathfrak{Emb}(\mathcal{L}) \to \mathfrak{Emb}(\Sigma)$ where $\mathcal{L}$ is the class of all linear orders. This is mainly due to the fact that we always can find enough indiscernible sequences, whereas arbitrary indiscernible systems do not need to exist. Note that $\mathcal{L}$ is hereditary and every infinite linear order is $\aleph_0$-universal.

Definition 3.9. Let $\mathfrak{Lin} := \mathfrak{Emb}(\mathcal{L})$ where $\mathcal{L}$ is the class of all linear orders.

(a) A strongly local functor $F : \mathfrak{Lin} \to \mathfrak{Emb}(\Sigma)$ is called an *Ehrenfeucht-Mostowski* functor. We say that F is an Ehrenfeucht-Mostowski functor *for a theory T* if F is an Ehrenfeucht-Mostowski functor such that $F(I) \vDash T$, for every linear order I.

(b) Let T be a first-order theory. An *Ehrenfeucht-Mostowski model* of T is a model of the form $F(I)$ where F is some Ehrenfeucht-Mostowski functor for T and I is a linear order.

(c) Let $F : \mathfrak{Lin} \to \mathfrak{Emb}(\Sigma)$ be an Ehrenfeucht-Mostowski functor. The *average type* of F is the set

$$\begin{aligned}\operatorname{Av}(F) := \big\{ \varphi(\bar{x}) \in \mathrm{FO}^{<\omega}[\Sigma] \bigm| \\ F(\mathfrak{I}) \vDash \varphi(s_{\mathfrak{I}}(\bar{c})) \text{ for all } \mathfrak{I} \in \mathcal{K} \text{ and } \bar{c} \in [I]^{<\omega} \big\} .\end{aligned}$$

Note that, by Proposition 3.7 and Lemma 3.5, the average type of an Ehrenfeucht-Mostowski function is complete.

Lemma 3.10. *If $F : \mathfrak{Lin} \to \mathfrak{Emb}(\Sigma)$ is an Ehrenfeucht-Mostowski functor, then* $\operatorname{Av}(F)$ *is a complete type.*

Theorem 3.11 (Ehrenfeucht-Mostowski). *Let $\mathfrak{M}$ be a model of a Skolem theory T. For every sequence $(a^i)_{i \in I}$ of distinct elements in $\mathfrak{M}$ there exists an Ehrenfeucht-Mostowski functor F for T such that*

$$\operatorname{Av}((a^i)_i/\varnothing) \subseteq \operatorname{Av}(F) .$$

Proof. By Proposition E5.3.6, there exists an elementary extension $\mathfrak{N} \geq \mathfrak{M}$ containing an indiscernible sequence $(c_n)_{n<\omega}$ with

$$\operatorname{Av}((a^i)_{i \in I}/\varnothing) \subseteq \operatorname{Av}((c_n)_{n<\omega}/\varnothing) .$$

Let $s : \omega \to N$ be the function mapping $n < \omega$ to c_n and set

$$\mathfrak{U} := \langle\!\langle \operatorname{rng} s \rangle\!\rangle_{\mathfrak{N}} .$$

Note that the function s is injective, since $x_0 \neq x_1 \in \operatorname{Av}((a^i)_i/\varnothing)$. Furthermore, we have $\mathfrak{U} \preceq \mathfrak{N}$ since T is a Skolem theory. Hence, we can use Proposition 3.8 to find an Ehrenfeucht-Mostowski functor F with $F(\omega) = \mathfrak{U}$ and $s_\omega = s$. It follows that $\operatorname{Av}((a^i)_i/\varnothing) \subseteq \operatorname{Av}((c_n)_n/\varnothing) = \operatorname{Av}(F)$. □

Corollary 3.12. *If a first-order theory T has infinite models then there exists an Ehrenfeucht-Mostowski functor for T.*

Proof. Let T^+ be a Skolemisation of T. It is sufficient to find an Ehrenfeucht-Mostowski functor F for T^+ since we can obtain the desired Ehrenfeucht-Mostowski functor for T by composing F with a suitable reduct functor.

Let $\mathfrak{M}^+$ be an infinite model of T^+ that contains an indiscernible sequence $(a^n)_{n<\omega}$ of distinct elements. By Theorem 3.11, there exists an Ehrenfeucht-Mostowski functor F with $\operatorname{Av}((a^n)_n) \subseteq \operatorname{Av}(F)$. We claim that F is the desired Ehrenfeucht-Mostowski functor for T^+. As $(a^n)_n$ is indiscernible, its average type $\operatorname{Av}((a^n)_n)$ is complete and, therefore, equal to $\operatorname{Av}(F)$. Consequently, $F(\omega) \vDash T^+$. Since T^+ is a Skolem theory, it follows by Lemma 3.7 that $F(I) \vDash T^+$, for every I. □

We use Ehrenfeucht-Mostowski functors to construct models of a theory with certain properties. As a first simple application, we build models with many automorphisms.

Lemma 3.13. *Let T be a complete first-order theory with infinite models. For every cardinal $\kappa \geq |T|$, there exists a model $\mathfrak{M}$ of T of size $|M| = \kappa$ with 2^κ automorphisms.*

Proof. According to Corollary 3.12, there is an Ehrenfeucht-Mostowski functor $F : \mathfrak{Lin} \to \operatorname{Mod}(T)$ for T. We will construct a linear order I of size $|I| = \kappa$ with 2^κ automorphisms. It follows that $F(I)$ is the desired model of T.

Let $I := \mathbb{Z} \cdot \kappa$ be the product of the order $\mathbb{Z}$ of the integers and the well-order κ. For every set $X \subseteq \kappa$, we can define an automorphism $\pi_X : I \to I$ by

$$\pi_X\langle k, \alpha\rangle := \begin{cases} \langle k+1, \alpha\rangle & \text{if } \alpha \in X\,, \\ \langle k, \alpha\rangle & \text{if } \alpha \notin X\,. \end{cases}$$

Since $\pi_X \neq \pi_Y$, for $X \neq Y$, it follows that I has at least 2^κ automorphisms. □

One important application of Ehrenfeucht-Mostowski models rests on the fact that such models realise few types.

Theorem 3.14. *Let T be a Skolem theory over the signature Σ and let $\mathfrak{M}$ be an Ehrenfeucht-Mostowski model of T.*

(a) *For every finite sequence of sorts $\bar{s}$, $\mathfrak{M}$ realises at most $|\Sigma| \oplus \aleph_0$ types in $S^{\bar{s}}(T)$.*

(b) *Let s be a sort and $U \subseteq M$. If the spine of $\mathfrak{M}$ is well-ordered then $\mathfrak{M}$ realises at most $|\Sigma| \oplus |U| \oplus \aleph_0$ types in $S^s(U)$.*

Proof. (a) Suppose that $\mathfrak{M} = F(I)$ for some Ehrenfeucht-Mostowski functor F. Fix a finite tuple $\bar{s}$ of sorts and let $\bar{a} \in M^{\bar{s}}$ be a tuple of elements of the corresponding sorts. For each index l there exists a term $t_l(\bar{x})$ and an increasing tuple $\bar{\imath}^l \subseteq I$ such that $a_l = t_l^{\mathfrak{M}}(s_I[\bar{\imath}^l])$. By adding redundant variables we may assume that all the tuples $\bar{\imath}^l$ are equal. We denote this tuple by $\bar{\imath}$. If $\bar{k} \subseteq I$ is another increasing tuple of the same length then it follows from indiscernibility of the spine s_I that

$$\begin{aligned} &\mathfrak{M} \vDash \varphi(t_0(s_I[\bar{\imath}]), \dots, t_{n-1}(s_I[\bar{\imath}])) \\ \text{iff}\quad &\mathfrak{M} \vDash \varphi(t_0(s_I[\bar{k}]), \dots, t_{n-1}(s_I[\bar{k}]))\,, \end{aligned}$$

for every formula φ. Setting $b_l := t_l(s_I[\bar{k}])$ we obtain $\text{tp}(\bar{a}) = \text{tp}(\bar{b})$. Hence, the type of $\bar{a}$ is uniquely determined by the terms t_l. Since

$$|T^{<\omega}_{s_l}[\Sigma]| = |\Sigma| \oplus \aleph_0$$

it follows that $\mathfrak{M}$ realises at most $|\Sigma| \oplus \aleph_0$ types from $S^{\bar{s}}(T)$.

(b) Suppose that $\mathfrak{M} = F(\alpha)$, for some ordinal α, and let $U \subseteq M$. Each element $c \in U$ can be written as $c = t_c^{\mathfrak{M}}(s_\alpha(\bar{\imath}_c))$, for some term t_c and indices $\bar{\imath}_c \subseteq \alpha$. The set $W := \bigcup_{c \in U} \bar{\imath}_c$ has size $|W| \leq |U| \oplus \aleph_0$. Let $u(\bar{x}) \in T^{<\omega}[\Sigma]$ be a term and $\bar{k} \subseteq \alpha$. By indiscernibility of s_α the type of $u^{\mathfrak{M}}(\bar{k})$ is determined by the relative position of $\bar{k}$ with respect to the elements of W. Since α is well-ordered, there are at most $|W| \oplus \aleph_0$ ways in which $\bar{k}$ can lie relative to W. Consequently, the elements $u^{\mathfrak{M}}(\bar{k})$ with $\bar{k} \subseteq \alpha$ realise at most $|W| \oplus \aleph_0$ complete types over U. Therefore, at most

$$|T^{<\omega}_s[\Sigma]| \oplus |W| \oplus \aleph_0 \leq |\Sigma| \oplus |U| \oplus \aleph_0$$

complete s-types over U are realised in $\mathfrak{M}$. □

Corollary 3.15. *Let T be a complete first-order theory with infinite models. For every cardinal $\kappa \geq |T|$, T has an Ehrenfeucht-Mostowski model $\mathfrak{M}$ of size $|M| = \kappa$ such that, for every set $U \subseteq M$ and every finite tuple $\bar{s}$ of sorts, $\mathfrak{M}$ realises at most $|U| \oplus |T|$ types from $S^{\bar{s}}(U)$.*

Proof. According to Corollary 3.12, there is an Ehrenfeucht-Mostowski functor $F : \mathfrak{Lin} \to \text{Mod}(T)$ for T. Let $\mathfrak{M} := F(\kappa)$. Then $|M| = \kappa$ and, by Theorem 3.14 (b), $\mathfrak{M}$ realises at most $|U| \oplus |T|$ types in $S^s(U)$, for every $U \subseteq M$. For a finite tuple $\bar{s} = s_0 \dots s_{n-1}$ it follows by induction that $\mathfrak{M}$ realises at most $(|U| \oplus |T|)^n = |U| \oplus |T|$ types in $S^{\bar{s}}(U)$. □

Theorem 3.16. *Let Σ be a signature. If a theory T over Σ is κ-categorical for some $\kappa \geq |\Sigma| \oplus \aleph_0$, then T is λ-stable, for every cardinal $|\Sigma| \oplus \aleph_0 \leq \lambda < \kappa$.*

Proof. Let $\mathfrak{M}$ be the Ehrenfeucht-Mostowski model from Corollary 3.15. For a contradiction, suppose that there is some set U of size $|U| = \lambda$ with $|S^s(U)| > \lambda$. Let $\mathfrak{N}$ be a model of T containing U that realises λ^+ of these types. By the Theorem of Löwenheim and Skolem we can choose $\mathfrak{N}$ to be of size $|N| = \lambda^+ \leq \kappa$. Hence, $\mathfrak{N}$ has an elementary extension $\mathfrak{N}_+$ of size $|N_+| = \kappa$. As T is κ-categorical this implies $\mathfrak{N}_+ \cong \mathfrak{M}$ and there exists an elementary embedding $h : \mathfrak{N} \to \mathfrak{M}$. Hence, $\mathfrak{M}$ contains a subset $h[U]$ of size λ such that more than λ types over $h[U]$ are realised in $\mathfrak{M}$. This contradicts our choice of $\mathfrak{M}$. □

Corollary 3.17. *Let T be a theory over a countable signature. If T is κ-categorical for some uncountable cardinal κ then T is $\aleph_0$-stable.*

The next proposition generalises Lemma E4.1.6.

Proposition 3.18. *Let T be a countable, complete theory. If there is some finite sequence $\bar{s}$ of sorts such that $S^{\bar{s}}(T)$ is uncountable then, for each infinite cardinal κ, T has at least $2^{\aleph_0}$ pairwise non-isomorphic models of cardinality κ.*

Proof. Let κ be an infinite cardinal and fix $\bar{s}$ such that $S^{\bar{s}}(T)$ is uncountable. By Corollary B5.7.5, it follows that $|S^{\bar{s}}(T)| = 2^{\aleph_0}$. Note that this also implies that T has infinite models. Let $\bar{c}$ be a tuple of new constant symbols of sorts $\bar{s}$. For each $\mathfrak{p}(\bar{x}) \in S^{\bar{s}}(T)$ we form the theory $T_{\mathfrak{p}} := T \cup \mathfrak{p}(\bar{c})$. Let $T^+_{\mathfrak{p}}$ be a Skolemisation of $T_{\mathfrak{p}}$. We can use Theorem 3.11 to find an Ehrenfeucht-Mostowski model $\mathfrak{A}_{\mathfrak{p}}$ of $T^+_{\mathfrak{p}}$ with a spine $s_{\mathfrak{p}} : \kappa \to A_{\mathfrak{p}}$. It follows that

$$\kappa \leq |A_{\mathfrak{p}}| \leq \kappa \oplus |T^+_{\mathfrak{p}}| = \kappa \oplus \aleph_0 = \kappa\,.$$

By Theorem 3.14 $\mathfrak{A}_{\mathfrak{p}}$ realises only countably many $\bar{s}$-types. Therefore, so does $\mathfrak{B}_{\mathfrak{p}} := \mathfrak{A}_{\mathfrak{p}}|_{\Sigma}$. Furthermore, the tuple $\bar{c}^{\mathfrak{A}_{\mathfrak{p}}}$ realises the type $\mathfrak{p}$ in $\mathfrak{B}_{\mathfrak{p}}$.

We claim that there are $2^{\aleph_0}$ pairwise non-isomorphic models among the $\mathfrak{B}_{\mathfrak{p}}$. Suppose otherwise. Then there exists a set $I \subseteq S^{\bar{s}}(T)$ of size $|I| < 2^{\aleph_0}$ such that every $\mathfrak{B}_{\mathfrak{p}}$ is isomorphic to some $\mathfrak{B}_{\mathfrak{q}}$ with $\mathfrak{q} \in I$. Since every type in $S^{\bar{s}}(T)$ is realised in some $\mathfrak{B}_{\mathfrak{p}}$, but each $\mathfrak{B}_{\mathfrak{p}}$ realises only countably many types, it follows that

$$|S^{\bar{s}}(T)| \leq |I| \otimes \aleph_0 < 2^{\aleph_0}.$$

Contradiction. □

Definable linear orders in an Ehrenfeucht-Mostowski model $F(I)$ are closely related to the order induced by I. We start with a technical lemma.

Lemma 3.19. *Let $\langle A, < \rangle$ be an infinite dense linear order and suppose that $\sqsubset$ is a linear order on $[A]^n$ with the following property. For all tuples $\bar{a}, \bar{a}', \bar{b}, \bar{b}' \in [A]^n$ such that $\bar{a}\bar{b}$ and $\bar{a}'\bar{b}'$ have the same order type with respect to $<$, we have*

$$\bar{a} \sqsubset \bar{b} \quad \text{iff} \quad \bar{a}' \sqsubset \bar{b}' .$$

Then there exist a linear order $\triangleleft$ on $[n]$ and a map $\sigma : [n] \to \{-1, 1\}$ such that,

$$\bar{a} \sqsubset \bar{b}$$

iff *there is some $l \in [n]$ with $a_l <^{\sigma(l)} b_l$ and $a_i = b_i$, for $i \triangleleft l$,*

where $<^1 := <$ and $<^{-1} := >$.

Proof. We start by defining linear orders $\prec_i$ on A, for $i < n$, by

$$a \prec_i b \quad :\text{iff} \quad \bar{c}[i/a] \sqsubset \bar{c}[i/b], \quad \text{for some } \bar{c} \in [A]^n \text{ with } c_{i-1} < a < c_{i+1} \text{ and } c_{i-1} < b < c_{i+1} .$$

(Recall that, according to Definition B3.1.12, $\bar{c}[i/a]$ denotes the tuple obtained from $\bar{c}$ by replacing c_i by a.) Note that, by our assumption on $\sqsubset$, if $a \prec_i b$ holds then we have $\bar{c}[i/a] \sqsubset \bar{c}[i/b]$ for *all* tuples $\bar{c}$ satisfying the above conditions. Furthermore, since we can always find such a tuple and $\sqsubset$ is linear it follows that $a \prec_i b$ or $b \prec_i a$. Finally, if $a \prec_i b$ holds for some $a < b$ then it holds for all $a < b$. Therefore, we have $\prec_i \, = \, <$ or $\prec_i \, = \, <^{-1}$. Let $\sigma : [n] \to \{1, -1\}$ be the function with $\prec_i \, = \, <^{\sigma(i)}$.

We define the ordering $\triangleleft$ on $[n]$ by

$$i \triangleleft j \quad \text{iff} \quad i \neq j \text{ and there are } a \prec_i a', \ b \prec_j b', \text{ and } \bar{c} \text{ such that } \bar{c}[i/a, j/b'] \sqsubset \bar{c}[i/a', j/b] \text{ and these tuples are increasing.}$$

By assumption on $\sqsubset$ it follows that the definition of $i \lhd j$ does not depend on the choice of a, a', b, b' and $\bar{c}$. If there are some elements satisfying the definition above then we have $\bar{c}[i/a, j/b'] \sqsubset \bar{c}[i/a', j/b]$ for *all* elements as above. Consequently, $i \lhd j$ implies $j \ntriangleleft i$. Furthermore, since $\sqsubset$ is linear we have $i \lhd j$ or $j \lhd i$, for all i, j. In order to show that $\lhd$ is a linear order it therefore remains to prove that it is transitive.

Suppose that $i \lhd j \lhd k$. We have to show that $i \lhd k$. If $i = k$ we would have $i \lhd j$ and $j \lhd i$, which is impossible. Hence, $i \neq k$. Choose elements $a <_i a'$, $b <_k b'$, and $\bar{c}$ such that the tuples $\bar{c}[i/a, k/b']$ and $\bar{c}[i/a', k/b]$ are increasing. We claim that $\bar{c}[i/a, k/b'] \sqsubset \bar{c}[i/a', k/b]$. Since A is dense we can find some element $d <_j c_j$ such that $\bar{c}[i/a', j/d, k/b]$ is increasing. Then $i \lhd j$ implies that

$$\bar{c}[i/a, k/b'] = \bar{c}[i/a, j/c_j, k/b'] \sqsubset \bar{c}[i/a', j/d, k/b'] .$$

Similarly, $j \lhd k$ implies

$$\bar{c}[i/a', j/d, k/b'] \sqsubset \bar{c}[i/a', j/c_j, k/b] = \bar{c}[i/a', k/b] .$$

Therefore, we have

$$\bar{c}[i/a, k/b'] \sqsubset \bar{c}[i/a', k/b] ,$$

as desired.

It remains to prove that the ordering $\sqsubset$ coincides with the ordering $\sqsubset^{\sigma}_{\lhd}$ induced by $\lhd$ and σ as in the claim above. Since both relations are linear orders it is sufficient to prove that $\bar{a} \sqsubset^{\sigma}_{\lhd} \bar{b}$ implies $\bar{a} \sqsubset \bar{b}$.

For $\bar{a}, \bar{b} \in [A]^n$, let $d(\bar{a}, \bar{b})$ be the number of indices i with $a_i \neq b_i$. We prove the claim by induction on $d := d(\bar{a}, \bar{b})$. If $d = 0$ then $\bar{a} \not\sqsubset^{\sigma}_{\lhd} \bar{b}$ and there is nothing to prove.

Suppose that $d = 1$ and let l be the unique index with $a_l \neq b_l$. Then we have

$$\bar{a} \sqsubset \bar{b} \quad \text{iff} \quad a_l <_l b_l \quad \text{iff} \quad a_l <^{\sigma(l)} b_l \quad \text{iff} \quad \bar{a} \sqsubset^{\sigma}_{\lhd} \bar{b} .$$

Suppose that $d = 2$. Let l and j be the indices where $\bar{a}$ and $\bar{b}$ differ and suppose that $l \lhd j$. By definition of $\sqsubseteq^{\sigma}_{\lhd}$ we have $a_l <_l b_l$. Hence, if $b_j <_j a_j$ then $l \lhd j$ implies that

$$\bar{a} = \bar{a}[l/a_l, j/a_j] \sqsubset \bar{a}[l/b_l, j/b_j] = \bar{b}\,,$$

and we are done. Suppose therefore that $a_j <_j b_j$. Let $k_0 := \min\{l, j\}$ and $k_1 := \max\{l, j\}$ (with respect to the natural ordering on $[n]$). If $a_{k_0} < b_{k_0}$ then $\bar{a}[k_1/b_{k_1}] \in [A]^n$ and, by inductive hypothesis, we have

$$\bar{a} \sqsubset \bar{a}[k_1/b_{k_1}] = \bar{b}[k_0/a_{k_0}] \sqsubset \bar{b}\,.$$

Similarly, $b_{k_0} < a_{k_0}$ implies that

$$\bar{a} \sqsubset \bar{a}[k_0/b_{k_0}] = \bar{b}[k_1/a_{k_1}] \sqsubset \bar{b}\,.$$

Finally, suppose that $d > 2$. Let l be the $\lhd$-minimal index with $a_l \neq b_l$ and let k be the $<$-maximal one. First, consider the case that $k \neq l$. If $a_k <_k b_k$ then we have

$$\bar{a} \sqsubset^{\sigma}_{\lhd} \bar{a}[k/b_k] \sqsubset^{\sigma}_{\lhd} \bar{b}\,,$$

and the claim follows by inductive hypothesis. Therefore, suppose that $b_k <_k a_k$. Since A is dense we can find some element c with $a_l <_l c <_l b_l$ and $a_{l-1}, b_{l-1} < c < a_{l+1}, b_{l+1}$. Then

$$\bar{a} \sqsubset^{\sigma}_{\lhd} \bar{a}[l/c, k/b_k] \sqsubset^{\sigma}_{\lhd} \bar{b}\,,$$

and the claim follows by inductive hypothesis.

It remains to consider the case that $k = l$. Let k' be the $<$-minimal index with $a_{k'} \neq b_{k'}$. Then $k' \neq l$ and we can use a dual argument to show that $\bar{a} \sqsubset \bar{b}$. □

Theorem 3.20. *Let $F : \mathfrak{Lin} \to \mathfrak{Emb}(\Sigma)$ be an Ehrenfeucht-Mostowski functor and $t(x^0, \dots, x^{n-1})$ a term over Σ. Suppose that $\chi(x, y)$ is a quantifier-free formula such that $\mathrm{Av}(F)$ implies that χ linearly orders all elements of the form $t(s_I[\bar{\imath}])$ with $\bar{\imath} \in [I]^n$.*

Then there exist a linear order $\triangleleft$ *on* $[n]$ *and a map* $\sigma : [n] \to \{-1, 1\}$ *such that, for every linear order* I *and all tuples* $\bar{\imath}, \bar{\jmath} \in I^n$,

$$F(I) \vDash \chi\big(t(s_I[\bar{\imath}]), t(s_I[\bar{\jmath}])\big)$$

iff *there is some* $l \in [n]$ *with* $i_l <^{\sigma(l)} j_l$ *and* $i_s = j_s$, *for* $s \triangleleft l$,

where $<^1 := <$ *and* $<^{-1} := >$.

Proof. Note that we can embed every model $F(I)$ into a model $F(J)$ where J is a dense order. Since χ is quantifier-free it is therefore sufficient to consider the case of a dense order I. Define

$$\bar{\imath} \sqsubset \bar{\jmath} \quad :\text{iff} \quad F(I) \vDash \chi\big(t(s_I[\bar{\imath}]), t(s_I[\bar{\jmath}])\big)\,.$$

According to Lemma 3.19 the order $\sqsubset$ has the desired form. □

E7. Abstract elementary classes

1. Abstract elementary classes

For every algebraic logic L, we can form the category $\mathfrak{Emb}_L(\Sigma)$ of L-embeddings. This is a subcategory of the category $\mathfrak{Emb}(\Sigma)$ of all embeddings. It has the same objects but fewer morphisms. In this section we investigate to which extend these two categories determine L.

Definition 1.1. Suppose that $\mathcal{K}$ is a class of Σ-structures that is closed under isomorphisms and let $\mathcal{E}$ be a class of embeddings between structures in $\mathcal{K}$.

(a) The pair $\langle \mathcal{K}, \mathcal{E} \rangle$ forms an *abstract elementary class* if it satisfies the following conditions.

(i) $\mathcal{E}$ is closed under composition and it contains all isomorphisms between structures in $\mathcal{K}$.

(ii) $f, f \circ g \in \mathcal{E}$ implies $g \in \mathcal{E}$, for all embeddings f and g.

(iii) The subcategory of $\mathfrak{Emb}(\mathcal{K})$ induced by $\mathcal{E}$ has direct limits and, for every directed diagram $D : I \to \mathcal{E}$, the direct limits of D in $\mathcal{E}$ and in $\mathfrak{Emb}(\Sigma)$ coincide.

(iv) There exists a cardinal $\ln(\mathcal{K}) \geq |\Sigma| \oplus \aleph_0$ such that, for every structure $\mathfrak{M} \in \mathcal{K}$ and every set $X \subseteq M$, we can find a substructure $\mathfrak{C} \in \mathcal{K}$ of size $|C| \leq |X| \oplus \ln(\mathcal{K})$ such that $\langle\!\langle X \rangle\!\rangle_{\mathfrak{M}} \subseteq \mathfrak{C} \subseteq \mathfrak{M}$ and the inclusion map $\mathfrak{C} \to \mathfrak{M}$ belongs to $\mathcal{E}$.

The cardinal $\ln(\mathcal{K})$ is called the *Löwenheim number* of $\mathcal{K}$.

(b) Let $\langle \mathcal{K}, \mathcal{E} \rangle$ be an abstract elementary class. The elements of $\mathcal{E}$ are called *$\mathcal{K}$-embeddings*. Usually, we drop the class $\mathcal{E}$ from our notation and just write $\mathcal{K}$ for $\langle \mathcal{K}, \mathcal{E} \rangle$.

(c) Let $\langle \mathcal{K}, \mathcal{E} \rangle$ be an abstract elementary class and let $\mathfrak{A} \subseteq \mathfrak{B}$ be structures in $\mathcal{K}$. We define

$$\mathfrak{A} \leq_{\mathcal{K}} \mathfrak{B} \quad : \text{iff} \quad \text{the inclusion map } i : \mathfrak{A} \to \mathfrak{B} \text{ belongs to } \mathcal{E} .$$

If $\mathfrak{A} \leq_{\mathcal{K}} \mathfrak{B}$ then we call $\mathfrak{A}$ a $\mathcal{K}$-*substructure* of $\mathfrak{B}$.

(d) The pair $\langle \mathcal{K}, \mathcal{E} \rangle$ forms an *algebraic class* if

(i) $\mathcal{E} = \operatorname{Emb}(\mathcal{K})$ is the set of all embeddings and

(ii) $\mathcal{K}$ is closed under isomorphisms, substructures, and direct limits of embeddings.

Example. (a) Every algebraic class $\langle \mathcal{K}, \mathcal{E} \rangle$ of Σ-structures is an abstract elementary class with Löwenheim number $\operatorname{ln}(\mathcal{K}) = |\Sigma| \oplus \aleph_0$.

(b) Let $L := \mathrm{FO}_{\kappa\aleph_0}$, let $T \subseteq L^0[\Sigma]$ be a theory, and let $\mathcal{E}$ be the class of all $L^{<\omega}$-embeddings between models of T. Then $\langle \operatorname{Mod}(T), \mathcal{E} \rangle$ is an abstract elementary class and the relation $\leq_{\mathcal{K}}$ coincides with the $L^{<\omega}$-substructure relation $\leq_{L^{<\omega}}$. The same holds for many other algebraic logics L.

Exercise 1.1. In (b) of the above example we have taken for $\mathcal{E}$ all embeddings that preserve every formula with finitely many free variables. What goes wrong if we take only those embeddings that also preserve formulae with infinitely many free variables?

Exercise 1.2. Let $\langle \mathcal{K}_i, \mathcal{E}_i \rangle$, $i \in I$, be a family of abstract elementary classes over the signature Σ. Show that the intersection $\langle \bigcap_i \mathcal{K}_i, \bigcap_i \mathcal{E}_i \rangle$ is an abstract elementary class with Löwenheim number $\sup_i \operatorname{ln}(\mathcal{K}_i)$.

Remark. (a) We have defined the $\mathcal{K}$-substructure relation $\leq_{\mathcal{K}}$ in terms of the class $\mathcal{E}$ of $\mathcal{K}$-embeddings. Conversely, $\leq_{\mathcal{K}}$ determines $\mathcal{E}$ since an embedding $h : \mathfrak{A} \to \mathfrak{B}$ belongs to $\mathcal{E}$ if and only if $\operatorname{rng} h \leq_{\mathcal{K}} \mathfrak{B}$.

(b) Let $\langle \mathcal{K}, \mathcal{E} \rangle$ be an abstract elementary class and let $\mathcal{K}_0 \subseteq \mathcal{K}$ be the subclass of all structures of size at most $\operatorname{ln}(\mathcal{K})$. Every structure $\mathfrak{M} \in \mathcal{K}$ can be written as a direct limit $D : I \to \mathcal{E}$ of its $\mathcal{K}$-substructures in $\mathcal{K}_0$. Hence, $\mathcal{K}$ is the class of all direct limits of structures in $\mathcal{K}_0$. In particular, $\mathcal{K}_0$ and the restriction of $\mathcal{E}$ to $\mathcal{K}_0$ completely determine $\langle \mathcal{K}, \mathcal{E} \rangle$.

We have seen that many algebraic logics give rise to an abstract elementary class. Conversely, we can show that every such class arises from an algebraic logic in this way. To do so, we need the notion of a Galois type.

Definition 1.2. Let $\langle \mathcal{K}, \mathcal{E} \rangle$ be an abstract elementary class. Let $\mathfrak{M} \in \mathcal{K}$ be a structure and $U \subseteq M$ a set of parameters.

We define the *Galois type* of a tuple $\bar{a} \subseteq M$ over U by

$$\mathrm{tp}_{\mathrm{Aut}}(\bar{a}/\mathfrak{M}, U) := [\bar{a}, \mathfrak{M}, U]_{\approx}$$

where the equivalence relation $\approx$ is the transitive closure of the following relation $\sim$ on triples $\langle \bar{a}, \mathfrak{M}, U \rangle$ with $U, \bar{a} \subseteq M$. We set

$$\langle \bar{a}, \mathfrak{A}, U \rangle \sim \langle \bar{b}, \mathfrak{B}, V \rangle$$

iff $U = V$ and, for some $\mathfrak{M} \in \mathcal{K}$, there are $\mathcal{K}$-embeddings $f : \mathfrak{A}_0 \to \mathfrak{M}$ and $g : \mathfrak{B}_0 \to \mathfrak{M}$ where $\mathfrak{A}_0 \leq_{\mathcal{K}} \mathfrak{A}$ and $\mathfrak{B}_0 \leq_{\mathcal{K}} \mathfrak{B}$ are $\mathcal{K}$-substructures with $U \cup \bar{a} \subseteq A_0$ and $U \cup \bar{b} \subseteq B_0$ such that

$$f \upharpoonright U = g \upharpoonright U \quad \text{and} \quad f(\bar{a}) = g(\bar{b}) .$$

We write $S^{\bar{s}}_{\mathrm{Aut}}(U)$ for the set of all Galois types of $\bar{s}$-tuples over U.

Remark. (a) Let T be a first-order theory and $\mathrm{Mod}(T)$ the corresponding abstract elementary class. Then the Galois type of a tuple coincides with its first-order type.

(b) If an abstract elementary class $\mathcal{K}$ stems from an algebraic logic L then no L-formula can distinguish between tuples of the same Galois type. Hence, tuples with the same Galois type also have the same L-type. In general the converse fails.

(c) Below we will not consider Galois types over arbitrary parameters U. The set U will always be either empty or the universe of some $\mathcal{K}$-substructure $\mathfrak{U}$.

Proposition 1.3. *Let $\langle \mathcal{K}, \mathcal{E} \rangle$ be an abstract elementary class of Σ-structures. There exists an algebraic logic L, a fragment $\Delta \subseteq L^{<\omega}[\Sigma]$, and a formula $\chi \in \Delta$ such that*

$$\mathcal{K} = \mathrm{Mod}_L(\chi) \quad \textit{and} \quad \mathcal{E} \textit{ is the class of all } \Delta\textit{-embeddings.}$$

Proof. For a set X of variables, we denote by Φ_X the set of all Galois types of X-tuples over the empty set. We start by defining the functor L. For a signature Γ and a set X of variables, we set

$$L[\Gamma, X] := \wp(\Phi_X) \times \mathfrak{Sig}(\Sigma, \Gamma)\,,$$

and, for a morphism $\lambda \in \mathfrak{Sig}(\Gamma, \Gamma')$, we set

$$L[\lambda] : \langle \Psi, \mu \rangle \mapsto \langle \Psi, \lambda \circ \mu \rangle\,.$$

For a formula $\langle \Psi, \mu \rangle \in L[\Gamma, X]$, a Γ-structure $\mathfrak{A}$, and a tuple $\bar{a} \in A^X$, we define the satisfaction relation by

$$\mathfrak{A} \vDash \langle \Psi, \mu \rangle(\bar{a}) \quad :\text{iff} \quad \mathrm{tp}_{\mathrm{Aut}}(\bar{a}/\mathfrak{A}|_\mu, \varnothing) \in \Psi\,.$$

Finally, we set

$$\Delta := \{\, \langle \Psi, \mu \rangle \in L^{<\omega}[\Sigma] \mid \mu = \mathrm{id} \,\} \quad \text{and} \quad \chi := \langle \Phi_\varnothing, \mathrm{id} \rangle\,. \qquad \square$$

This proposition provides a syntax for each abstract elementary class. But because of the high degree of generality in the definition of an algebraic logic, this result is of little practical use. A more concrete way of equipping an abstract elementary class with a kind of syntax is given by the notion of a Skolem expansion.

Definition 1.4. Let $\langle \mathcal{K}, \mathcal{E} \rangle$ be an abstract elementary class of Σ-structures.

(a) An *expansion* of $\mathcal{K}$ is an abstract elementary class $\langle \mathcal{K}_+, \mathcal{E}_+ \rangle$ of Σ_+-structures, for some $\Sigma_+ \supseteq \Sigma$, such that

$$\mathrm{pr}_\Sigma(\mathcal{K}_+) = \mathcal{K}\,, \quad \mathrm{pr}_\Sigma(\mathcal{E}_+) = \mathcal{E}\,, \quad \text{and} \quad \ln(\mathcal{K}_+) = \ln(\mathcal{K})\,,$$

where $\mathrm{pr}_\Sigma : \mathfrak{Emb}(\Sigma_+) \to \mathfrak{Emb}(\Sigma)$ is the reduct functor.

(b) An expansion $\langle \mathcal{K}_+, \mathcal{E}_+ \rangle$ of $\langle \mathcal{K}, \mathcal{E} \rangle$ is a *Skolem expansion* if $\langle \mathcal{K}_+, \mathcal{E}_+ \rangle$ is an algebraic class.

Algebraic classes and, hence, Skolem expansions are very nicely behaved abstract elementary classes. For instance, the membership of a structure in such a class only depends on its finitely generated substructures.

Lemma 1.5. *Let $\mathcal{K}$ be an algebraic class and $\mathfrak{M}$ a structure. Then*

$$\mathfrak{M} \in \mathcal{K} \quad \text{iff} \quad \mathrm{Sub}_{\aleph_0}(\mathfrak{M}) \subseteq \mathcal{K} .$$

Proof. ($\Rightarrow$) Suppose that $\mathcal{K}$ is algebraic, $\mathfrak{M} \in \mathcal{K}$, and $\mathfrak{A} \subseteq \mathfrak{M}$. Since $\mathcal{K}$ is algebraic, we have $\mathfrak{A} \leq_\mathcal{K} \mathfrak{M}$. This implies that $\mathfrak{A} \in \mathcal{K}$.

($\Leftarrow$) Each structure $\mathfrak{M}$ can be written as direct limit $\mathfrak{M} = \varinjlim D$ where $D : I \to \mathrm{Sub}_{\aleph_0}(\mathfrak{M})$ is the diagram of the finitely generated substructures of $\mathfrak{M}$. By assumption we have $D(i) \in \mathcal{K}$, for every $i \in I$. Since $\mathcal{K}$ is algebraic it is closed under direct limits of embeddings. Consequently, we have $\mathfrak{M} = \varinjlim D \in \mathcal{K}$. □

As a corollary it follows that every algebraic class is $\forall_{\infty\aleph_0}$-axiomatisable.

Proposition 1.6. *Let Σ be a signature and set $\kappa := |\Sigma| \oplus \aleph_0$. Every algebraic class $\mathcal{K}$ of Σ-structures is $\forall_{(2^\kappa)^+\aleph_0}$-axiomatisable.*

Proof. Let

$$\mathcal{C}_n := \{ \langle \mathfrak{A}, \bar{a} \rangle \mid \mathfrak{A} \in \mathcal{K} \text{ is generated by } \bar{a} \in A^n \}$$

be the class of all structures in $\mathcal{K}$ that are generated by a set of size n. Note that every structure in $\mathcal{C}_n$ has size at most $\kappa = |\Sigma| \oplus \aleph_0$. Consequently, $\mathcal{C}_n$ contains, up to isomorphism, at most 2^κ structures. For every $\langle \mathfrak{A}, \bar{a} \rangle \in \mathcal{C}_n$, we can write down a quantifier-free formula $\varphi_{\mathfrak{A},\bar{a}}(\bar{x}) \in \mathrm{QF}^n_{\kappa^+\aleph_0}[\Sigma]$ such that

$$\mathfrak{B} \vDash \varphi_{\mathfrak{A},\bar{a}}(\bar{b}) \quad \text{iff} \quad \langle \langle\!\langle \bar{b} \rangle\!\rangle_\mathfrak{B}, \bar{b} \rangle \cong \langle \mathfrak{A}, \bar{a} \rangle .$$

By Lemma 1.5, it follows that the $\forall^{\circ}_{(2^\kappa)^+\aleph_0}[\Sigma]$-formula

$$\bigwedge_{n<\omega} \forall x_0 \cdots \forall x_{n-1} \bigvee_{\langle\mathfrak{A},\bar{a}\rangle\in\mathcal{C}_n} \varphi_{\mathfrak{A},\bar{a}}(\bar{x})$$

axiomatises $\mathcal{K}$. □

If we can show that every abstract elementary class has a Skolem expansion, it follows that each such class is a projective $\forall_{\infty\aleph_0}$-class.

Theorem 1.7. *Let $\mathcal{K}$ be an abstract elementary class of Σ-structures. There exists a Skolem expansion $\mathcal{K}_+$ of $\mathcal{K}$ over a signature $\Sigma_+ \supseteq \Sigma$ of size $|\Sigma_+| =$* $\ln(\mathcal{K})$.

Proof. Let $\lambda := \ln(\mathcal{K})$ and set $\Sigma_+ := \Sigma \cup \{ f^n_\alpha \mid n < \omega,\ \alpha < \lambda \}$ where the f^n_α are new n-ary function symbols. We call a Σ_+-expansion $\mathfrak{M}_+$ of a structure $\mathfrak{M} \in \mathcal{K}$ *admissible* if

$$\mathfrak{A}|_\Sigma \leq_{\mathcal{K}} \mathfrak{M}, \quad \text{for every } \mathfrak{A} \subseteq \mathfrak{M}_+ .$$

We claim that the desired Skolem expansion $\langle \mathcal{K}_+, \mathcal{E}_+ \rangle$ is given by

$$\mathcal{K}_+ := \{ \mathfrak{M}_+ \mid \mathfrak{M}_+ \text{ an admissible expansion of some } \mathfrak{M} \in \mathcal{K} \},$$
$$\mathcal{E}_+ := \mathrm{Emb}(\mathcal{K}_+) .$$

Clearly, we have $\ln(\mathcal{K}_+) = |\Sigma_+| = \ln(\mathcal{K})$. Hence, it remains to prove the following claims.

Claim. (a) *For every pair $\mathfrak{A} \leq_{\mathcal{K}} \mathfrak{B}$ in $\mathcal{K}$, there exist admissible expansions $\mathfrak{A}_+$ and $\mathfrak{B}_+$ such that $\mathfrak{A}_+ \subseteq \mathfrak{B}_+$. In particular, we have* $\mathrm{pr}_\Sigma(\mathcal{K}_+) = \mathcal{K}$.

(b) $\mathrm{pr}_\Sigma(\mathcal{E}_+) = \mathcal{E}$.

(c) *$\mathcal{K}_+$ is closed under direct limits.*

(a) By induction on $n < \omega$, we can fix, for every subset $X \subseteq B$ of size n, a $\mathcal{K}$-substructure $\mathfrak{B}_X \leq_{\mathcal{K}} \mathfrak{B}$ of size at most λ containing $X \cup \bigcup_{Y \subset X} B_Y$.

Furthermore, if $X \subseteq A$ then we choose $\mathfrak{B}_X$ such that $B_X \subseteq A$. By construction, we have $\mathfrak{B}_X \subseteq \mathfrak{B}_Y$, for $X \subseteq Y$. Since $\mathfrak{B}_X, \mathfrak{B}_Y \leq_{\mathcal{K}} \mathfrak{B}$ this implies that $\mathfrak{B}_X \leq_{\mathcal{K}} \mathfrak{B}_Y$.

For every $\bar{a} \in B^n$, $n < \omega$, fix an enumeration $(c^{\bar{a}}_\alpha)_{\alpha<\lambda}$ (possibly with repetitions) of $B_{\bar{a}}$. To obtain the desired expansion $\mathfrak{B}_+$ we set $f^n_\alpha(\bar{a}) := c^{\bar{a}}_\alpha$, for $\bar{a} \in B^n$. Note that our construction ensures that A induces a substructure of $\mathfrak{B}_+$ since $\mathfrak{B}_X \subseteq \mathfrak{A}$, for $X \subseteq A$, implies that $\langle\!\langle X \rangle\!\rangle_{\mathfrak{B}_+} \subseteq A$. Therefore, we can set $\mathfrak{A}_+ := \mathfrak{B}_+|_A$.

To see that $\mathfrak{A}_+$ and $\mathfrak{B}_+$ are admissible, note that, by construction, we have $\mathfrak{B}_X \subseteq \langle\!\langle X \rangle\!\rangle_{\mathfrak{B}_+}|_\Sigma$, for every finite $X \subseteq B$. If $\mathfrak{C} \subseteq \mathfrak{B}_+$ is an arbitrary substructure then

$$\mathfrak{C}|_\Sigma = \varinjlim_{X \subseteq C \text{ finite}} \langle\!\langle X \rangle\!\rangle_{\mathfrak{C}}|_\Sigma = \varinjlim_{X \subseteq C \text{ finite}} \langle\!\langle X \rangle\!\rangle_{\mathfrak{B}_+}|_\Sigma = \varinjlim_{X \subseteq C \text{ finite}} \mathfrak{B}_X \,.$$

We have already seen that the $\mathfrak{B}_X$ form a directed system of $\mathcal{K}$-embeddings such that $\mathfrak{B}_X \leq_{\mathcal{K}} \mathfrak{B}$. Hence, the limit also satisfies $\mathfrak{C}|_\Sigma \leq_{\mathcal{K}} \mathfrak{B}$, as desired. Furthermore, if $\mathfrak{C} \subseteq \mathfrak{A}_+ \subseteq \mathfrak{B}_+$ then $\mathfrak{C}|_\Sigma, \mathfrak{A} \leq_{\mathcal{K}} \mathfrak{B}$ implies that $\mathfrak{C}|_\Sigma \leq_{\mathcal{K}} \mathfrak{A}$. Thus, $\mathfrak{A}_+$ and $\mathfrak{B}_+$ are admissible.

(b) ($\subseteq$) Let $h : \mathfrak{A}_+ \to \mathfrak{B}_+$ be a $\mathcal{K}_+$-embedding and set $C := \operatorname{rng} h$. Then C induces a substructure $\mathfrak{C}_+ \subseteq \mathfrak{B}_+$ and h induces an isomorphism $h' : \mathfrak{A}_+ \cong \mathfrak{C}_+$. The structure $\mathfrak{B}_+$ is an admissible expansion of some structure $\mathfrak{B} \in \mathcal{K}$. Hence, $\mathfrak{C}_+|_\Sigma \leq_{\mathcal{K}} \mathfrak{B}$ and the inclusion map $i : \mathfrak{C}_+|_\Sigma \to \mathfrak{B}$ belongs to $\mathcal{E}$. Since $\mathcal{E}$ contains all isomorphisms and it is closed under composition, it follows that $\operatorname{pr}_\Sigma(h) = i \circ \operatorname{pr}_\Sigma(h') \in \mathcal{E}$.

($\supseteq$) Let $h : \mathfrak{C} \to \mathfrak{B}$ be a $\mathcal{K}$-embedding. Setting $\mathfrak{A} := \operatorname{rng} h$ we can use (a) to find admissible expansions $\mathfrak{A}_+ \subseteq \mathfrak{B}_+$ of $\mathfrak{A}$ and $\mathfrak{B}$. Let $\mathfrak{C}_+$ be the expansion of $\mathfrak{C}$ that corresponds to $\mathfrak{A}_+$ via the isomorphism $h : \mathfrak{C} \cong \mathfrak{A}$. Then h induces an embedding $h_+ : \mathfrak{C}_+ \to \mathfrak{B}_+$. Since $\mathcal{K}_+$ is closed under isomorphisms we have $\mathfrak{C}_+ \in \mathcal{K}_+$. Hence, $h_+ \in \mathcal{E}_+$.

(c) Let $D : I \to \mathcal{K}_+$ be a directed diagram with limit $\mathfrak{M}_+ := \varinjlim D$. We have to show that $\mathfrak{M}_+ \in \mathcal{K}_+$. Let $p : \mathcal{K}_+ \to \mathcal{K}$ be the canonical projection functor and set $\mathfrak{M} := \mathfrak{M}_+|_\Sigma$. Then $p \circ D : I \to \mathcal{K}$ is a directed diagram with limit $\varinjlim (p \circ D) = \mathfrak{M}_+|_\Sigma = \mathfrak{M}$. By (b), it follows that $p \circ D$ is in fact a diagram $I \to \mathcal{E}$. Hence, the limit $\mathfrak{M}$ is in $\mathcal{K}$. We claim that $\mathfrak{M}_+$ is an admissible expansion of $\mathfrak{M}$. Let $\mathfrak{A} \subseteq \mathfrak{M}_+$ be a substructure. For every finite set $X \subseteq M$, there exists some i with $X \subseteq D(i)$. Since $D(i)$ is an admissible expansion it follows that

$$\langle\!\langle X \rangle\!\rangle_{D(i)}|_\Sigma \leq_\mathcal{K} D(i)|_\Sigma \leq_\mathcal{K} \varinjlim (p \circ D) = \mathfrak{M} .$$

The substructure $\mathfrak{A}$ is the direct limit of its finitely generated substructures $\mathfrak{X}$. We have just seen that $\mathfrak{X}|_\Sigma \leq_\mathcal{K} \mathfrak{M}$, for all such $\mathfrak{X}$. By the definition of a direct limit, it follows that $\mathfrak{A}|_\Sigma = \varinjlim \mathfrak{X}|_\Sigma \leq_\mathcal{K} \mathfrak{M}$. □

The existence of Skolem expansions enables us to apply the theory of Ehrenfeucht-Mostowski functors to abstract elementary classes. We will make extensive use of these functors in Section 4 below. As an example we use them in the remainder of this section to compute the Hanf number of a class.

Lemma 1.8. *Let $\mathcal{K}$ be an algebraic class of Σ-structures and set $\kappa := |\Sigma| \oplus \aleph_0$ and $\lambda := \beth_{(2^\kappa)^+}$. If $\mathcal{K}$ contains a structure of size at least λ then there exists an Ehrenfeucht-Mostowski functor $F : \mathfrak{Lin} \to \mathfrak{Emb}(\mathcal{K})$.*

Proof. Fix a structure $\mathfrak{M} \in \mathcal{K}$ of size $|M| \geq \lambda$ and let $(a_i)_{i<\lambda}$ be a sequence of distinct elements of M. Since $|S^{<\omega}(\varnothing)| \leq 2^\kappa$ we can apply Theorem E5.3.7 to $(a_i)_i$ to obtain an elementary extension $\mathfrak{M}_+ \succeq_{\mathrm{FO}} \mathfrak{M}$ that contains an indiscernible sequence $(b_i)_{i<\omega}$ such that, for all $n < \omega$ and every $\bar{\imath} \in [\omega]^n$, there is some $\bar{k} \in [\lambda]^n$ with

$$\mathrm{tp}(b[\bar{\imath}]) = \mathrm{tp}(a[\bar{k}]) .$$

Note that this implies in particular that $\langle\!\langle b[\bar{\imath}] \rangle\!\rangle_{\mathfrak{M}_+} \cong \langle\!\langle a[\bar{k}] \rangle\!\rangle_{\mathfrak{M}} \in \mathcal{K}$. By Proposition E6.3.8, there exists a unique strongly local functor $F : \mathfrak{Lin} \to \mathfrak{Emb}(\Sigma)$ such that $F(\omega) \cong \langle\!\langle (b_i)_i \rangle\!\rangle_{\mathfrak{M}_+}$. We claim that the range of F is contained in $\mathcal{K}$.

Let I be a linear order and consider a finitely generated substructure $\mathfrak{A} \subseteq F(I)$. Then there is a finite subset $I_0 \subseteq I$ such that $\mathfrak{A} \subseteq F(I_0)$. Consequently, for some $n < \omega$, $\mathfrak{A}$ is isomorphic to a substructure of

$$F(n) \cong \langle\!\langle b_0 \dots b_{n-1} \rangle\!\rangle_{\mathfrak{M}_+} \subseteq \mathfrak{M}_+ \in \mathcal{K} .$$

Since $\mathcal{K}$ is closed under substructures and isomorphisms, it follows that $\mathfrak{A} \in \mathcal{K}$. Hence, we have $\mathrm{Sub}_{\aleph_0}(F(I)) \subseteq \mathcal{K}$ which, by Lemma 1.5, implies that $F(I) \in \mathcal{K}$. Thus, $F : \mathfrak{Lin} \to \mathfrak{Emb}(\mathcal{K})$ is the desired Ehrenfeucht-Mostowski functor. □

Using Skolem expansions we can extend this result to arbitrary abstract elementary classes.

Remark. Let $\langle \mathcal{K}, \mathcal{E} \rangle$ be an abstract elementary class, $\mathcal{K}_+$ a Skolem expansion of $\mathcal{K}$, and $F_+ : \mathfrak{Lin} \to \mathfrak{Emb}(\mathcal{K}_+)$ an Ehrenfeucht-Mostowski functor. Composing F_+ with the reduct functor $\mathrm{pr}_\Sigma : \mathfrak{Emb}(\Sigma_+) \to \mathfrak{Emb}(\Sigma)$ we obtain a functor $F := \mathrm{pr}_\Sigma \circ F_+ : \mathfrak{Lin} \to \mathfrak{Emb}(\Sigma)$. By definition of a Skolem expansion, F is actually a functor $\mathfrak{Lin} \to \mathcal{E}$, i.e., it maps every embedding $I \to J$ of linear orders to a $\mathcal{K}$-embedding $F(I) \to F(J)$.

Definition 1.9. Let $\mathcal{K}$ be an abstract elementary class of Σ-structures and $\mathcal{K}_+$ a Skolem expansion of $\mathcal{K}$. An *Ehrenfeucht-Mostowski functor for $\mathcal{K}$* is a functor $F : \mathfrak{Lin} \to \mathfrak{Emb}(\mathcal{K})$ of the form $F = \mathrm{pr}_\Sigma \circ F_+$, where $F_+ : \mathfrak{Lin} \to \mathfrak{Emb}(\mathcal{K}_+)$ is an ordinary Ehrenfeucht-Mostowski functor.

Corollary 1.10. *Let $\mathcal{K}$ be an abstract elementary class and set $\kappa := 2^{\mathrm{ln}(\mathcal{K})}$. If $\mathcal{K}$ contains a structure of size at least $\beth_{\kappa^+}$, then there exists an Ehrenfeucht-Mostowski functor for $\mathcal{K}$.*

As promised we apply these results to compute the Hanf number of an abstract elementary class.

Definition 1.11. Let $\mathcal{K}$ be an arbitrary class of Σ-structures. The *Hanf number* of $\mathcal{K}$ is

$$\mathrm{hn}(\mathcal{K}) := \sup \{ |M|^+ \mid \mathfrak{M} \in \mathcal{K} \} .$$

If this supremum does not exist then we set $\mathrm{hn}(\mathcal{K}) := \infty$. In this case the class $\mathcal{K}$ is called *unbounded.*

Proposition 1.12. *Let $\mathcal{K}$ be an abstract elementary class of Σ-structures and set $\kappa := 2^{\mathrm{ln}(\mathcal{K})}$. We either have*

$$\mathrm{hn}(\mathcal{K}) \leq \beth_{\kappa^+} \quad \textit{or} \quad \mathrm{hn}(\mathcal{K}) = \infty .$$

Proof. Suppose that $\text{hn}(\mathcal{K}) > \beth_{\kappa^+}$. By Corollary 1.10, there exists an Ehrenfeucht-Mostowski functor $F : \mathfrak{Lin} \to \mathfrak{Emb}(\mathcal{K})$ for $\mathcal{K}$. For every cardinal λ, we have $F(\lambda) \in \mathcal{K}$. This implies that

$$\text{hn}(\mathcal{K}) > |F(\lambda)| = \lambda \oplus \text{ln}(\mathcal{K}) .$$

Consequently, $\text{hn}(\mathcal{K}) = \infty$. □

With this proposition we are finally able to provide the missing part of the proof of Theorem C5.2.7. (Except that we do not obtain a strict inequality $\text{hn}_1(\text{FO}_{\kappa^+\aleph_0}) < \beth_{(2^\kappa)^+}$.)

Corollary 1.13. $\text{hn}_1(\text{FO}_{\kappa^+\aleph_0}) \leq \beth_{(2^\kappa)^+}$.

2. *Amalgamation and saturation*

In this section we consider saturated structures in abstract elementary classes. As we have already seen in the first-order case, an important ingredient in the construction of such structures is the amalgamation property.

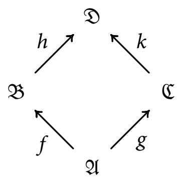
Definition 2.1. Let $\langle \mathcal{K}, \mathcal{E} \rangle$ be an abstract elementary class.

(a) For a cardinal κ, we set

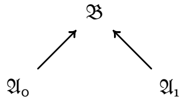
$$\mathcal{K}_\kappa := \{ \mathfrak{M} \in \mathcal{K} \mid |M| = \kappa \} \quad \text{and} \quad \mathcal{K}_{<\kappa} := \{ \mathfrak{M} \in \mathcal{K} \mid |M| < \kappa \} .$$

We define $\mathcal{K}_{>\kappa}$, $\mathcal{K}_{\leq\kappa}$, and $\mathcal{K}_{\geq\kappa}$ analogously.

(b) $\mathcal{K}$ has the *amalgamation property* if, for all $\mathcal{K}$-embeddings $f : \mathfrak{A} \to \mathfrak{B}$ and $g : \mathfrak{A} \to \mathfrak{C}$, there exist $\mathcal{K}$-embeddings $h : \mathfrak{B} \to \mathfrak{D}$ and $k : \mathfrak{C} \to \mathfrak{D}$ with $h \circ f = k \circ g$.



(c) $\mathcal{K}$ has the *joint embedding property* if, for all $\mathfrak{A}_0, \mathfrak{A}_1 \in \mathcal{K}$, there are $\mathcal{K}$-embeddings $\mathfrak{A}_0 \to \mathfrak{B}$ and $\mathfrak{A}_1 \to \mathfrak{B}$, for some $\mathfrak{B} \in \mathcal{K}$.



(d) An *amalgamation class* is an abstract elementary class with the amalgamation property. A *Jónsson class* is an abstract elementary class with the amalgamation property and the joint embedding property.

Example. Let T be an $\forall\exists$-theory and $\mathcal{K}$ the class of all existentially closed models of T. Then $\langle \mathcal{K}, \mathrm{Emb}(\mathcal{K}) \rangle$ forms an abstract elementary class with the amalgamation property.

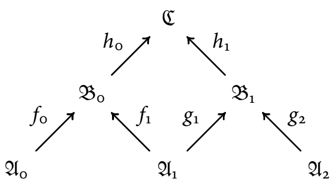
In the same way that the class of all algebraically closed fields can be decomposed into the classes of algebraically closed fields of characteristic p, for the various p, we can write each amalgamation class as a union of Jónsson classes.

Lemma 2.2. *Every amalgamation class $\mathcal{K}$ is a disjoint union of at most $2^{\ln(\mathcal{K})}$ Jónsson classes.*

Proof. We define an equivalence relation on $\mathcal{K}$ by

$$\mathfrak{A} \sim \mathfrak{B} \quad :\text{iff} \quad \text{there are } \mathcal{K}\text{-embeddings } \mathfrak{A} \to \mathfrak{C} \text{ and } \mathfrak{B} \to \mathfrak{C}, \text{ for some } \mathfrak{C} \in \mathcal{K}.$$

Clearly, $\sim$ is reflexive and symmetric. For transitivity, let us assume that $\mathfrak{A}_0 \sim \mathfrak{A}_1 \sim \mathfrak{A}_2$. Then there are structures $\mathfrak{B}_0, \mathfrak{B}_1 \in \mathcal{K}$ and $\mathcal{K}$-embeddings $f_i : \mathfrak{A}_i \to \mathfrak{B}_0$, for $i \in \{0,1\}$, and $g_i : \mathfrak{A}_i \to \mathfrak{B}_1$, for $i \in \{1,2\}$.



By the amalgamation property, we can find some structure $\mathfrak{C} \in \mathcal{K}$ and $\mathcal{K}$-embeddings $h_i : \mathfrak{B}_i \to \mathfrak{C}$, for $i < 2$, such that $h_0 \circ f_1 = h_1 \circ g_1$. Consequently, there are $\mathcal{K}$-embeddings $h_0 \circ f_0 : \mathfrak{A}_0 \to \mathfrak{C}$ and $h_1 \circ g_2 : \mathfrak{A}_2 \to \mathfrak{C}$. This implies that $\mathfrak{A}_0 \sim \mathfrak{A}_2$.

By definition, every $\sim$-class is a Jónsson class. Furthermore, $\mathfrak{A} \nsim \mathfrak{B}$ implies that there is no $\mathcal{K}$-embedding $\mathfrak{A} \to \mathfrak{B}$. Hence, $\mathcal{K}$ is the disjoint union of all $\sim$-classes. Finally, every $\sim$-class contains a structure of size at most $\ln(\mathcal{K})$. Consequently, there are at most $2^{\ln(\mathcal{K})}$ such classes. □

For amalgamation classes, the definition of a Galois type can be simplified quite a bit.

Lemma 2.3. *Let $\mathcal{K}$ be an amalgamation class, $\mathfrak{A}, \mathfrak{B}, \mathfrak{U} \in \mathcal{K}$ structures with $\mathfrak{U} \leq_{\mathcal{K}} \mathfrak{A}, \mathfrak{B}$, and let $\bar{a} \subseteq A$ and $\bar{b} \subseteq B$. Then we have*

$$\operatorname{tp}_{\mathrm{Aut}}(\bar{a}/\mathfrak{A}, U) = \operatorname{tp}_{\mathrm{Aut}}(\bar{b}/\mathfrak{B}, U)$$

if and only if there exists a structure $\mathfrak{M} \in \mathcal{K}$ of size $|M| \leq |A| \oplus |B| \oplus \ln(\mathcal{K})$ and $\mathcal{K}$-embeddings $g : \mathfrak{A} \to \mathfrak{M}$ and $h : \mathfrak{B} \to \mathfrak{M}$ such that

$$g \restriction U = h \restriction U \quad \text{and} \quad g(\bar{a}) = h(\bar{b}) .$$

Proof. ($\Leftarrow$) is trivial. For ($\Rightarrow$), suppose that the Galois types are equal. Recall the relation $\sim$ from Definition 1.2. There exists a finite sequence $\langle \mathfrak{C}_0, \bar{c}_0 \rangle, \dots, \langle \mathfrak{C}_n, \bar{c}_n \rangle$ of structures such that

$$\langle \mathfrak{C}_0, \bar{c}_0 \rangle = \langle \mathfrak{A}, \bar{a} \rangle , \quad \langle \mathfrak{C}_n, \bar{c}_n \rangle = \langle \mathfrak{B}, \bar{b} \rangle ,$$

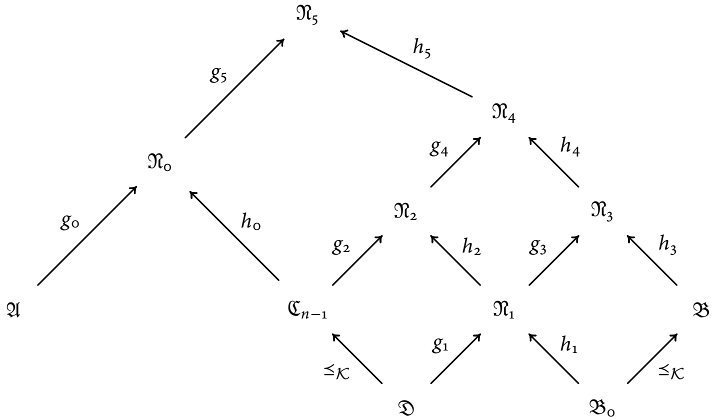
$$\text{and} \quad \langle \bar{c}_i, \mathfrak{C}_i, U \rangle \sim \langle \bar{c}_{i+1}, \mathfrak{C}_{i+1}, U \rangle , \quad \text{for all } i < n .$$

We prove the claim by induction on n. For $n = 0$, we have $\mathfrak{A} = \mathfrak{B}$ and $\bar{a} = \bar{b}$, and there is nothing to do. Hence, suppose that $n > 0$. By inductive hypothesis, there exist a structure $\mathfrak{N}_0 \in \mathcal{K}$ and $\mathcal{K}$-embeddings $g_0 : \mathfrak{A} \to \mathfrak{N}_0$ and $h_0 : \mathfrak{C}_{n-1} \to \mathfrak{N}_0$ such that

$$g_0 \restriction U = h_0 \restriction U \quad \text{and} \quad g_0(\bar{a}) = h_0(\bar{c}_{n-1}) .$$

Furthermore, by definition of $\sim$, we can find a structure $\mathfrak{N}_1 \in \mathcal{K}$, $\mathcal{K}$-substructures $\mathfrak{D} \leq_{\mathcal{K}} \mathfrak{C}_{n-1}$ and $\mathfrak{B}_0 \leq_{\mathcal{K}} \mathfrak{B}$ with $U \cup \bar{c}_{n-1} \subseteq D$ and $U \cup \bar{b} \subseteq B_0$, and $\mathcal{K}$-embeddings $g_1 : \mathfrak{D} \to \mathfrak{N}_1$ and $h_1 : \mathfrak{B}_0 \to \mathfrak{N}_1$ such that

$$g_1 \upharpoonright U = h_1 \upharpoonright U \quad \text{and} \quad g_1(\bar{c}_{n-1}) = h_1(\bar{b}) .$$



By the amalgamation property, there exist structures $\mathfrak{N}_2, \mathfrak{N}_3, \mathfrak{N}_4, \mathfrak{N}_5 \in \mathcal{K}$ such that we can complete the above diagram. Setting $g := g_5 \circ g_0$ and $h := h_5 \circ h_4 \circ h_3$ it follows that

$$g \upharpoonright U = h \upharpoonright U \quad \text{and} \quad g(\bar{a}) = h(\bar{b}) .$$

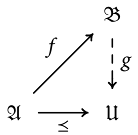
Choosing a $\mathcal{K}$-substructure $\mathfrak{M} \leq_{\mathcal{K}} \mathfrak{N}_5$ of size $|M| \leq |A| \oplus |B| \oplus \ln(\mathcal{K})$ with $\operatorname{rng} g \cup \operatorname{rng} h \subseteq M$ the claim follows. □

Next, we introduce a notion of saturation for abstract elementary classes.

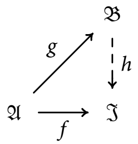
Definition 2.4. Let $\mathcal{K}$ be an abstract elementary class and let $\kappa \geq \ln(\mathcal{K})$ be a cardinal.

(a) A structure $\mathfrak{U} \in \mathcal{K}$ is *κ-universal* (for $\mathcal{K}$) if, for all $\mathfrak{A} \in \mathcal{K}_{<\kappa}$, there exists a $\mathcal{K}$-embedding $\mathfrak{A} \to \mathfrak{U}$. We call $\mathfrak{U}$ *$\mathcal{K}$-universal* if it is $|U|^+$-universal for $\mathcal{K}$.

(b) Similarly, we say that a structure $\mathfrak{U} \in \mathcal{K}$ is κ-universal *over a substructure* $\mathfrak{A} \leq_{\mathcal{K}} \mathfrak{U}$ if, for all $\mathcal{K}$-embeddings $f : \mathfrak{A} \to \mathfrak{B}$ with $|B| < \kappa$, there exists a $\mathcal{K}$-embedding $g : \mathfrak{B} \to \mathfrak{U}$ such that $g \circ f = \mathrm{id}_A$.



(c) A structure $\mathfrak{J} \in \mathcal{K}$ is *κ-injective* (for $\mathcal{K}$), or *κ-model homogeneous,* if, for all $\mathcal{K}$-embeddings $f : \mathfrak{A} \to \mathfrak{J}$ and $g : \mathfrak{A} \to \mathfrak{B}$ with $|A|, |B| < \kappa$, there exists a $\mathcal{K}$-embedding $h : \mathfrak{B} \to \mathfrak{J}$ with $h \circ g = f$.



$\mathfrak{J}$ is called *$\mathcal{K}$-injective* if it is $|I|$-injective.

Remark. Note that a structure $\mathfrak{M}$ is κ-injective if and only if it is κ-universal over every substructure $\mathfrak{A} \leq_{\mathcal{K}} \mathfrak{M}$ of size $|A| < \kappa$.

We can characterise κ-injective structures also by a back-and-forth condition.

Definition 2.5. Let $\mathcal{K}$ be an abstract elementary class and $\mathfrak{A}, \mathfrak{B} \in \mathcal{K}$.

(a) We denote by $I^\kappa_{\mathcal{K}}(\mathfrak{A}, \mathfrak{B})$ the set of all $\mathcal{K}$-embeddings $f : \mathfrak{A}_0 \to \mathfrak{B}_0$ between $\mathcal{K}$-substructures $\mathfrak{A}_0 \leq_{\mathcal{K}} \mathfrak{A}$ and $\mathfrak{B}_0 \leq_{\mathcal{K}} \mathfrak{B}$ of size $|A_0|, |B_0| < \kappa$.

(b) We write

$$\begin{aligned} &\mathfrak{A} \sqsubseteq^\kappa_{\mathcal{K}} \mathfrak{B} \quad &&:\text{iff} \quad I^\kappa_{\mathcal{K}}(\mathfrak{A}, \mathfrak{B}) : \mathfrak{A} \sqsubseteq^\kappa_{\mathrm{iso}} \mathfrak{B}\,, \\ \text{and} \quad &\mathfrak{A} \cong^\kappa_{\mathcal{K}} \mathfrak{B} \quad &&:\text{iff} \quad I^\kappa_{\mathcal{K}}(\mathfrak{A}, \mathfrak{B}) : \mathfrak{A} \cong^\kappa_{\mathrm{iso}} \mathfrak{B}\,. \end{aligned}$$

In Lemma E1.2.2 we have characterised κ-saturated models in terms of the relation $\sqsubseteq^{\kappa}_{\mathrm{FO}}$. The next lemma gives a similar characterisation of κ-injective structures.

Lemma 2.6. *Let $\mathcal{K}$ be an abstract elementary class and $\kappa > \ln(\mathcal{K})$ a cardinal. A structure $\mathfrak{M} \in \mathcal{K}$ is κ-injective if and only if*

$$\mathfrak{A} \sqsubseteq^{\kappa}_{\mathcal{K}} \mathfrak{M}\,, \quad \text{for all } \mathfrak{A} \in \mathcal{K} \text{ with } I^{\kappa}_{\mathcal{K}}(\mathfrak{A}, \mathfrak{M}) \neq \varnothing\,.$$

Proof. ($\Leftarrow$) Suppose that $\mathfrak{A}, \mathfrak{B} \in \mathcal{K}_{<\kappa}$ are structures with $\mathfrak{A} \leq_{\mathcal{K}} \mathfrak{B}$, and let $f : \mathfrak{A} \to \mathfrak{M}$ be a $\mathcal{K}$-embedding. Then $f \in I^{\kappa}_{\mathcal{K}}(\mathfrak{B}, \mathfrak{M})$. Since $|B| < \kappa$, we can use Lemma C4.4.9 (b) to find a $\mathcal{K}$-embedding $g \in I^{\kappa}_{\mathcal{K}}(\mathfrak{B}, \mathfrak{M})$ with $\operatorname{dom} g = B$ and $g \upharpoonright A = f$.

($\Rightarrow$) By assumption, $I^{\kappa}_{\mathcal{K}}(\mathfrak{A}, \mathfrak{M})$ is nonempty. It has the forth property since $\mathfrak{M}$ is κ-injective. Furthermore, $I^{\kappa}_{\mathcal{K}}(\mathfrak{M}, \mathfrak{A})$ is $\ln(\mathcal{K})^+$-bounded. Finally, the closure of $\mathcal{K}$-embeddings under direct limits implies that $I^{\kappa}_{\mathcal{K}}(\mathfrak{A}, \mathfrak{M})$ is κ-complete. ☐

As usual we can use Lemma C4.4.9 to prove that, up to isomorphism, $\mathcal{K}$-injective structures are uniquely determined by their cardinality.

Proposition 2.7. *Let $\mathfrak{A}, \mathfrak{B} \in \mathcal{K}$ be two $\mathcal{K}$-injective structures with $|A| = |B|$. Then*

$$I_{\mathcal{K}}(\mathfrak{A}, \mathfrak{B}) \neq \varnothing \quad \text{implies} \quad \mathfrak{A} \cong \mathfrak{B}\,.$$

The existence of κ-injective structures implies a weak form of the amalgamation property.

Lemma 2.8. *Let $\mathcal{K}$ be an abstract elementary class and suppose that $\mathfrak{M} \in \mathcal{K}$ is κ-injective, for some $\kappa > \ln(\mathcal{K})$.*

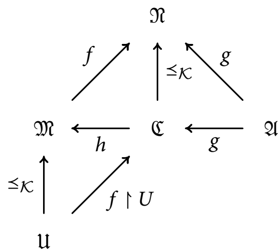
(a) *The class of all $\mathcal{K}$-substructures $\mathfrak{A} \leq_{\mathcal{K}} \mathfrak{M}$ with $|A| < \kappa$ has the amalgamation property.*

(b) *If $\mathcal{K}$ has the joint embedding property, then $\mathfrak{M}$ is κ^+-universal.*

(c) *If $\mathcal{K}$ has the joint embedding property, then the subclass $\mathcal{K}_{<\kappa}$ has the amalgamation property.*

Proof. (a) Let $f : \mathfrak{A} \to \mathfrak{B}$ and $g : \mathfrak{A} \to \mathfrak{C}$ be $\mathcal{K}$-embeddings with $\mathfrak{A}, \mathfrak{B}, \mathfrak{C} \leq_{\mathcal{K}} \mathfrak{M}$ and $|A|, |B|, |C| < \kappa$. Replacing $\mathfrak{A}$ by an isomorphic copy, we may assume that $g = \mathrm{id}_A$. Since $\mathfrak{M}$ is κ-injective, there exists a $\mathcal{K}$-embedding $h : \mathfrak{B} \to \mathfrak{M}$ with $h \circ f = \mathrm{id}_A$. Let $\mathfrak{D} \leq_{\mathcal{K}} \mathfrak{M}$ be a substructure containing $C \cup \operatorname{rng} h$. Then we can use $h : \mathfrak{B} \to \mathfrak{D}$ and $\mathrm{id}_C : \mathfrak{C} \to \mathfrak{D}$ to complete the amalgamation diagram.

(b) As a first step, we show that $\mathfrak{M}$ is κ-universal. Let $\mathfrak{A}$ be some structure of size $|A| < \kappa$. We can use the joint embedding property to find $\mathcal{K}$-embeddings $f : \mathfrak{M} \to \mathfrak{N}$ and $g : \mathfrak{A} \to \mathfrak{N}$, for some $\mathfrak{N} \in \mathcal{K}$.



Choose a $\mathcal{K}$-substructure $\mathfrak{U} \leq_{\mathcal{K}} \mathfrak{M}$ of size $|U| < \kappa$ and let $\mathfrak{C} \leq_{\mathcal{K}} \mathfrak{N}$ be a $\mathcal{K}$-substructure of size $|C| < \kappa$ with $f[U] \cup g[A] \subseteq C$. Since $\mathfrak{M}$ is κ-injective, there exists a $\mathcal{K}$-embedding $h : \mathfrak{C} \to \mathfrak{M}$ with $h \circ f \restriction U = \mathrm{id}_U$. The composition $h \circ g$ is a $\mathcal{K}$-embedding $\mathfrak{A} \to \mathfrak{M}$.

It remains to show that $\mathfrak{M}$ is even κ^+-universal. Let $\mathfrak{A}$ be a structure of size $|A| = \kappa$. Fix an increasing chain $(\mathfrak{C}_\alpha)_{\alpha<\kappa}$ of $\mathcal{K}$-substructures $\mathfrak{C}_\alpha \leq_{\mathcal{K}} \mathfrak{A}$ of size $|C_\alpha| < \kappa$ such that $\mathfrak{A} = \bigcup_{\alpha<\kappa} \mathfrak{C}_\alpha$. By induction on α, we construct $\mathcal{K}$-embeddings $f_\alpha : \mathfrak{C}_\alpha \to \mathfrak{M}$ such that $f_\beta \restriction C_\alpha = f_\alpha$, for all $\alpha \leq \beta$. We have already shown that $\mathfrak{M}$ is κ-universal. Hence, there exists a $\mathcal{K}$-embedding $f_0 : \mathfrak{C}_0 \to \mathfrak{M}$ which we can start our induction with. For limit ordinals δ, we set $f_\delta := \bigcup_{\alpha<\delta} f_\alpha$. For the successor step, suppose that we have already defined $f_\alpha : \mathfrak{C}_\alpha \to \mathfrak{M}$. Since $\mathfrak{M}$ is κ-injective, there exists a $\mathcal{K}$-embedding $f_{\alpha+1} : \mathfrak{C}_{\alpha+1} \to \mathfrak{M}$ such that $f_{\alpha+1} \restriction C_\alpha = f_\alpha$.

Having defined the family $(f_\alpha)_\alpha$ we can use the properties of a direct limit to find a $\mathcal{K}$-embedding $h : \bigcup_\alpha \mathfrak{C}_\alpha \to \mathfrak{M}$ such that $h \restriction C_\alpha = f_\alpha$, for all α. This is the desired $\mathcal{K}$-embedding $\mathfrak{A} \to \mathfrak{M}$.

(c) Let $f : \mathfrak{A} \to \mathfrak{B}$ and $g : \mathfrak{A} \to \mathfrak{C}$ be $\mathcal{K}$-embeddings with $|A|, |B|, |C| < \kappa$. By (b), we may assume that $\mathfrak{A}, \mathfrak{B}, \mathfrak{C} \leq_{\mathcal{K}} \mathfrak{M}$. Hence, we can use (a) to complete f and g to an amalgamation diagram. □

κ-injective structures generalise the characterisation of κ-saturated structures in terms of the relation $\sqsubseteq^{\kappa}_{\mathrm{FO}}$. We can also generalise the original definition of κ-saturation in terms of types. It turns out that, for amalgamation classes, these two notions coincide.

Definition 2.9. Let $\mathcal{K}$ be an abstract elementary class.

(a) A structure $\mathfrak{M} \in \mathcal{K}$ is *κ-Galois saturated* if it realises every Galois type in $S^{<\omega}_{\mathrm{Aut}}(U)$ where $\mathfrak{U} \leq_{\mathcal{K}} \mathfrak{M}$ is a substructure of size $|U| < \kappa$. As usual we say that $\mathfrak{M}$ is *Galois saturated* if it is $|M|$-Galois saturated.

(b) $\mathcal{K}$ is *κ-Galois stable* if $|S^{<\omega}_{\mathrm{Aut}}(U)| \leq \kappa$, for all $\mathfrak{U} \in \mathcal{K}_{\leq\kappa}$.

Remark. Note that in the definition of κ-Galois stability we only count the Galois types over $\mathcal{K}$-substructures, not over arbitrary subsets. In general, this does make a difference.

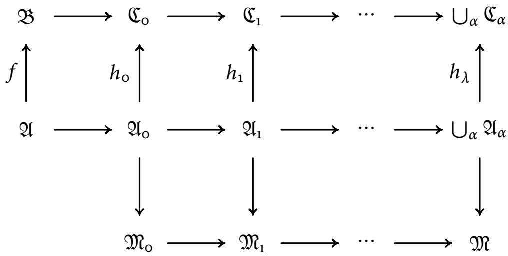
The following lemma is the main ingredient in showing that κ-Galois saturated structures are κ-injective. We state it in a slightly more general form than needed here, since we will use it again in Section 3.

Lemma 2.10. *Let $\mathcal{K}$ be an amalgamation class and $\gamma \geq \ln(\mathcal{K})$ an ordinal. Suppose that $(\mathfrak{M}_\alpha)_{\alpha<\gamma}$ is an increasing chain such that each structure $\mathfrak{M}_{\alpha+1}$ realises every Galois type $\mathfrak{p} \in S^{<\omega}_{\mathrm{Aut}}(U)$ where $\mathfrak{U} \leq \mathfrak{M}_\alpha$ is some substructure of size $|U| \leq |\gamma|$.*

Then the limit $\mathfrak{M} := \bigcup_{\alpha<\gamma} \mathfrak{M}_\alpha$ is $|\gamma|^+$-universal over every substructure $\mathfrak{A} \leq_{\mathcal{K}} \mathfrak{M}_0$ of size $|A| \leq |\gamma|$.

Proof. Let $\mathfrak{A} \leq_{\mathcal{K}} \mathfrak{M}_0$ be of size $|A| \leq |\gamma|$. To show that $\mathfrak{M}$ is $|\gamma|^+$-universal over $\mathfrak{A}$, we consider a $\mathcal{K}$-embedding $f : \mathfrak{A} \to \mathfrak{B}$ with $|B| \leq |\gamma|$. Set $\lambda := |B| \oplus \ln(\mathcal{K})$ and fix an enumeration $(b_\alpha)_{\alpha<\lambda}$ of B. We construct two increasing chains $(\mathfrak{A}_\alpha)_{\alpha<\lambda}$ and $(\mathfrak{C}_\alpha)_{\alpha<\lambda}$ of structures with $\mathfrak{B} \leq_{\mathcal{K}} \mathfrak{C}_\alpha$ and $\mathfrak{A} \leq_{\mathcal{K}} \mathfrak{A}_\alpha \leq_{\mathcal{K}} \mathfrak{M}_\alpha$, and an increasing chain $(h_\alpha)_{\alpha<\lambda}$ of $\mathcal{K}$-embeddings $h_\alpha : \mathfrak{A}_\alpha \to \mathfrak{C}_\alpha$ such that

$$|A_\alpha| \leq \lambda\,, \quad f \subseteq h_\alpha\,, \quad \text{and} \quad b_\alpha \in \operatorname{rng} h_{\alpha+1}\,.$$

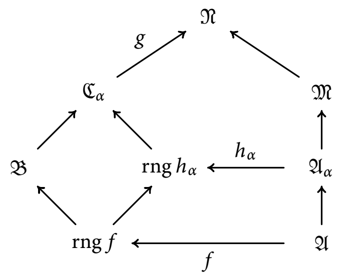


Then we obtain the desired embedding $g : \mathfrak{B} \to \mathfrak{M}$ by taking the limit $h_\lambda := \bigcup_{\alpha<\lambda} h_\alpha$ and setting $g := h_\lambda^{-1} \upharpoonright B$.

We start with $\mathfrak{A}_0 := \mathfrak{A}$, $\mathfrak{C}_0 := \mathfrak{B}$, and $h_0 := f$. For limit ordinals δ, we take limits:

$$\mathfrak{A}_\delta := \bigcup_{\alpha<\delta} \mathfrak{A}_\alpha\,, \quad \mathfrak{C}_\delta := \bigcup_{\alpha<\delta} \mathfrak{C}_\alpha\,, \quad \text{and} \quad h_\delta := \bigcup_{\alpha<\delta} h_\alpha\,.$$

For the successor step, suppose that $h_\alpha : \mathfrak{A}_\alpha \to \mathfrak{C}_\alpha$ has already been defined. If $b_\alpha \in \operatorname{rng} h_\alpha$, we simply set $h_{\alpha+1} := h_\alpha$. Otherwise, we use amalgamation to find a $\mathcal{K}$-extension $\mathfrak{N} \geq_{\mathcal{K}} \mathfrak{M}$ and a $\mathcal{K}$-embedding $g : \mathfrak{C}_\alpha \to \mathfrak{N}$ with $g \circ h_\alpha = \mathrm{id}$.



By assumption on $\mathfrak{M}_{\alpha+1}$, there is some element $c \in M_{\alpha+1}$ with

$$\operatorname{tp}_{\mathrm{Aut}}(c/\mathfrak{N}, A_\alpha) = \operatorname{tp}_{\mathrm{Aut}}(g(b_\alpha)/\mathfrak{N}, A_\alpha)\,.$$

By Lemma 2.3, this implies that there is a $\mathcal{K}$-extension $\mathfrak{N}^+ \geq_{\mathcal{K}} \mathfrak{N}$ and a $\mathcal{K}$-embedding $\sigma : \mathfrak{N} \to \mathfrak{N}^+$ such that

$$\sigma \upharpoonright A_\alpha = \mathrm{id} \quad \text{and} \quad \sigma(g(b_\alpha)) = c\,.$$

We choose a $\mathcal{K}$-substructure $\mathfrak{A}_{\alpha+1} \leq_{\mathcal{K}} \mathfrak{M}_{\alpha+1}$ of size $|A_{\alpha+1}| \leq \lambda$ containing A_α and c. Let $\mathfrak{C}'_{\alpha+1} \leq_{\mathcal{K}} \mathfrak{N}^+$ be a $\mathcal{K}$-substructure containing $\mathrm{rng}(\sigma \circ g)$ and $A_{\alpha+1}$, and let $\mathfrak{C}_{\alpha+1}$ be the isomorphic copy of $\mathfrak{C}'_{\alpha+1}$ where each element of $\mathrm{rng}(\sigma \circ g)$ is replaced by its preimage. We denote the corresponding isomorphism $\mathfrak{C}'_{\alpha+1} \to \mathfrak{C}_{\alpha+1}$ by π. It follows that $\mathfrak{C}_\alpha \leq_{\mathcal{K}} \mathfrak{C}_{\alpha+1}$. We claim that the restriction $h_{\alpha+1} := \pi \upharpoonright A_{\alpha+1}$ is the desired $\mathcal{K}$-embedding $\mathfrak{A}_{\alpha+1} \to \mathfrak{C}_{\alpha+1}$. Note that

$$b_\alpha = \pi((\sigma \circ g)(b_\alpha)) = \pi(c) \in \mathrm{rng}\, h_{\alpha+1}\,.$$

Furthermore, $\sigma \upharpoonright A_\alpha = \mathrm{id}_{A_\alpha} = g \circ h_\alpha \upharpoonright A_\alpha$ implies for $a \in A_\alpha$ that

$$h_{\alpha+1}(a) = \pi(a) = \pi(\sigma(a)) = \pi(\sigma((g \circ h_\alpha)(a))) = h_\alpha(a)\,.$$

Hence, $h_\alpha \subseteq h_{\alpha+1}$. □

Theorem 2.11. *Let $\mathcal{K}$ be an amalgamation class and $\kappa > \mathrm{ln}(\mathcal{K})$. A structure $\mathfrak{M} \in \mathcal{K}$ is κ-Galois saturated if and only if it is κ-injective.*

Proof. ($\Leftarrow$) Let $\mathfrak{U} \leq_{\mathcal{K}} \mathfrak{M}$ be a substructure of size $|U| < \kappa$ and let $\mathfrak{p} \in S^{<\omega}_{\mathrm{Aut}}(U)$ be a type. There exists an extension $\mathfrak{A} \geq_{\mathcal{K}} \mathfrak{U}$ realising $\mathfrak{p}$. We can choose $\mathfrak{A}$ of size $|A| \leq |U| \oplus \mathrm{ln}(\mathcal{K}) < \kappa$. Since $\mathfrak{M}$ is κ-injective, we can extend the $\mathcal{K}$-embedding $\mathfrak{U} \to \mathfrak{M}$ to a $\mathcal{K}$-embedding $\mathfrak{A} \to \mathfrak{M}$. Consequently, $\mathfrak{p}$ is realised in $\mathfrak{M}$.

($\Rightarrow$) Suppose that $f : \mathfrak{A} \to \mathfrak{B}$ is a $\mathcal{K}$-embedding with $\mathfrak{A} \leq_{\mathcal{K}} \mathfrak{M}$ and $\lambda := |B| < \kappa$. For $\alpha < \lambda$, we set $\mathfrak{M}_\alpha := \mathfrak{M}$. Then $(\mathfrak{M}_\alpha)_{\alpha<\lambda}$ is an increasing chain satisfying the hypothesis of Lemma 2.10. It follows that the limit $\bigcup_{\alpha<\lambda} \mathfrak{M}_\alpha = \mathfrak{M}$ is λ^+-universal over $\mathfrak{A}$. Consequently, there exists a $\mathcal{K}$-embedding $g : \mathfrak{B} \to \mathfrak{M}$ with $g \circ f \upharpoonright A = \mathrm{id}$. □

The next lemma shows that Galois saturated structures are strongly homogeneous.

Lemma 2.12. *Let $\mathcal{K}$ be an amalgamation class, suppose that $\mathfrak{M} \in \mathcal{K}$ is a Galois saturated structure of size $|M| = \kappa$, and let $\mathfrak{U} \leq_{\mathcal{K}} \mathfrak{M}$ be a substructure of size* $\ln(\mathcal{K}) \leq |U| < \kappa$. *For $\bar{a}, \bar{b} \in M^{<\kappa}$, we have*

$$\mathrm{tp}_{\mathrm{Aut}}(\bar{a}/\mathfrak{M}, U) = \mathrm{tp}_{\mathrm{Aut}}(\bar{b}/\mathfrak{M}, U)$$

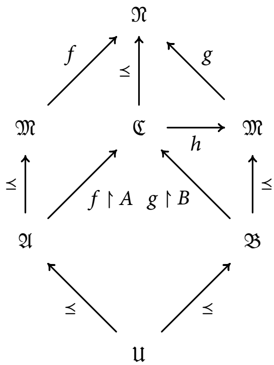
if and only if there exists an automorphism $\pi \in \mathrm{Aut}\,\mathfrak{M}$ with $\pi \upharpoonright U = \mathrm{id}_U$ and $\pi(\bar{a}) = \bar{b}$.

Proof. It is sufficient to find an embedding $p \in \mathcal{I}^{\kappa}_{\mathcal{K}}(\mathfrak{M}, \mathfrak{M})$ with $p \upharpoonright U = \mathrm{id}_U$ and $p(\bar{a}) = \bar{b}$. Since $\mathfrak{M} \cong^{\kappa}_{\mathcal{K}} \mathfrak{M}$ we can then use Lemma C4.4.9 to extend p to the desired isomorphism $\pi : \mathfrak{M} \to \mathfrak{M}$.

Fix $\mathcal{K}$-substructures $\mathfrak{U} \leq_{\mathcal{K}} \mathfrak{A} \leq_{\mathcal{K}} \mathfrak{M}$ and $\mathfrak{U} \leq_{\mathcal{K}} \mathfrak{B} \leq_{\mathcal{K}} \mathfrak{M}$ of size $|A|, |B| < \kappa$ with $\bar{a} \subseteq A$ and $\bar{b} \subseteq B$. Since

$$\mathrm{tp}_{\mathrm{Aut}}(\bar{a}/\mathfrak{M}, U) = \mathrm{tp}_{\mathrm{Aut}}(\bar{b}/\mathfrak{M}, U)\,,$$

we can use Lemma 2.3 to find $\mathcal{K}$-embeddings $f, g : \mathfrak{M} \to \mathfrak{N}$ with $f \upharpoonright U = g \upharpoonright U$ and $f(\bar{a}) = g(\bar{b})$.



Let $\mathfrak{C} \leq_{\mathcal{K}} \mathfrak{M}$ be a $\mathcal{K}$-substructure of size $|C| < \kappa$ with $f[A] \cup g[B] \subseteq C$. Since $\mathfrak{M}$ is κ-injective, there exists a $\mathcal{K}$-embedding $h : \mathfrak{C} \to \mathfrak{M}$ with $h \circ g \upharpoonright B = \mathrm{id}_B$. Setting $p := h \circ f \upharpoonright A$ we have

$$p \upharpoonright U = h \circ f \upharpoonright U = h \circ g \upharpoonright U = \mathrm{id}_U\,,$$

and $\quad p(\bar{a}) = h(f(\bar{a})) = h(g(\bar{b})) = \bar{b}\,.$ $\square$

When amalgamation is available we can construct κ-Galois saturated structures in the same way as κ-saturated ones. The main step in the inductive construction is the following lemma.

Lemma 2.13. *Let $\mathcal{K}$ be an amalgamation class. Every $\mathfrak{M} \in \mathcal{K}$ has an extension $\mathfrak{M}^+ \geq_{\mathcal{K}} \mathfrak{M}$ that realises every Galois type over $\mathfrak{M}$. If $\mathcal{K}$ is κ-stable, for $\kappa := |M| \oplus \ln(\mathcal{K})$, then we can choose $\mathfrak{M}^+$ of size $|M^+| \leq \kappa$.*

Proof. Let $(\mathfrak{p}_i)_{i<\lambda}$ be an enumeration of $S^{<\omega}_{\mathrm{Aut}}(M)$. For every $i < \lambda$, we can find an extension $\mathfrak{A}_i \geq_{\mathcal{K}} \mathfrak{M}$ of size $|A_i| \leq |M| \oplus \ln(\mathcal{K}) = \kappa$ realising $\mathfrak{p}_i$. We construct $\mathfrak{M}^+$ as the limit of an increasing chain $(\mathfrak{B}_i)_{i<\lambda}$ where the structure $\mathfrak{B}_\alpha$ realises all types $\mathfrak{p}_i$ with $i < \alpha$. We start with $\mathfrak{B}_0 := \mathfrak{M}$. For limit ordinals δ, we set $\mathfrak{B}_\delta := \bigcup_{i<\delta} \mathfrak{B}_i$. For successor ordinals $\alpha = \beta + 1$, we use the amalgamation property to find an extension $\mathfrak{B}_\alpha \geq_{\mathcal{K}} \mathfrak{B}_\beta$ of size $|B_\alpha| \leq |B_\beta| \oplus |A_\beta| \oplus \ln(\mathcal{K})$ such that there exists a $\mathcal{K}$-embedding $h : \mathfrak{A}_\beta \to \mathfrak{B}_\alpha$ with $h \upharpoonright M = \mathrm{id}$.

We obtain the desired extension of $\mathfrak{M}$ by setting $\mathfrak{M}^+ := \bigcup_{i<\lambda} \mathfrak{B}_i$. By induction on α, it follows that $|B_\alpha| \leq \kappa \otimes |\alpha + 1|$. In particular, $|M^+| \leq \kappa \otimes \lambda$. Hence, if $\mathcal{K}$ is κ-stable then we have $\lambda \leq \kappa$ and $|M^+| = \kappa$. $\square$

Iterating the construction of the preceding lemma, we obtain the desired Galois saturated extension. For the proof that the limit really is Galois saturated, we need the following technical lemma.

Definition 2.14. Let $\mathfrak{p} \in S^{<\omega}_{\mathrm{Aut}}(B)$ be a Galois type and let $f : \mathfrak{A} \to \mathfrak{B}$ be a $\mathcal{K}$-embedding. We define the *restriction* $\mathfrak{p}|_f$ of $\mathfrak{p}$ along f as follows.

Fix a structure $\mathfrak{N} \geq_{\mathcal{K}} \mathfrak{B}$ containing a tuple $\bar{a} \subseteq N$ with

$$\mathfrak{p} = \mathrm{tp}_{\mathrm{Aut}}(\bar{a}/\mathfrak{N}, B)\,.$$

Let $\mathfrak{M}$ be the isomorphic copy of $\mathfrak{N}$ obtained by replacing all elements of $\mathrm{rng}\, f$ by their preimages in A, and let $\pi : \mathfrak{N} \to \mathfrak{M}$ be the corresponding isomorphism. We set

$$\mathfrak{p}|_f := \mathrm{tp}_{\mathrm{Aut}}(\pi(\bar{a})/\mathfrak{M}, A)\,.$$

If $\mathfrak{A} \leq_{\mathcal{K}} \mathfrak{B}$ and $f : \mathfrak{A} \to \mathfrak{B}$ is the inclusion map, then we also write $\mathfrak{p}|_A$ for $\mathfrak{p}|_f$.

Lemma 2.15. *Let $\mathcal{K}$ be an amalgamation class and $f : \mathfrak{A} \to \mathfrak{B}$ a $\mathcal{K}$-embedding. For every Galois type $\mathfrak{p} \in S^{<\omega}_{\mathrm{Aut}}(A)$, there is a Galois type $\mathfrak{q} \in S^{<\omega}_{\mathrm{Aut}}(B)$ with $\mathfrak{q}|_f = \mathfrak{p}$.*

Proof. We fix an extension $\mathfrak{C} \geq_{\mathcal{K}} \mathfrak{A}$ and a tuple $\bar{a} \subseteq C$ such that $\mathfrak{p} = \mathrm{tp}_{\mathrm{Aut}}(\bar{a}/\mathfrak{C}, A)$. By the amalgamation property, we can find an extension $\mathfrak{D} \geq_{\mathcal{K}} \mathfrak{B}$ such that there exists a $\mathcal{K}$-embedding $h : \mathfrak{C} \to \mathfrak{D}$ with $h \restriction A = f$. We can set $\mathfrak{q} := \mathrm{tp}_{\mathrm{Aut}}(h(\bar{a})/\mathfrak{D}, B)$. □

Lemma 2.16. *Let $\mathcal{K}$ be an amalgamation class, γ an ordinal, and suppose that $(\mathfrak{A}_\alpha)_{\alpha<\gamma}$ is an increasing chain of structures $\mathfrak{A}_\alpha \in \mathcal{K}$ such that $\mathfrak{A}_{\alpha+1}$ realises every type in $S^{<\omega}_{\mathrm{Aut}}(A_\alpha)$, for all α. Then their union $\bigcup_{\alpha<\gamma} \mathfrak{A}_\alpha$ is* $\mathrm{cf}(\gamma)$*-Galois saturated.*

Proof. Let $\mathfrak{U} \leq_{\mathcal{K}} \bigcup_{\alpha<\gamma} \mathfrak{A}_\alpha$ be a substructure of size $|U| < \mathrm{cf}(\gamma)$ and fix a type $\mathfrak{p} \in S^{<\omega}_{\mathrm{Aut}}(U)$. There exists an index $\alpha < \gamma$ with $U \subseteq A_\alpha$. Hence, $\mathfrak{U} \leq_{\mathcal{K}} \mathfrak{A}_\alpha$ and, by Lemma 2.15, we can find a type $\mathfrak{q} \in S^{<\omega}_{\mathrm{Aut}}(A_\alpha)$ with $\mathfrak{q}|_U = \mathfrak{p}$. By construction, $\mathfrak{q}$ is realised in $\bigcup_{\alpha<\gamma} \mathfrak{A}_\alpha \geq_{\mathcal{K}} \mathfrak{A}_{\alpha+1}$. Hence, so is $\mathfrak{p}$. □

Proposition 2.17. *Let $\mathcal{K}$ be an amalgamation class and suppose that κ is a regular cardinal. Every structure $\mathfrak{M} \in \mathcal{K}$ has a κ-Galois saturated extension $\mathfrak{M}^+ \geq_{\mathcal{K}} \mathfrak{M}$.*

Proof. We construct an increasing chain $(\mathfrak{A}_\alpha)_{\alpha<\kappa}$ as follows. We start with $\mathfrak{A}_0 := \mathfrak{M}$. For limit ordinals δ, we set $\mathfrak{A}_\delta := \bigcup_{\alpha<\delta} \mathfrak{A}_\alpha$. For the successor step, we use Lemma 2.13 to find an extension $\mathfrak{A}_{\alpha+1} \geq_{\mathcal{K}} \mathfrak{A}_\alpha$ realising all Galois types over $\mathfrak{A}_\alpha$. By Lemma 2.16, it follows that the limit $\mathfrak{M}^+ := \bigcup_{\alpha<\kappa} \mathfrak{A}_\alpha$ is κ-Galois saturated. □

As usual the existence of Galois saturated structures depends on an additional hypothesis like stability.

Theorem 2.18. *Let $\mathcal{K}$ be a Jónsson class and suppose that κ is a regular cardinal with* $\ln(\mathcal{K}) \leq \kappa < \mathrm{hn}(\mathcal{K})$. *If $\mathcal{K}$ is κ-stable then every structure $\mathfrak{M} \in \mathcal{K}$ of size $|M| \leq \kappa$ has a Galois saturated $\mathcal{K}$-extension of size κ.*

Proof. We construct an increasing chain $(\mathfrak{A}_\alpha)_{\alpha<\kappa}$ of structures $\mathfrak{A}_\alpha \in \mathcal{K}$ of size $|A_\alpha| = \kappa$ as follows. Since $\kappa < \mathrm{hn}(\mathcal{K})$ we have $\mathcal{K}_\kappa \neq \varnothing$. Using amalgamation and the joint embedding property, we can find a structure $\mathfrak{A}_0 \in \mathcal{K}$ of size $|A_0| = \kappa$ with $\mathfrak{M} \leq_{\mathcal{K}} \mathfrak{A}_0$. For limit ordinals δ, we set $\mathfrak{A}_\delta := \bigcup_{\alpha<\delta} \mathfrak{A}_\alpha$. Note that $|A_\delta| \leq |\delta| \otimes \kappa = \kappa$. For the successor step, suppose that $\mathfrak{A}_\alpha$ has already been defined. We use Lemma 2.13 to find an extension $\mathfrak{A}_{\alpha+1} \geq_{\mathcal{K}} \mathfrak{A}_\alpha$ of size $|A_{\alpha+1}| = \kappa$ that realises all types over $\mathfrak{A}_\alpha$. By Lemma 2.16, it follows that the limit $\bigcup_{\alpha<\kappa} \mathfrak{A}_\alpha$ is κ-Galois saturated. □

3. *Limits of chains*

We have seen that we can inductively construct Galois saturated structures as limits of chains. In this section we take a close look at such chains. Our aim is Theorem 4.13, which states that, under certain conditions, the union of a chain of Galois saturated structures is again Galois saturated.

Definition 3.1. Let $\mathcal{K}$ be an abstract elementary class and γ an ordinal.

(a) An increasing chain $(\mathfrak{M}_\alpha)_{\alpha<\gamma}$ is a *weak γ-chain* if each $\mathfrak{M}_{\alpha+1}$ realises every Galois type over M_α. In this case we say that $\mathfrak{M} := \bigcup_\alpha \mathfrak{M}_\alpha$ is the *weak γ-limit* of the chain, or that $\mathfrak{M}$ is a *weak γ-limit* over $\mathfrak{M}_0$.

(b) An increasing chain $(\mathfrak{M}_\alpha)_{\alpha<\gamma}$ is a *strong γ-chain* if every $\mathfrak{M}_{\alpha+1}$ is $|M_{\alpha+1}|^+$-universal over $\mathfrak{M}_\alpha$. In this case we say that $\mathfrak{M} := \bigcup_\alpha \mathfrak{M}_\alpha$ is the *strong γ-limit* of the chain, or that $\mathfrak{M}$ is a *strong γ-limit* over $\mathfrak{M}_0$.

The following observation is just a restatement of Lemma 2.16.

Lemma 3.2. *Let $\mathcal{K}$ be an amalgamation class. Every weak γ-limit is* $\mathrm{cf}(\gamma)$-*Galois saturated.*

Lemma 3.3. *Suppose that $\mathcal{K}$ is an amalgamation class and $\gamma \geq$* $\ln(\mathcal{K})$ *an ordinal. Let $\mathfrak{M}$ be a weak γ-limit over $\mathfrak{A} \leq_{\mathcal{K}} \mathfrak{M}$. Then $\mathfrak{M}$ is $|\gamma|^+$-universal over every $\mathcal{K}$-substructure $\mathfrak{A}_0 \leq_{\mathcal{K}} \mathfrak{A}$ of size $|A_0| \leq |\gamma|$.*

Proof. Let $(\mathfrak{M}_\alpha)_{\alpha<\gamma}$ be a weak γ-chain with limit $\mathfrak{M}$ and $\mathfrak{M}_0 = \mathfrak{A}$. This chain satisfies the hypothesis of Lemma 2.10. □

Corollary 3.4. *Suppose that $\mathcal{K}$ is an amalgamation class, let $\kappa \geq \ln(\mathcal{K})$ be a cardinal, and γ an ordinal. Let $(\mathfrak{M}_\alpha)_{\alpha<\kappa\gamma}$ be a weak $\kappa\gamma$-chain with $|\bigcup_{\alpha<\kappa\gamma} \mathfrak{M}_\alpha| \leq \kappa$. Then the subsequence $(\mathfrak{M}_{\kappa\alpha})_{\alpha<\gamma}$ is a strong γ-chain.*

Proof. Let $\alpha < \gamma$. The sequence $(\mathfrak{M}_{\kappa\alpha+\beta})_{\beta<\kappa}$ is a weak κ-chain over $\mathfrak{M}_{\kappa\alpha}$ with limit $\mathfrak{N} := \bigcup_{\beta<\kappa} \mathfrak{M}_{\kappa\alpha+\beta} \leq \mathfrak{M}_{\kappa(\alpha+1)}$. By the preceding lemma, $\mathfrak{N}$ is κ^+-universal over $\mathfrak{M}_{\kappa\alpha}$. Hence, so is its extension $\mathfrak{M}_{\kappa(\alpha+1)} \geq_{\mathcal{K}} \mathfrak{N}$. As $|M_{\kappa(\alpha+1)}| \leq \kappa$, the claim follows. □

Corollary 3.5. *Suppose that $\mathcal{K}$ is an amalgamation class. Let $\mathfrak{A} \in \mathcal{K}$ be a structure of size $\kappa := |A| \geq \ln(\mathcal{K})$ and let $\gamma < \kappa^+$ be an ordinal. If $\mathcal{K}$ is κ-Galois stable, then there exists a strong γ-limit $\mathfrak{M} \in \mathcal{K}$ over $\mathfrak{A}$ of size $|M| = \kappa$.*

Proof. By Corollary 3.4, it is sufficient to construct a weak $\kappa\gamma$-chain $(\mathfrak{M}_\alpha)_{\alpha<\kappa\gamma}$ over $\mathfrak{A}$ such that $|M_\alpha| = \kappa$, for all α. We define such a chain by induction on α starting with $\mathfrak{M}_0 := \mathfrak{A}$. For the inductive step, note that, given $\mathfrak{M}_\alpha$, we can use Lemma 2.13 to find a structure $\mathfrak{M}_{\alpha+1}$ with the desired properties. □

The next lemma implies that, in the definition of a strong γ-chain $(\mathfrak{M}_\alpha)_\alpha$, we could also require universality of $\mathfrak{M}_{\alpha+1}$ over every $\mathcal{K}$-substructure of $\mathfrak{M}_\alpha$.

Lemma 3.6. *Suppose that $\mathcal{K}$ is an amalgamation class and let $\mathfrak{A} \in \mathcal{K}$ be a structure of size $\ln(\mathcal{K}) \leq |A| < \kappa$. If $\mathfrak{M}$ is κ-universal over $\mathfrak{A} \leq_{\mathcal{K}} \mathfrak{M}$, then it is also κ-universal over every substructure $\mathfrak{A}_0 \leq_{\mathcal{K}} \mathfrak{A}$.*

Proof. Let $\mathfrak{A}_0 \leq_{\mathcal{K}} \mathfrak{A}$ and consider a $\mathcal{K}$-embedding $f : \mathfrak{A}_0 \to \mathfrak{C}$ with $|C| < \kappa$. By amalgamation, we can find a $\mathcal{K}$-extension $\mathfrak{C}_+ \geq_{\mathcal{K}} \mathfrak{C}$ of size $|C_+| = |C| \oplus |A| < \kappa$ and a $\mathcal{K}$-embedding $f_+ : \mathfrak{A} \to \mathfrak{C}_+$ such that $f_+ \restriction A_0 = f$. As $\mathfrak{M}$ is κ-universal over $\mathfrak{A}$, there exists a $\mathcal{K}$-embedding $h_+ : \mathfrak{C}_+ \to \mathfrak{M}$ with $h_+ \circ f_+ = \mathrm{id}_A$. Setting $h := h_+ \restriction C$ it follows that $h \circ f = h_+ \circ f_+ \restriction A_0 = \mathrm{id}_{A_0}$, as desired. □

Lemma 3.7. *Let $\mathcal{K}$ be an amalgamation class. If a structure $\mathfrak{M} \in \mathcal{K}$ realises all Galois types over $\mathfrak{U} \leq_{\mathcal{K}} \mathfrak{M}$, then it also realises all Galois type over every $\mathfrak{U}_0 \leq_{\mathcal{K}} \mathfrak{U}$.*

Proof. Let $\mathfrak{U}_0 \leq_{\mathcal{K}} \mathfrak{U}$ and $\mathfrak{p} \in S^{<\omega}_{\mathrm{Aut}}(\mathfrak{U}_0)$. By Lemma 2.15, there exists a type $\mathfrak{q} \in S^{<\omega}_{\mathrm{Aut}}(\mathfrak{U})$ with $\mathfrak{q}|_{U_0} = \mathfrak{p}$. By assumption, $\mathfrak{M}$ realises $\mathfrak{q}$. Hence, it also realises $\mathfrak{p}$. □

We conclude this section with a result stating that a strong limit is unique up to isomorphism.

Theorem 3.8. *Let $\mathcal{K}$ be an amalgamation class, $\mathfrak{A}, \mathfrak{A}' \in \mathcal{K}$ structures of size $|A|, |A'| \geq \ln(\mathcal{K})$, and let δ, δ' be limit ordinals with $\mathrm{cf}(\delta) = \mathrm{cf}(\delta')$.*

If $\mathfrak{M}$ is a strong δ-limit over $\mathfrak{A}$ and $\mathfrak{M}'$ is a strong δ'-limit over $\mathfrak{A}'$ with $|M| = |M'|$, then we can extend every isomorphism $f : \mathfrak{A} \to \mathfrak{A}'$ to an isomorphism $\pi : \mathfrak{M} \to \mathfrak{M}'$.

Proof. Fix strong chains $(\mathfrak{M}_\alpha)_{\alpha<\delta}$ and $(\mathfrak{M}'_\alpha)_{\alpha<\delta'}$ such that

$$\bigcup_{\alpha<\delta} \mathfrak{M}_\alpha = \mathfrak{M}, \quad \bigcup_{\alpha<\delta'} \mathfrak{M}'_\alpha = \mathfrak{M}', \quad \mathfrak{M}_0 = \mathfrak{A}, \quad \mathfrak{M}'_0 = \mathfrak{A}'.$$

Set $\beta := \mathrm{cf}(\delta)$ and let $h : \beta \to \delta$ and $h' : \beta \to \delta'$ be strictly increasing functions with $h(0) = 0$ and $h'(0) = 0$. We can choose h and h' such that, for every $\alpha < \beta$, $h(\alpha + 1)$ and $h'(\alpha + 1)$ are successor ordinals.

Since $|M| = |M'|$ we can find increasing chains $(\mathfrak{N}_\alpha)_{\alpha<\beta}$ and $(\mathfrak{N}'_\alpha)_{\alpha<\beta}$ of $\mathcal{K}$-substructures $\mathfrak{N}_\alpha \leq_{\mathcal{K}} \mathfrak{M}_{h(\alpha)}$ and $\mathfrak{N}'_\alpha \leq_{\mathcal{K}} \mathfrak{M}'_{h'(\alpha)}$ such that

$$\bigcup_{\alpha<\beta} \mathfrak{N}_\alpha = \mathfrak{M}, \quad \bigcup_{\alpha<\beta} \mathfrak{N}'_\alpha = \mathfrak{M}', \quad \mathfrak{N}_0 = \mathfrak{A}, \quad \mathfrak{N}'_0 = \mathfrak{A}',$$

$$\text{and} \quad |N_\alpha| = |N'_\alpha| = \min\left\{|M_{h(\alpha)}|, |M'_{h'(\alpha)}|\right\}.$$

We construct an increasing chain $(p_\alpha)_{\alpha<\beta}$ of isomorphisms $p_\alpha : \mathfrak{B}_\alpha \to \mathfrak{B}'_\alpha$ such that

$$\mathfrak{N}_\alpha \leq_{\mathcal{K}} \mathfrak{B}_\alpha \leq_{\mathcal{K}} \mathfrak{M}_{h(\alpha)},$$
$$\mathfrak{N}'_\alpha \leq_{\mathcal{K}} \mathfrak{B}'_\alpha \leq_{\mathcal{K}} \mathfrak{M}'_{h'(\alpha)+1},$$

and $\quad |B_\alpha| = |N_\alpha|$.

Then the limit $\pi := \bigcup_{\alpha<\beta} p_\alpha$ is the desired isomorphism $\pi : \mathfrak{M} \to \mathfrak{M}'$.

We start with $p_0 := f : \mathfrak{A} \to \mathfrak{A}'$. For limit ordinals γ, we set $p_\gamma := \bigcup_{\alpha<\gamma} p_\alpha$. For the successor step, suppose that $p_\alpha : \mathfrak{B}_\alpha \to \mathfrak{B}'_\alpha$ has already been defined. We fix a substructure $\mathfrak{C}' \leq_{\mathcal{K}} \mathfrak{M}'_{h'(\alpha+1)}$ such that

$$N'_{\alpha+1} \cup B'_\alpha \subseteq C' \quad \text{and} \quad |C'| = |N'_{\alpha+1}|.$$

By Lemma 3.6, $\mathfrak{M}_{h(\alpha+1)}$ is $|M_{h(\alpha+1)}|^+$-universal over $\mathfrak{B}_\alpha \leq_{\mathcal{K}} \mathfrak{M}_{h(\alpha)}$. Since $|C'| \leq |M_{h(\alpha+1)}|$, it therefore follows that there is a $\mathcal{K}$-embedding $g : \mathfrak{C}' \to \mathfrak{M}_{h(\alpha+1)}$ such that $g \circ p_\alpha = \mathrm{id}_{B_\alpha}$. Fix a $\mathcal{K}$-substructure $\mathfrak{C} \leq_{\mathcal{K}} \mathfrak{M}_{h(\alpha+1)}$ such that

$$N_{\alpha+1} \cup \operatorname{rng} g \subseteq C \quad \text{and} \quad |C| = |N_{\alpha+1}|.$$

As above, $\mathfrak{M}'_{h'(\alpha+1)+1}$ is $|M'_{h'(\alpha+1)+1}|^+$-universal over $\mathfrak{C}' \leq_{\mathcal{K}} \mathfrak{M}'_{h'(\alpha+1)}$, and we have $|C| \leq |M'_{h'(\alpha+1)+1}|$. Hence, we can find a $\mathcal{K}$-embedding $g' : \mathfrak{C} \to \mathfrak{M}'_{h'(\alpha+1)+1}$ such that $g' \circ g = \mathrm{id}_{C'}$. We take this embedding g' for our isomorphism $p_{\alpha+1} : \mathfrak{B}_{\alpha+1} \to \mathfrak{B}'_{\alpha+1}$. Then

$$\mathfrak{N}_{\alpha+1} \leq_{\mathcal{K}} \mathfrak{B}_{\alpha+1} \leq_{\mathcal{K}} \mathfrak{M}_{h(\alpha+1)},$$
$$\mathfrak{N}'_{\alpha+1} \leq_{\mathcal{K}} \mathfrak{B}'_{\alpha+1} \leq_{\mathcal{K}} \mathfrak{M}'_{h'(\alpha+1)+1},$$

and $\quad |B_{\alpha+1}| = |N_{\alpha+1}|$.

Furthermore, for $a \in B_\alpha$, we have

$$\begin{aligned} p_{\alpha+1}(a) = g'(a) &= g'((g \circ p_\alpha)(a)) \\ &= (g' \circ g)(p_\alpha(a)) = p_\alpha(a). \end{aligned}$$

Hence, $p_\alpha \subseteq p_{\alpha+1}$. ☐

Corollary 3.9. *Suppose that $\mathcal{K}$ is an amalgamation class with $\kappa \geq \ln(\mathcal{K})$, and let $\mathfrak{M}$ be a weak $\kappa\delta$-limit over $\mathfrak{A}$ of size $|M| = \kappa$ where δ is a limit ordinal with $\delta < \kappa^+$. Every strong $\kappa\delta$-limit over $\mathfrak{A}$ is isomorphic to $\mathfrak{M}$.*

Proof. By Corollary 3.4, $\mathfrak{M}$ is a strong δ-limit over $\mathfrak{A}$. Since δ is a limit ordinal we have $\mathrm{cf}(\delta) = \mathrm{cf}(\kappa\delta)$. Consequently, the claim follows from Theorem 3.8. ☐

4. *Categoricity and stability*

In this section we study the consequences of categoricity and stability for an abstract elementary class. We will see that Ehrenfeucht-Mostowski functors provide an invaluable tool in this context.

Lemma 4.1. *Let $\mathcal{K}$ be a κ-categorical abstract elementary class with the joint embedding property where $\kappa \geq \ln(\mathcal{K})$. The structure $\mathfrak{M} \in \mathcal{K}$ of size κ is $\mathcal{K}$-universal.*

Proof. Let $\mathfrak{A} \in \mathcal{K}$ be of size $|A| \leq \kappa$. By the joint embedding property, we can find $\mathcal{K}$-embeddings $f : \mathfrak{A} \to \mathfrak{N}$ and $g : \mathfrak{M} \to \mathfrak{N}$ into some structure $\mathfrak{N} \in \mathcal{K}$ of size $|N| \leq |M| \oplus |A| \oplus \ln(\mathcal{K}) = \kappa$. Since $\mathcal{K}$ is κ-categorical, there exits an isomorphism $\pi : \mathfrak{N} \to \mathfrak{M}$. It follows that $\pi \circ f$ is a $\mathcal{K}$-embedding $\mathfrak{A} \to \mathfrak{M}$. □

We start by showing that categoricity implies stability. This generalises Theorem E6.3.16.

Lemma 4.2. *Suppose that $\mathcal{K}$ is unbounded and κ-categorical, for $\kappa \geq \ln(\mathcal{K})$, and let $\mathfrak{M} \in \mathcal{K}$ be the structure of size $|M| = \kappa$. For every $\mathfrak{U} \leq_{\mathcal{K}} \mathfrak{M}$, $\mathfrak{M}$ realises at most $|U| \oplus \ln(\mathcal{K})$ Galois types over U.*

Proof. By Corollary 1.10, there exists an Ehrenfeucht-Mostowski functor $F = \mathrm{pr}_\Sigma \circ F_+$ for $\mathcal{K}$. Then $|F(\kappa)| = \kappa$ implies $F(\kappa) \cong \mathfrak{M}$. W.l.o.g. we may assume that this isomorphism is the identity. Fix a substructure $\mathfrak{U} \leq_{\mathcal{K}} \mathfrak{M}$. There is some $I \subseteq \kappa$ of size $|I| \leq |U|$ such that $\mathfrak{U} \subseteq F(I)$. Every finite tuple $\bar{a} \subseteq M = F_+(\kappa)|_\Sigma$ is of the form $a_l = t_l[\bar{\imath}]$ where t_l is a term of the expansion $F_+(\kappa)$ with parameters $\bar{\imath} \subseteq \kappa$. By enlarging the tuples $\bar{\imath}$ we may assume that these parameters are the same for every a_l. If $a'_l = t_l[\bar{\imath}']$ are elements where $\bar{\imath}$ and $\bar{\imath}'$ have the same order type over I, then we can find a linear order L extending κ and an automorphism π of L that fixes I and maps $\bar{\imath}$ to $\bar{\imath}'$. Hence, $F_+(\pi)$ is an automorphism of $F_+(L)$ fixing U and mapping $\bar{a}$ to $\bar{a}'$. Consequently, $\mathrm{tp}_{\mathrm{Aut}}(\bar{a}/\mathfrak{M}, U) = \mathrm{tp}_{\mathrm{Aut}}(\bar{a}'/\mathfrak{M}, U)$.

It follows that the number of Galois types over $\mathfrak{U}$ realised in $\mathfrak{M}$ is bounded by the number of terms $t(\bar{x})$, times the number of order types of finite tuples $\bar{\imath} \subseteq \kappa$ over I. There are at most $\ln(\mathcal{K})$ such terms and, since κ is well-ordered, at most $|I|$ such order types. □

Theorem 4.3. *An unbounded κ-categorical Jónsson class $\mathcal{K}$ is λ-Galois stable, for every cardinal* $\ln(\mathcal{K}) \leq \lambda < \kappa$.

Proof. For a contradiction, suppose that $\mathcal{K}$ is not λ-Galois stable, for some $\ln(\mathcal{K}) \leq \lambda < \kappa$. Fix a structure $\mathfrak{U} \in \mathcal{K}$ of size $|U| = \lambda$ such that $|S^{<\omega}_{\mathrm{Aut}}(U)| > \lambda$. By Proposition 2.17, we can find a $\mathcal{K}$-extension $\mathfrak{A} \geq_{\mathcal{K}} \mathfrak{U}$ of size $|A| = \lambda^+$ realising λ^+ types from $S^{<\omega}_{\mathrm{Aut}}(U)$.

Let $\mathfrak{M} \in \mathcal{K}$ be a structure of size κ. We have seen in Lemma 4.1 that $\mathfrak{M}$ is κ^+-universal. Hence, there exists a $\mathcal{K}$-embedding $f : \mathfrak{A} \to \mathfrak{M}$. It follows that $\mathfrak{M}$ realises at least λ^+ Galois types over $f[U]$. This contradicts Lemma 4.2. □

Lemma 4.4. *Let $\mathcal{K}$ be an amalgamation class. If $\mathcal{K}$ is κ-categorical for $\kappa > \ln(\mathcal{K})$, then the structure $\mathfrak{M} \in \mathcal{K}$ of size κ is* $\mathrm{cf}(\kappa)$*-Galois saturated.*

Proof. Starting with an arbitrary structure $\mathfrak{A}_0 \in \mathcal{K}_{<\kappa}$ we use Lemma 2.13 to construct a strictly increasing chain $(\mathfrak{A}_\alpha)_{\alpha<\kappa}$ of structures $\mathfrak{A}_\alpha \in \mathcal{K}$ of size $|A_\alpha| < \kappa$ such that $\mathfrak{A}_{\alpha+1}$ realises every Galois type over A_α.

By Lemma 2.16, the union $\mathfrak{A}_\kappa := \bigcup_{\alpha<\kappa} \mathfrak{A}_\alpha$ is $\mathrm{cf}(\kappa)$-Galois saturated. Since $|A_\kappa| = \kappa$ and $\mathcal{K}$ is κ-categorical, we have $\mathfrak{A}_\kappa \cong \mathfrak{M}$. Hence, $\mathfrak{M}$ is $\mathrm{cf}(\kappa)$-Galois saturated. □

Corollary 4.5. *Let $\mathcal{K}$ be an unbounded Jónsson class. If $\mathcal{K}$ is κ-categorical, for $\kappa > \ln(\mathcal{K})$, then $\mathcal{K}$ contains Galois saturated structures of size λ, for every regular cardinal λ with* $\ln(\mathcal{K}) \leq \lambda \leq \kappa$.

Proof. For $\lambda = \kappa$, we have already proved the claim in Lemma 4.4. For $\lambda < \kappa$, it follows from Theorems 4.3 and 2.18. □

Next, we consider an analogue of the notion of an indiscernible sequence for abstract elementary classes. The following result is comparable to Theorem E5.3.13.

Lemma 4.6. *Let $\mathcal{K}$ be an amalgamation class and let $F : \mathfrak{Lin} \to \mathfrak{Emb}(\mathcal{K})$ be an Ehrenfeucht-Mostowski functor for $\mathcal{K}$ with spine s. Suppose that I is a linear order, $\bar{\imath} \in [I]^{<\omega}$ a finite tuple, and $\sigma : \bar{\imath} \to \bar{\imath}$ a permutation such that*

$$\mathrm{tp}_{\mathrm{Aut}}\big(s_I(\bar{\imath}) \big/ F(I), \varnothing\big) \neq \mathrm{tp}_{\mathrm{Aut}}\big(s_I(\sigma(\bar{\imath})) \big/ F(I), \varnothing\big).$$

Then $\mathcal{K}$ is not κ-stable, for any $\kappa \geq \ln(\mathcal{K})$.

Proof. We can write each permutation as a product of transpositions. Hence, suppose that $\sigma = \sigma_0 \circ \cdots \circ \sigma_n$, where each $\sigma_l : \bar{\imath} \to \bar{\imath}$ is a permutation of $\bar{\imath}$ interchanging two consecutive components of $\bar{\imath}$. There is at least one index l such that

$$\mathrm{tp}_{\mathrm{Aut}}\big(s_I(\bar{\imath}) \big/ F(I), \varnothing\big) \neq \mathrm{tp}_{\mathrm{Aut}}\big(s_I(\sigma_l(\bar{\imath})) \big/ F(I), \varnothing\big),$$

since, otherwise, we would have

$$\mathrm{tp}_{\mathrm{Aut}}\big(s_I(\bar{\imath}) \big/ F(I), \varnothing\big) = \mathrm{tp}_{\mathrm{Aut}}\big(s_I(\sigma(\bar{\imath})) \big/ F(I), \varnothing\big).$$

Replacing σ by σ_l we may therefore assume that $\bar{\imath} = \bar{k}ij\bar{m}$ and $\sigma(\bar{\imath}) = \bar{k}ji\bar{m}$ where $\bar{k} < i < j < \bar{m}$.

Let J be a linear order of size $|J| > \kappa$ containing a dense subset $J_0 \subseteq J$ of size $|J_0| = \kappa$. Set $\mathfrak{M} := F(J)$ and $\mathfrak{U} := F(J_0)$. Since $|U| = \kappa$, it is sufficient to show that

$$\mathrm{tp}_{\mathrm{Aut}}(s_J(x)/\mathfrak{M}, U) \neq \mathrm{tp}_{\mathrm{Aut}}(s_J(y)/\mathfrak{M}, U), \quad \text{for all } x \neq y \text{ in } J.$$

Fix elements $x < y$ in J. To prove that the Galois types of $s_J(x)$ and $s_J(y)$ over U are different, we choose indices $w, \bar{u}, \bar{v} \subseteq J_0$ such that $x < w < y$ and the tuples $\bar{u}xy\bar{v}$ and $\bar{k}ij\bar{m}$ have the same order type. It follows that

$$\begin{aligned}
\mathrm{tp}_{\mathrm{Aut}}(s_J(xw\bar{u}\bar{v})/\mathfrak{M}, \varnothing) &= \mathrm{tp}_{\mathrm{Aut}}(s_I(ij\bar{k}\bar{m})/F(I), \varnothing) \\
&\neq \mathrm{tp}_{\mathrm{Aut}}(s_I(ji\bar{k}\bar{m})/F(I), \varnothing) \\
&= \mathrm{tp}_{\mathrm{Aut}}(s_J(yw\bar{u}\bar{v})/\mathfrak{M}, \varnothing).
\end{aligned}$$

Since $s_J(w\bar{u}\bar{v}) \subseteq U$ the claim follows. □

We have already seen that κ-categorical classes are stable and, therefore, they contain Galois saturated structures of all regular cardinals below κ. We conclude this section with some results about the existence of Galois saturated structures of *singular* cardinality.

Lemma 4.7. *Let $\mathcal{K}$ be a κ-categorical amalgamation class, let $F : \mathfrak{Lin} \to \mathfrak{Emb}(\mathcal{K})$ be an Ehrenfeucht-Mostowski functor for $\mathcal{K}$, let $\lambda > \ln(\mathcal{K})$ be a cardinal, and set $\mathcal{C}_\lambda := \{ \mu^+ \mid \mu < \lambda \}$. Then $F(I)$ is λ-Galois saturated, for every $\mathcal{C}_\lambda$-universal linear order I of size $\lambda \leq |I| < \mathrm{cf}(\kappa)$.*

Proof. It is sufficient to show that $F(I)$ is μ^+-Galois saturated, for every $\mu < \lambda$. Since I is $\mathcal{C}_\lambda$-universal there is some embedding $h : \mu^+ \to I$. Set $A := \Downarrow \operatorname{rng} h$, $B := I \smallsetminus A$, and $J := A + \kappa + B$. Then $|F(J)| = \kappa$. Since $\mu^+ < \mathrm{cf}(\kappa)$ it therefore follows by Lemma 4.4 that $F(J)$ is μ^+-Galois saturated.

To show that also $F(I)$ is μ^+-Galois saturated, we consider a substructure $\mathfrak{U} \leq_{\mathcal{K}} F(I)$ of size $|U| = \mu$ and a type $\mathfrak{p} \in S^{<\omega}_{\mathrm{Aut}}(U)$. Let $\mathfrak{q} \in S^{<\omega}_{\mathrm{Aut}}(F(h)[U])$ be the type with $\mathfrak{q}|_{F(h)} = \mathfrak{p}$. Then $\mathfrak{q}$ is realised by some tuple $\bar a \subseteq F(J)$. Each a_l is denoted by a term $t_l[\bar\imath \bar k]$ (in the Skolem expansion) with parameters $\bar\imath \subseteq I$ and $\bar k \subseteq J \smallsetminus I$. By enlarging the tuples of parameters we may assume without loss of generality that the parameters $\bar\imath \bar k$ are the same for every l. Let $J_0 \subseteq J$ be a set of size $|J_0| = \mu$ such that $F(h)[U] \cup \bar\imath \subseteq F(J_0)$. Since μ^+ is regular, there is some $\alpha < \mu^+$ such that $J_0 \cap A \subseteq h[\downarrow\alpha]$. Hence, there is some tuple $\bar k' \subseteq \operatorname{rng} h$ such that $\bar k$ and $\bar k'$ have the same order type over $J_0 \cup \bar\imath$. Setting $b_l := t_l[\bar\imath \bar k']$ it follows that $\mathrm{tp}_{\mathrm{Aut}}(\bar b/F(I), U) = \mathfrak{p}$. □

In the following $\lambda^{<\omega}$ denotes the linear order $\langle \lambda^{<\omega}, \leq_{\mathrm{lex}} \rangle$ where $\leq_{\mathrm{lex}}$ is the lexicographic order on $\lambda^{<\omega}$.

Proposition 4.8. *Let $\mathcal{K}$ be an unbounded amalgamation class that is κ-categorical, for some regular cardinal $\kappa > \ln(\mathcal{K})$. If $F : \mathfrak{Lin} \to \mathfrak{Emb}(\mathcal{K})$ is an Ehrenfeucht-Mostowski functor for $\mathcal{K}$, then*

(a) *$F(\lambda)$ is Galois saturated, for every $\ln(\mathcal{K}) < \lambda \leq \kappa$;*

(b) *$F(\lambda^{<\omega}\alpha)$ is Galois saturated, for every cardinal $\ln(\mathcal{K}) < \lambda \leq \kappa$ and every ordinal $\alpha < \lambda^+$.*

Proof. For $\lambda < \kappa$, the claims follow from Lemma 4.7 since the orders $\lambda^{<\omega}\alpha$ and λ are both $\mathcal{C}_\lambda$-universal. For $\lambda = \kappa$, note that $F(\kappa^{<\omega}\alpha) \cong F(\kappa)$ is the only structure in $\mathcal{K}$ of size κ. This structure is Galois saturated by Corollary 4.5. □

We can use structures of the form $F(\lambda^{<\omega}\alpha)$ to build strong δ-chains. We start by proving an universality lemma for the order $\lambda^{<\omega}$.

Lemma 4.9. *Let λ be a cardinal. For every ordinal $\beta < \lambda^+$, there exists an embedding $g : \beta \to \lambda^{<\omega}$.*

Proof. We define g by induction on β. If $\beta \leq \lambda$ then we can set $g(\alpha) := \langle\alpha\rangle$, for all $\alpha < \beta$. For the successor step, suppose that $\beta = \gamma + 1$ and let $g_0 : \gamma \to \lambda^{<\omega}$ be the embedding obtained by inductive hypothesis. We define $g : \beta \to \lambda^{<\omega}$ by

$$g(\alpha) := \begin{cases} \langle 0\rangle \cdot g_0(\alpha) & \text{for } \alpha < \gamma\,, \\ \langle 1\rangle & \text{for } \alpha = \gamma\,. \end{cases}$$

If β is a limit ordinal, we fix an increasing chain $(\gamma_i)_{i<\lambda}$ of ordinals $\lambda \leq \gamma_i < \beta$ with $\sup_i \gamma_i = \beta$. By inductive hypothesis, there are embeddings $g_i : \gamma_i \to \lambda^{<\omega}$. We define $g : \beta \to \lambda^{<\omega}$ by

$$g(\alpha) := \langle i\rangle \cdot g_i(\alpha) \quad \text{where } i \text{ is the least index with } \alpha < \gamma_i\,. \qquad \square$$

Lemma 4.10. *Let $\mathcal{K}$ be a κ-categorical amalgamation class where κ is regular, let $\ln(\mathcal{K}) \leq \lambda < \kappa$ be a cardinal, and $\delta < \lambda^+$ a limit ordinal. Suppose that $F : \mathfrak{Lin} \to \mathfrak{Emb}(\mathcal{K})$ is an Ehrenfeucht-Mostowski functor for $\mathcal{K}$.*

(a) *$(F(\lambda^{<\omega}\alpha))_{\alpha<\delta}$ is a strong δ-chain over $F(\lambda^{<\omega})$.*

(b) *If $\mathfrak{M}$ is a strong δ-limit over $F(\lambda^{<\omega})$ of size $|\mathfrak{M}| = \lambda$, then $\mathfrak{M} \cong F(\lambda^{<\omega}\delta)$.*

Proof. (b) follows immediately by (a) and Theorem 3.8.

(a) We have to show that $F(\lambda^{<\omega}(\alpha+1))$ is λ^+-universal over $F(\lambda^{<\omega}\alpha)$. Let $f : F(\lambda^{<\omega}\alpha) \to \mathfrak{C}$ be a $\mathcal{K}$-embedding with $|C| \leq \lambda$. Since $\mathcal{K}$ is κ-categorical, we know by Lemma 4.4 that $F(\lambda^{<\omega}\kappa)$ is Galois saturated. In particular, $F(\lambda^{<\omega}\kappa)$ is λ^+-universal over $F(\lambda^{<\omega}\alpha)$. Hence, we can find a $\mathcal{K}$-embedding $g : \mathfrak{C} \to F(\lambda^{<\omega}\kappa)$ such that $g \circ f = \mathrm{id}$. There exists a set $I \subseteq \lambda^{<\omega}\kappa$ of size $|I| = \lambda$ such that $\operatorname{rng} g \subseteq F(I)$. Setting $I_0 := I \cap \lambda^{<\omega}\alpha$ and $I_1 := I \setminus \lambda^{<\omega}\alpha$, we obtain a partition $I = I_0 \cup I_1$ with $I_0 < I_1$. Since I_1 is a well-order with $\operatorname{ord}(I_1) < \lambda^+$, we can apply Lemma 4.9 to find an embedding $\sigma_1 : I_1 \to \lambda^{<\omega}$. Using σ_1, we define an embedding $\sigma : I \to \lambda^{<\omega}(\alpha+1)$ by

$$\sigma(i) := \begin{cases} i & \text{if } i \in I_0 , \\ \lambda^{<\omega}\alpha + \sigma_1(i) & \text{if } i \in I_1 . \end{cases}$$

Setting $h := F(\sigma) \circ g$ we obtain a $\mathcal{K}$-embedding $h : \mathfrak{C} \to F(\lambda^{<\omega}(\alpha+1))$ with

$$h \circ f = F(\sigma) \circ g \circ f = F(\sigma) \circ \mathrm{id}_{F(\lambda^{<\omega}\alpha)} = \mathrm{id}_{F(\lambda^{<\omega}\alpha)} . \qquad \square$$

Using these technical results about Ehrenfeucht-Mostowski functors we can prove the following two theorems on the existence of Galois saturated structures.

Theorem 4.11. *Suppose that $\mathcal{K}$ is an unbounded κ-categorical Jónsson class where κ is regular. Let $\mathfrak{A} \in \mathcal{K}$ be a structure of size $|A| = \lambda$ where $\ln(\mathcal{K}) < \lambda < \kappa$, and let $\delta < \lambda^+$ be a limit ordinal. Every strong δ-limit $\mathfrak{M}$ over $\mathfrak{A}$ of size $|M| = \lambda$ is Galois saturated.*

Proof. Let $F : \mathfrak{Lin} \to \mathfrak{Emb}(\mathcal{K})$ be an Ehrenfeucht-Mostowski functor for $\mathcal{K}$ and let $(\mathfrak{M}_\alpha)_{\alpha<\delta}$ be a strong δ-chain over $\mathfrak{A}$ with limit $\mathfrak{M}$. According to Proposition 4.8, the structure $F(\lambda^{<\omega})$ is Galois saturated and has size λ. By Lemma 2.8 (b), $F(\lambda^{<\omega})$ is λ^+-universal. Hence, there exists a $\mathcal{K}$-embedding $f : \mathfrak{A} \to F(\lambda^{<\omega})$. Since $\mathfrak{M}_1$ is λ^+-universal over $\mathfrak{M}_0 = \mathfrak{A}$, there also exists a $\mathcal{K}$-embedding $g : F(\lambda^{<\omega}) \to \mathfrak{M}_1$ with $g \circ f = \mathrm{id}_A$.

Replacing the sequence $(\mathfrak{M}_\alpha)_\alpha$ by isomorphic copies, we may therefore assume that

$$\mathfrak{A} \leq_{\mathcal{K}} F(\lambda^{<\omega}) \leq_{\mathcal{K}} \mathfrak{M}_1 .$$

Since $\mathfrak{M}_2$ is λ^+-universal over $\mathfrak{M}_1$, it is also λ^+-universal over $F(\lambda^{<\omega})$. Let $(\mathfrak{M}'_\alpha)_{\alpha<\delta}$ be the sequence obtained from $(\mathfrak{M}_\alpha)_{\alpha<\delta}$ by replacing the first two entries $\mathfrak{M}_0, \mathfrak{M}_1$ by the single entry $F(\lambda^{<\omega})$. Then $(\mathfrak{M}'_\alpha)_{\alpha<\delta}$ is also a strong δ-chain with limit $\mathfrak{M}$. By Lemma 4.10 (b), we have $\mathfrak{M} \cong F(\lambda^{<\omega}\delta)$. Since $\lambda^{<\omega}\delta$ is $\mathcal{C}_\lambda$-universal, it follows by Lemma 4.7 that $\mathfrak{M}$ is λ-Galois saturated. □

Using the fact that Galois saturated structures of the same cardinality are isomorphic, we obtain the following strengthening of Theorem 3.8.

Corollary 4.12. *Suppose that $\mathcal{K}$ is an unbounded Jónsson class that is κ-categorical, for some regular cardinal κ. Let λ be a cardinal with* $\ln(\mathcal{K}) < \lambda < \kappa$ *and let $\delta, \delta' < \lambda^+$ be limit ordinals. If $\mathfrak{M}, \mathfrak{M}', \mathfrak{A}, \mathfrak{A}' \in \mathcal{K}$ are structures of size λ such that $\mathfrak{M}$ is a strong δ-limit over $\mathfrak{A}$ and $\mathfrak{M}'$ is a strong δ'-limit over $\mathfrak{A}'$, then $\mathfrak{M} \cong \mathfrak{M}'$.*

Our final theorem concerns unions of Galois saturated structures. One can show that we can do without the assumption that λ is a limit cardinal, but the proof is much more involved for regular cardinals λ.

Theorem 4.13. *Let $\mathcal{K}$ be an unbounded κ-categorical Jónsson class where κ is regular, and let λ be a limit cardinal with* $\ln(\mathcal{K}) < \lambda < \kappa$. *If $(\mathfrak{M}_\alpha)_{\alpha<\delta}$ is an increasing chain of Galois saturated structures $\mathfrak{M}_\alpha \in \mathcal{K}$ of size $|M_\alpha| = \lambda$ with $\delta < \lambda^+$, then the union $\bigcup_{\alpha<\delta} \mathfrak{M}_\alpha$ is also Galois saturated.*

Proof. Let $\mathfrak{N} := \bigcup_{\alpha<\delta} \mathfrak{M}_\alpha$ be the limit. Then $|N| \leq |\delta| \otimes \lambda = \lambda$. To show that $\mathfrak{N}$ is Galois saturated fix a structure $\mathfrak{U} \leq_{\mathcal{K}} \mathfrak{N}$ of size $\mu := |U| < \lambda$ and some type $\mathfrak{p} \in S^{<\omega}_{\mathrm{Aut}}(U)$. W.l.o.g. we may assume that $\mu \geq \ln(\mathcal{K})$. Note that λ being a limit implies that $\mu^{++} < \lambda$.

The set $I := \{\, \alpha < \delta \mid (M_{\alpha+1} \setminus M_\alpha) \cap U \neq \emptyset \,\}$ has size $|I| \leq |U| = \mu$. Consequently, there exists a cofinal strictly increasing map $f : \mu_0 \to I$

where $\mu_0 := \mathrm{cf}(\mu) \leq \mu$. We construct a strong μ_0-chain $(\mathfrak{N}_\alpha)_{\alpha<\mu_0}$ where each $\mathfrak{N}_\alpha \leq_{\mathcal{K}} \mathfrak{M}_{f(\alpha)}$ has size $|N_\alpha| = \mu^+$ and, for all $\alpha < \mu_0$, we have

$$U \cap M_{f(\alpha+1)} \subseteq N_{\alpha+1} \subseteq M_{f(\alpha+1)} .$$

We define $\mathfrak{N}_\alpha$ by induction on α. We start with an arbitrary structure $\mathfrak{N}_0 \leq_{\mathcal{K}} \mathfrak{M}$ of size $|N_0| = \mu^+$. For limit ordinals γ, we set $\mathfrak{N}_\gamma := \bigcup_{\alpha<\gamma} \mathfrak{N}_\alpha$.

For the successor step, suppose that $\mathfrak{N}_\alpha$ has already been defined. We construct a weak μ^+-chain $(\mathfrak{B}_\beta)_{\beta<\mu^+}$ with $|B_\beta| = \mu^+$ as follows. We start with an arbitrary structure $\mathfrak{B}_0 \leq_{\mathcal{K}} \mathfrak{M}_{f(\alpha+1)}$ of size $|B_0| = \mu^+$ such that $N_\alpha \cup (U \cap M_{f(\alpha+1)}) \subseteq B_0$. Then we use Lemma 2.13 to inductively define $\mathfrak{B}_\beta$, for $0 < \beta < \mu^+$. Since $\mathcal{K}$ is μ^+-Galois stable, we can choose all $\mathfrak{B}_\beta$ of size $|B_\beta| = \mu^+$. Since $\mathfrak{M}_{f(\beta+1)}$ is μ^{++}-Galois saturated, we can further choose $\mathfrak{B}_\beta$ such that $\mathfrak{B}_\beta \leq_{\mathcal{K}} \mathfrak{M}_{f(\beta+1)}$. Let $\mathfrak{N}_{\alpha+1} := \bigcup_{\beta<\mu^+} \mathfrak{B}_\beta$ be the limit. By Lemma 3.3, $\mathfrak{N}_{\alpha+1}$ is μ^{++}-universal over $\mathfrak{B}_0 \geq_{\mathcal{K}} \mathfrak{N}_\alpha$.

We have constructed a strong μ_0-chain $(\mathfrak{N}_\alpha)_{\alpha<\mu_0}$ whose limit $\mathfrak{A} := \bigcup_{\alpha<\mu_0} \mathfrak{N}_\alpha$ has size $|A| = \mu_0 \otimes \mu^+ = \mu^+$. Since $|N_0| = \mu^+$ it follows by Theorem 4.11 that $\mathfrak{A}$ is Galois saturated. Consequently, $\mathfrak{p}$ is realised in $\mathfrak{A} \leq_{\mathcal{K}} \mathfrak{N}$. □

Part F.

Independence and Forking

F1. Geometries

1. Dependence relations

We have seen that a vector space or an algebraically closed field (of a given characteristic) is uniquely determined by, respectively, its dimension and its transcendence degree. In this chapter we try to generalise these two results. We investigate first-order theories whose models are uniquely determined by some kind of dimension. We start by introducing an abstract notion of dimension. As for vector spaces and algebraically closed fields, this notion is based on a closure operator. With these tools in hand we can then prove categoricity results for certain theories. Our first application will be Theorem 4.13, which states that two models of the same dimension are isomorphic.

Definition 1.1. (a) A *dependence relation* on a set A is a system $D \subseteq \wp(A)$ with the property that

$$X \in D \quad \text{iff} \quad X_0 \in D \text{ for some nonempty finite } X_0 \subseteq X\,.$$

A subset $X \subseteq A$ is *D-dependent* if $X \in D$. Otherwise, it is called *D-independent.*

(b) An element $a \in A$ *D-depends* on a set $X \subseteq A$ if $a \in X$ or there is an D-independent subset $I \subseteq X$ such that $I \cup \{a\}$ is D-dependent. The set of all elements D-depending on X is denoted by $\langle\!\langle X \rangle\!\rangle_D$.

(c) A dependence relation D on A is *transitive* if $\langle\!\langle \langle\!\langle X \rangle\!\rangle_D \rangle\!\rangle_D = \langle\!\langle X \rangle\!\rangle_D$, for all $X \subseteq A$.

Remark. Note that, if I is D-independent then we have $a \in \langle\!\langle I \rangle\!\rangle_D$ if and only if $I \cup \{a\}$ is D-dependent.

Example. (a) Let $\mathfrak{V}$ be a $\mathfrak{K}$-vector space. Then

$$D := \{ X \subseteq V \mid X \text{ is linearly dependent} \}$$

is a transitive dependence system on V.

(b) Let $\mathfrak{K}$ be a field. Then

$$D := \{ X \subseteq K \mid X \text{ is algebraically dependent} \}$$

is a transitive dependence system on K.

(c) Let $\mathfrak{G} = \langle V, E \rangle$ be an undirected graph. Then

$$D := \{ X \subseteq E \mid E \text{ contains a cycle} \}$$

is a transitive dependence system on E.

Lemma 1.2. *Let D be a transitive dependence relation on A. The function $c : X \mapsto \langle\!\langle X \rangle\!\rangle_D$ is a finitary closure operator with the exchange property.*

Proof. By definition c is finitary. To show that it is a closure operator note that we have $X \subseteq c(X)$ since all elements of X D-depend on X. As D is transitive we further have $c(c(X)) = c(X)$. Finally, if $X \subseteq Y$ then every element D-depending on X also D-depends on Y. Hence, $c(X) \subseteq c(Y)$.

For the exchange property, suppose that $b \in c(X \cup \{a\}) \setminus c(X)$. Then there is a D-independent subset $I \subseteq X \cup \{a\}$ with $I \cup \{b\} \in D$. Let $I_0 := I \setminus \{a\}$. Note that $I' := I_0 \cup \{b\}$ is D-independent since, otherwise, we would have $b \in c(I_0) \subseteq c(X)$. Therefore, $I' \cup \{a\} \in D$ implies that $a \in c(I') \subseteq c(X \cup \{b\})$, as desired. □

Lemma 1.3. *Let D be a transitive dependence relation, I a D-independent set, and $I_0 \subseteq I$. If $a \in \langle\!\langle I \rangle\!\rangle_D \setminus \langle\!\langle I_0 \rangle\!\rangle_D$ then there exists an element $b \in I \setminus I_0$ such that $I' := (I \setminus \{b\}) \cup \{a\}$ is D-independent and $b \in \langle\!\langle I' \rangle\!\rangle_D$.*

Proof. Since $a \in \langle\!\langle I \rangle\!\rangle_D$ there is some D-independent subset $J \subseteq I$ such that $J \cup \{a\} \in D$. Choose J minimal. Since $a \notin \langle\!\langle I_0 \rangle\!\rangle_D$ we have $J \nsubseteq I_0$.

Fix some element $b \in J \setminus I_0$ and set $J' := J \setminus \{b\}$ and $I' := I \setminus \{b\}$. By minimality of J we have $J' \cup \{a\} \notin D$. Consequently, $b \in \langle\!\langle J' \cup \{a\} \rangle\!\rangle_D \subseteq \langle\!\langle I' \cup \{a\} \rangle\!\rangle_D$.

It remains to prove that $I' \cup \{a\}$ is D-independent. For a contradiction, suppose that $I' \cup \{a\} \in D$. Then $a \in \langle\!\langle I' \rangle\!\rangle_D$. Since D is transitive it follows that $b \in \langle\!\langle I' \cup \{a\} \rangle\!\rangle_D \subseteq \langle\!\langle I' \rangle\!\rangle_D$. Consequently, $I = I' \cup \{b\}$ is not D-independent. Contradiction. □

We can characterise transitive dependence systems in terms of closure operators with the exchange property.

Proposition 1.4. (a) *If c is a finitary closure operator on A with the exchange property, then*

$$D := \{\, X \subseteq A \mid \text{there is some } a \in X \text{ with } a \in c(X \setminus \{a\}) \,\}$$

is a transitive dependence relation with $c(X) = \langle\!\langle X \rangle\!\rangle_D$, for all X.

(b) *A subset $D \subseteq \wp(A)$ is a transitive dependence relation if and only if the function $c : X \mapsto \langle\!\langle X \rangle\!\rangle_D$ is a finitary closure operator with the exchange property.*

Proof. (a) To show that D is a dependence relation let $X \in D$. We have to find a finite subset $X_0 \subseteq X$ with $X_0 \in D$. By definition, there is some element $a \in X$ with $a \in c(X \setminus \{a\})$. Since c is finitary it follows that there is some $X_0 \subseteq X \setminus \{a\}$ with $a \in c(X_0)$. Consequently, $X_0 \cup \{a\} \in D$.

It remains to show that D is transitive and that $c(X) = \langle\!\langle X \rangle\!\rangle_D$. We start with the latter. Let $a \in c(X)$ and choose a minimal subset $X_0 \subseteq X$ with $a \in c(X_0)$. Then there is no $b \in X_0$ with $b \in c(X_0 \setminus \{b\})$ since, otherwise, $c(X_0) = c(X_0 \setminus \{b\})$ and X_0 would not be minimal. It follows that X_0 is D-independent while $X_0 \cup \{a\}$ is not. Consequently, we have $a \in \langle\!\langle X \rangle\!\rangle_D$.

Conversely, suppose that $a \in \langle\!\langle X \rangle\!\rangle_D$. Then there is a D-independent subset $I \subseteq X$ with $I \cup \{a\} \in D$. Hence, we can find an element $b \in I \cup \{a\}$ such that $b \in c((I \cup \{a\}) \setminus \{b\})$. If $b = a$ then we have $a \in c(I) \subseteq c(X)$, as desired. Otherwise, let $I_0 := I \setminus \{b\}$. Since I is D-independent we have $b \notin c(I_0)$. Therefore, $b \in c(I_0 \cup \{a\}) \setminus c(I_0)$ implies that $a \in c(I_0 \cup \{b\}) \subseteq c(X)$.

Finally, note that $c \circ c = c$ implies that D is transitive.

(b) ($\Rightarrow$) was already proved in Lemma 1.2. ($\Leftarrow$) By (a), we only have to show that, if D and D' are sets such that $\langle\!\langle X \rangle\!\rangle_D = \langle\!\langle X \rangle\!\rangle_{D'}$, for all $X \subseteq A$, then we have $D = D'$. By symmetry, suppose that there is a set $X \in D \setminus D'$. Then there is a finite nonempty subset $X_0 \subseteq X$ with $X_0 \in D \setminus D'$. Choose X_0 such that its size is minimal and fix some element $a \in X_0$. By minimality we have $X_0 \setminus \{a\} \notin D$. This implies that $a \in \langle\!\langle X_0 \setminus \{a\} \rangle\!\rangle_D$. But $X_0 \notin D'$ implies $X_0 \setminus \{a\} \notin D'$. Therefore, $a \notin \langle\!\langle X_0 \setminus \{a\} \rangle\!\rangle_{D'} = \langle\!\langle X_0 \setminus \{a\} \rangle\!\rangle_D$. A contradiction. □

We can use this proposition to translate between dependence relations and closure operators. In the following we will use the terminology for both interchangeably, e.g., we will speak of independent sets with respect to a closure operator.

Using dependence relations or, equivalently, closure operators with the exchange property, we can introduce bases and dimensions as for vector spaces.

Definition 1.5. Let D be a dependence relation on A. A set $X \subseteq A$ is *D-spanning* if $\langle\!\langle X \rangle\!\rangle_D = A$. A *$D$-basis* is a D-spanning set which is D-independent.

Lemma 1.6. *Let D be a transitive dependence relation on A and $X \subseteq A$ a set. The following statements are equivalent:*

(1) *X is a maximal D-independent set.*

(2) *X is a minimal D-spanning set.*

(3) *X is a D-basis.*

Proof. (1) $\Rightarrow$ (2) Let X be maximal D-independent and suppose that there is some element $a \in A \setminus \langle\!\langle X \rangle\!\rangle_D$. Since X is D-independent we have $X \cup \{a\} \notin D$ and X is not maximal.

(2) $\Rightarrow$ (3) Let X be minimal D-spanning. For a contradiction suppose that $X \in D$. Let $X_0 \subseteq X$ be a minimal subset with $X_0 \in D$ and fix some element $a \in X_0$. By minimality, $I := X_0 \setminus \{a\}$ is D-independent. Hence,

$a \in \langle\!\langle I \rangle\!\rangle_D \subseteq \langle\!\langle X \setminus \{a\} \rangle\!\rangle_D$. By transitivity, it follows that $\langle\!\langle X \setminus \{a\} \rangle\!\rangle_D = \langle\!\langle X \rangle\!\rangle_D = A$. This contradicts the minimality of X.

(3) $\Rightarrow$ (1) Every D-basis X is D-independent. If X were not maximal, we could find an element $a \in A \setminus X$ such that $X \cup \{a\}$ were D-independent. But this would imply that $a \notin \langle\!\langle X \rangle\!\rangle_D = A$. A contradiction. □

Once we have shown that all bases have the same cardinality, we obtain a well-defined notion of dimension.

Lemma 1.7 (Exchange Lemma). *Let D be a transitive dependence relation on A. If I is D-independent and S is D-spanning then there exists a subset $S_0 \subseteq S$ with $I \cap S_0 = \varnothing$ such that $I \cup S_0$ is a D-basis.*

Proof. The set $F := \{ J \mid J \text{ is } D\text{-independent with } I \subseteq J \subseteq I \cup S \}$ is inductively ordered by $\subseteq$ since $\bigcup C \in D$ would imply that there is a finite subset $C_0 \subseteq C$ with $\bigcup C_0 \in D$. Consequently, F has a maximal element B. By maximality, every element of $S \setminus B$ D-depends on B. Hence, $S \subseteq \langle\!\langle B \rangle\!\rangle_D$ implies that $\langle\!\langle B \rangle\!\rangle_D \supseteq \langle\!\langle S \rangle\!\rangle_D = A$, and B is a D-basis. Setting $S_0 := B \setminus I$ yields the desired subset of S. □

Lemma 1.8. *Let D be a transitive dependence relation on A. If I, J are D-independent sets with $J \subseteq \langle\!\langle I \rangle\!\rangle_D$ then $|J| \leq |I|$.*

Proof. Since D induces a transitive dependence relation on $\langle\!\langle I \rangle\!\rangle_D$ we may assume that $A = \langle\!\langle I \rangle\!\rangle_D$ and that I is a D-basis.

First, suppose that J is finite. We prove the claim by induction on $|J \setminus I|$. If $J \subseteq I$ then there is nothing to do. Hence, suppose that there is some element $a \in J \setminus I$, and set $H := I \cap J$. Since J is D-independent we have $a \in \langle\!\langle I \rangle\!\rangle_D \setminus \langle\!\langle H \rangle\!\rangle_D$. By Lemma 1.3, we can find an element $b \in I \setminus H$ such that $I_0 := (I \setminus \{b\}) \cup \{a\}$ is D-independent and $b \in \langle\!\langle I_0 \rangle\!\rangle_D$. By transitivity of D it follows that $J \subseteq \langle\!\langle I \rangle\!\rangle_D \subseteq \langle\!\langle I_0 \cup \{b\} \rangle\!\rangle_D = \langle\!\langle I_0 \rangle\!\rangle_D$. Since $|J \setminus I_0| < |J \setminus I|$ we can apply the induction hypothesis to conclude that $|J| \leq |I_0| = |I|$.

It remains to consider the case that J is infinite. If I were finite, we could choose a subset $J_0 \subseteq J$ of size $|J_0| = |I| + 1$. This would contradict the finite case proved above. Hence, I is also infinite. Since the operator $X \mapsto \langle\!\langle X \rangle\!\rangle_D$ is finitary we have

$$J \subseteq \bigcup \{ \langle\!\langle I_0 \rangle\!\rangle_D \mid I_0 \subseteq I \text{ is finite} \} .$$

If $I_0 \subseteq I$ is finite, we have seen above that $|J \cap \langle\!\langle I_0 \rangle\!\rangle_D| \leq |I_0|$. Consequently,

$$J = \bigcup \{ J \cap \langle\!\langle I_0 \rangle\!\rangle_D \mid I_0 \subseteq I \text{ is finite} \}$$

implies that

$$|J| \leq \sum \Big\{ |J \cap \langle\!\langle I_0 \rangle\!\rangle_D| \;\Big|\; I_0 \subseteq I \text{ is finite} \Big\} \leq |I|^{<\omega} = |I| . \qquad \square$$

Theorem 1.9. *Let D be a transitive dependence relation on A.*

(a) *For every D-independent set I and every D-spanning set $S \supseteq I$ there exists a D-basis B with $I \subseteq B \subseteq S$.*

(b) *There exists a D-basis and all D-bases have the same cardinality*

Proof. (a) follows from Lemma 1.7.

(b) The existence of a D-basis follows from (a) by setting $I := \varnothing$ and $S := A$. The fact that two bases have the same cardinality follows from Lemma 1.8. $\square$

2. *Matroids and geometries*

It will be convenient to work with closure operators instead of dependence relations.

Definition 2.1. Let Ω be a set.

(a) A *matroid* is a pair $\langle \Omega, \mathrm{cl} \rangle$ where cl is a finitary closure operator on Ω with the exchange property.

(b) A matroid $\langle \Omega, \mathrm{cl} \rangle$ is a *geometry* if it satisfies

$$\mathrm{cl}(\varnothing) = \varnothing \quad \text{and} \quad \mathrm{cl}(\{a\}) = \{a\} , \quad \text{for every } a \in \Omega .$$

(c) Let $\langle \Omega, \mathrm{cl} \rangle$ be a matroid. For $U, I \subseteq \Omega$, we say that I is *independent over U* if

$$a \notin \mathrm{cl}(U \cup (I \setminus \{a\}))\,, \quad \text{for all } a \in I\,.$$

We call *I independent* if it is independent over the empty set.

A *basis* of a set $X \subseteq \Omega$ is an independent set $I \subseteq X$ with $\mathrm{cl}(I) \supseteq X$. The *dimension* of X is the cardinality of any basis of X. We denote it by $\dim_{\mathrm{cl}}(X)$. Similarly, we define a basis of X *over a set U* as a maximal set $I \subseteq X$ that is independent over U. The dimension $\dim_{\mathrm{cl}}(X/U)$ of X *over U* is the cardinality of any such set.

Example. Let $f : A \to B$ be a function and define

$$c(X) := f^{-1}[f[X]]\,, \quad \text{for } X \subseteq A\,.$$

Then $\langle A, c \rangle$ forms a matroid.

Remark. With any matroid $\langle \Omega, \mathrm{cl} \rangle$ we can associate the lattice $\langle \mathrm{fix}\,\mathrm{cl}, \subseteq \rangle$ of all closed sets and the closure space $\langle \Omega, \mathrm{fix}\,\mathrm{cl} \rangle$.

Exercise 2.1. Let $\langle \Omega, \mathrm{cl} \rangle$ be a matroid, $X \subseteq \Omega$, and let $C \subseteq \mathrm{fix}\,\mathrm{cl}$ be a maximal chain of closed sets such that $A \subseteq \mathrm{cl}(X)$, for all $A \in C$. Prove that $|C| = \dim_{\mathrm{cl}}(X) \oplus 1$.

Definition 2.2. Let $\mathfrak{V}$ be a vector space over a skew field $\mathfrak{S}$.

(a) The *linear matroid* associated with $\mathfrak{V}$ is the matroid $\langle V, \mathrm{cl} \rangle$ where $\mathrm{cl}(X) := \langle\!\langle X \rangle\!\rangle_{\mathfrak{V}}$ is the linear subspace spanned by X.

(b) The *affine geometry* associated with $\mathfrak{V}$ is the matroid $\langle V, \mathrm{cl} \rangle$ where

$$\mathrm{cl}(X) := \{\, s_0 x_0 + \dots + s_{n-1} x_{n-1} \mid n < \omega,\ s_i \in S,\ x_i \in X \text{ with } s_0 + \dots + s_{n-1} = 1 \,\}\,.$$

Example. Let $\mathfrak{V}$ be a vector space and let $x, y \in V$ be linearly independent. In the linear matroid the closure of $\{x, y\}$ is the plain through x, y, and the zero vector 0. In the affine geometry the closure of $\{x, y\}$ is the line through x and y.

Remark. (a) The linear matroid is not a geometry since $\text{cl}\,\varnothing = \{0\} \neq \varnothing$. Furthermore, the usual dimension of a linear subspace $U \subseteq V$ coincides with its dimension $\dim_{\text{cl}}(U)$ in the linear matroid as defined above.

(b) The affine geometry $\langle V, \text{cl}\rangle$ associated with a vector space $\mathfrak{V}$ is really a geometry. But note that the usual affine dimension of an affine subspace $U \subseteq V$ is one less than its dimension $\dim_{\text{cl}}(U)$ in the affine geometry as defined above.

The dimension function of a matroid has the following basic properties. In fact, we will show below that every function of this kind arises from a matroid.

Definition 2.3. Let Ω be a set. A function $\dim : \mathscr{P}(\Omega) \times \mathscr{P}(\Omega) \to \mathrm{Cn}$ is a *geometric dimension function* if, for all sets $A, B, U, V \subseteq \Omega$, the following conditions are satisfied:

(1) $\dim(A/U) \leq |A \setminus U|$.

(2) $\dim(A \cup U/U) = \dim(A/U)$.

(3) $A \subseteq B$ and $U \subseteq V$ implies $\dim(A/V) \leq \dim(B/U)$.

(4) If, for some ordinal γ, $(A_\alpha)_{\alpha<\gamma}$ is an increasing chain of sets $A_\alpha \subseteq \Omega$, then

$$\dim(A_{<\gamma}/U) = \sum_{\alpha<\gamma} \dim(A_\alpha/U \cup A_{<\alpha}),$$

where $A_{<\alpha} := \bigcup_{\beta<\alpha} A_\beta$.

(5) For every element $a \in \Omega$ with $\dim(a/U) = 0$, there is a finite subset $U_0 \subseteq U$ such that $\dim(a/U_0) = 0$.

First, let us show that the dimension function of a matroid has these properties.

Proposition 2.4. *The dimension function* $\dim_{\text{cl}}$ *associated with a matroid* $\langle \Omega, \text{cl}\rangle$ *is a geometric dimension function.*

Proof. We have to check five conditions.

(1) If I is a basis of A over U, then $I \subseteq A \setminus U$. Hence, $\dim_{cl}(A/U) = |I| \leq |A \setminus U|$.

(2) Every basis of $A \cup U$ over U is also a basis of A over U.

(3) Every set $I \subseteq A$ that is independent over V is also independent over U. Hence, $|I| \leq \dim_{cl}(B/U)$.

(4) Let I be a basis of U. We define an increasing sequence of sets $(J_\alpha)_{\alpha<\gamma}$ such that J_α is a basis of $U \cup A_\alpha$ with $I \subseteq J_\alpha$. We proceed by induction on $\alpha < \gamma$. Suppose that we have already defined J_β, for all $\beta < \alpha$. Set $J_{<\alpha} := I \cup \bigcup_{\beta<\alpha} J_\beta$. By inductive hypothesis, $J_{<\alpha}$ is a basis of $U \cup A_{<\alpha}$. We can use Theorem 1.9 to extend $J_{<\alpha}$ to a basis J_α of $U \cup A_\alpha$. It follows that $B_\alpha := J_\alpha \setminus J_{<\alpha}$ is a basis of A_α over $U \cup A_{<\alpha}$ and $J_{<\gamma} \setminus I$ is a basis of $A_{<\gamma}$ over U. Hence,

$$\dim_{cl}(A_{<\gamma}/U) = |J_{<\gamma} \setminus I| = \sum_{\alpha<\gamma} |B_\alpha| = \sum_{\alpha<\gamma} \dim_{cl}(A_\alpha/U \cup A_{<\alpha}) \,.$$

(5) If $\dim_{cl}(a/U) = 0$ then $a \in cl(U)$. Since cl has finite character, there is a finite subset $U_0 \subseteq U$ such that $a \in cl(U_0)$. This implies $\dim_{cl}(a/U_0) = 0$. □

Before proving that, conversely, every geometric dimension function arises from a matroid, let us collect some immediate consequences of the definition of a dimension function.

Lemma 2.5. *Let* $\dim : \mathscr{P}(\Omega) \times \mathscr{P}(\Omega) \to \mathrm{Cn}$ *be a geometric dimension function.*

(a) $\dim(A \cup B/U) = \dim(A/U \cup B) \oplus \dim(B/U)$

(b) *If* $(a_\alpha)_{\alpha<\kappa}$ *is an enumeration of A then*

$$\dim(A/U) = \sum_{\alpha<\kappa} \dim(a_\alpha/U \cup A_{<\alpha}) \,,$$

where $A_{<\alpha} := \{ a_\beta \mid \beta < \alpha \}$.

Proof. (a) Considering the two-element increasing sequence $B \subseteq A \cup B$, it follows from the axioms of a geometric dimension function that

$$\begin{aligned}\dim(A \cup B/U) &= \dim(A \cup B/U \cup B) \oplus \dim(B/U)\\ &= \dim\big(A \cup B \cup (U \cup B) \bigm/ U \cup B\big) \oplus \dim(B/U)\\ &= \dim(A/U \cup B) \oplus \dim(B/U)\,.\end{aligned}$$

(b) By (a) and the axioms of a geometric dimension function, we have

$$\begin{aligned}\dim(A/U) &= \sum_{\alpha<\kappa} \dim\big(\{a_\alpha\} \cup A_{<\alpha} \bigm/ U \cup A_{<\alpha}\big)\\ &= \sum_{\alpha<\kappa} \big[\dim(a_\alpha/U \cup A_{<\alpha}) \oplus \dim(A_{<\alpha}/U \cup A_{<\alpha})\big]\\ &= \sum_{\alpha<\kappa} \dim(a_\alpha/X \cup A_{<\alpha}) \oplus 0\,.\end{aligned}$$

□

Proposition 2.6. *Let* $\dim : \wp(\Omega) \times \wp(\Omega) \to \mathrm{Cn}$ *be a geometric dimension function. For* $X \subseteq \Omega$*, we define*

$$\mathrm{cl}(X) := \{\, a \in \Omega \mid \dim(a/X) = 0 \,\}\,.$$

Then $\langle \Omega, \mathrm{cl} \rangle$ *is a matroid such that* $\dim_{\mathrm{cl}} = \dim$.

Proof. First, let us show that cl is a closure operator. Note that, for every $a \in X$, $\dim(a/X) \leq |\{a\} \smallsetminus X| = 0$ implies that $a \in \mathrm{cl}(X)$. Consequently, $X \subseteq \mathrm{cl}(X)$.

For monotonicity, assume that $X \subseteq Y$ and let $a \in \mathrm{cl}(X)$. Then

$$\dim(a/Y) \leq \dim(a/X) = 0 \quad \text{implies} \quad a \in \mathrm{cl}(Y)\,.$$

It remains to show that $\mathrm{cl}(\mathrm{cl}(X)) = X$. Let $a \in \mathrm{cl}(\mathrm{cl}(X))$. Then $\dim(a/\mathrm{cl}(X)) = 0$. Furthermore, $\dim(b/X) = 0$ for each $b \in \mathrm{cl}(X)$. Let $(b_\alpha)_{\alpha<\kappa}$ be an enumeration of $\mathrm{cl}(X)$ and set $B_{<\alpha} := \{\, b_\beta \mid \beta < \alpha \,\}$. Then $B_{<\kappa} = \mathrm{cl}(X)$ and, by Lemma 2.5 (b), it follows that

$$\dim(B_{<\kappa}/X) = \sum_{\alpha<\kappa} \dim(b_\alpha/X \cup B_{<\alpha}) \leq \sum_{\alpha<\kappa} \dim(b_\alpha/X) = 0\,.$$

Consequently, Lemma 2.5 (a) implies

$$\begin{aligned}\dim(a/X) &\leq \dim\big(\mathrm{cl}(X) \cup \{a\} \,/\, X\big) \\ &= \dim(a/\,\mathrm{cl}(X)) \oplus \dim(\mathrm{cl}(X)/X) = 0 \oplus 0\,,\end{aligned}$$

as desired.

We have shown that cl is a closure operator. To prove that it has finite character, suppose that $a \in \mathrm{cl}(X)$. Then $\dim(a/X) = 0$. Hence, there is a finite subset $X_0 \subseteq X$ such that $\dim(a/X_0) = 0$. This implies $a \in \mathrm{cl}(X_0)$.

It remains to check that cl has the exchange property. Suppose that $b \in \mathrm{cl}(U \cup \{a\}) \setminus \mathrm{cl}(U)$. Then $\dim(b/U \cup \{a\}) = 0$. Since $b \notin \mathrm{cl}(U)$, we have $\dim(b/U) = 1$. Hence,

$$\begin{aligned} & \dim(a/U \cup \{b\}) \oplus 1 \\ = {} & \dim(a/U \cup \{b\}) \oplus \dim(b/U) \\ = {} & \dim(ab/U) \\ = {} & \dim(b/U \cup \{a\}) \oplus \dim(a/U) = \dim(a/U) \leq 1 \end{aligned}$$

implies that $\dim(a/U \cup \{b\}) = 0$. Consequently, $a \in \mathrm{cl}(U \cup \{b\})$.

We have shown that $\langle \Omega, \mathrm{cl} \rangle$ is a matroid. To conclude the proof, we must check that $\dim_{\mathrm{cl}} = \dim$. We proceed in two steps. First, we show that $\dim(I/U) = |I|$ for every set I that is cl-independent over U. Let I be such a set. By definition of cl, it follows that

$$\dim(a/U \cup (I \setminus \{a\})) = 1\,, \quad \text{for every } a \in I\,.$$

Set $\kappa := |I|$ and let $(a_\alpha)_{\alpha<\kappa}$ be an enumeration of I. Setting $I_{<\alpha} := \{\, a_\beta \mid \beta < \alpha \,\}$ it follows from Lemma 2.5 (b) that

$$\begin{aligned}\dim(I/U) &= \sum_{\alpha<\kappa} \dim(a_\alpha/U \cup I_{<\alpha}) \\ &\geq \sum_{\alpha<\kappa} \dim(a_\alpha/U \cup (I \setminus \{a_\alpha\})) = \kappa\,.\end{aligned}$$

Therefore, $\dim(I/U) \leq |I \setminus U| \leq \kappa$ implies $\dim(I/U) = \kappa$.

Finally, we prove that $\dim(\mathrm{cl}(X)/U) = \dim(X/U)$, for every set X. Let $(a_\alpha)_{\alpha<\kappa}$ be an enumeration of $\mathrm{cl}(X)$ and set $A_{<\alpha} := \{ a_\beta \mid \beta < \alpha \}$. Then

$$\begin{aligned}
\dim(\mathrm{cl}(X)/U) &= \dim(\mathrm{cl}(X)/X) \oplus \dim(X/U) \\
&= \sum_{\alpha<\kappa} \dim(a_\alpha/X \cup A_{<\alpha}) \oplus \dim(X/U) \\
&\leq \sum_{\alpha<\kappa} \dim(a_\alpha/X) \oplus \dim(X/U) \\
&= 0 \oplus \dim(X/U) .
\end{aligned}$$

To prove that $\dim_{\mathrm{cl}}(X/U) = \dim(X/U)$, let I be a cl-basis of X over U. Then $\dim(I/U) \leq \dim(X/U) \leq \dim(\mathrm{cl}(I)/U) = \dim(I/U)$ implies that

$$\dim_{\mathrm{cl}}(X/U) = |I| = \dim(I/U) = \dim(X/U) . \qquad \square$$

Note that it follows from Proposition 2.6 that a dimension function is uniquely determined by the set of all pairs A, U such that $\dim(A/U) = 0$.

Corollary 2.7. *Let $d, d' : \wp(\Omega)\times\wp(\Omega) \to \mathrm{Cn}$ be two geometric dimension functions. If*

$$d(A/U) = 0 \quad \text{iff} \quad d'(A/U) = 0 , \quad \text{for all } A, U \subseteq \Omega ,$$

then $d = d'$.

Proof. According to Proposition 2.6, we can associate with d and d' matroids $\langle \Omega, c\rangle$ and $\langle \Omega, c'\rangle$, respectively. Since $d(A/U) = 0$ if, and only if, $d'(A/U) = 0$, it follows that $c = c'$. Hence,

$$d = \dim_c = \dim_{c'} = d' . \qquad \square$$

3. *Modular geometries*

There is a general construction turning an arbitrary matroid into a geometry.

Definition 3.1. Let $\langle \Omega, \mathrm{cl} \rangle$ be a matroid and $U \subseteq \Omega$. The *localisation* of $\langle \Omega, \mathrm{cl} \rangle$ at U is the matroid $\langle \Omega, \mathrm{cl} \rangle_{(U)} := \langle \Omega_{(U)}, \mathrm{cl}_{(U)} \rangle$ where

$$\begin{aligned} \Omega_{(U)} &:= \{ \mathrm{cl}(U \cup \{a\}) \mid a \in \Omega \smallsetminus \mathrm{cl}(U) \}, \\ \mathrm{cl}_{(U)}(X) &:= \{ L \in \Omega_{(U)} \mid L \subseteq \mathrm{cl}(U \cup \bigcup X) \}. \end{aligned}$$

Lemma 3.2. *Every localisation of a matroid is a geometry.*

Exercise 3.1. Prove the preceding lemma.

Definition 3.3. Let $\mathfrak{V}$ be a vector space over a skew field $\mathfrak{S}$. The *projective geometry* associated with $\mathfrak{V}$ is the localisation $\langle V, \mathrm{cl} \rangle_{(0)}$ of the linear matroid at the subspace $\{0\}$.

Remark. This coincides with the usual definition of a projective space: the points are the lines $L \subseteq V$ through the origin.

Lemma 3.4. *Let $\langle \Omega, \mathrm{cl} \rangle$ be a matroid, $U, X \subseteq \Omega$ sets, and $\langle \Omega_{(U)}, \mathrm{cl}_{(U)} \rangle$ the localisation at U. Let*

$$X_{(U)} := \{ \mathrm{cl}(U \cup \{x\}) \mid x \in X \smallsetminus \mathrm{cl}(U) \}$$

be the image of X in $\Omega_{(U)}$.

$$\dim_{\mathrm{cl}}(X/U) = \dim_{\mathrm{cl}_{(U)}}(X_{(U)}) .$$

Proof. Let I be a basis of X over U. Then $I \cap \mathrm{cl}(U) = \emptyset$. Hence, if we can show that

$$I_{(U)} := \{ \mathrm{cl}(U \cup \{a\}) \mid a \in I \}$$

is a basis of $X_{(U)}$, then $|I_{(U)}| = |I|$ and the claim follows.

For $x \in X$, let $L_x := \mathrm{cl}(U \cup \{x\})$. To show that $I_{(U)}$ is independent, suppose that there is some $a \in I$ such that

$$\begin{aligned} L_a &\in \mathrm{cl}_{(U)}(I_{(U)} \smallsetminus \{L_a\}) \\ &= \{ L \in \Omega_{(U)} \mid L \subseteq \mathrm{cl}\big(U \cup \bigcup (I_{(U)} \smallsetminus \{L_a\})\big) \}. \end{aligned}$$

Since $a \in L_a$ it follows that

$$a \in \mathrm{cl}\big(U \cup \bigcup(I_{(U)} \smallsetminus \{L_a\})\big) \subseteq \mathrm{cl}(U \cup (I \smallsetminus \{a\}))\,.$$

Hence, I is not independent over U. A contradiction.

It remains to show that $X_{(U)} \subseteq \mathrm{cl}_{(U)}(I_{(U)})$. Let $L_x \in X_{(U)}$. Then

$$U \cup \{x\} \subseteq U \cup X \subseteq \mathrm{cl}(U \cup I) \quad \text{implies} \quad L_x \in \mathrm{cl}_{(U)}(I_{(U)})\,. \qquad \square$$

Some special properties of affine and projective geometries are worth singling out.

Definition 3.5. Let $\langle \Omega, \mathrm{cl} \rangle$ be a matroid.

(a) $\langle \Omega, \mathrm{cl} \rangle$ is *modular* if the lattice $\langle \mathrm{fix}\,\mathrm{cl}, \subseteq \rangle$ of its closed sets is modular. The matroid is *locally modular* if all of its localisations at a single point $a \in \Omega$ are modular.

(b) $\langle \Omega, \mathrm{cl} \rangle$ is *disintegrated* if $\mathrm{cl}(X) = X$, for all $X \subseteq \Omega$.

(c) $\langle \Omega, \mathrm{cl} \rangle$ is *locally finite* if the closure of every finite set is finite.

(d) A *morphism* between matroids is a continuous function between the corresponding closure spaces.

(e) $\langle \Omega, \mathrm{cl} \rangle$ is *homogeneous* if, for every finite set $U \subseteq \Omega$ and all $a, b \in \Omega \smallsetminus \mathrm{cl}(U)$, there is an isomorphism $\pi : \Omega \to \Omega$ with $\pi \restriction \mathrm{cl}(U) = \mathrm{id}$ and $\pi(a) = b$.

We have defined modularity of a matroid in terms of the corresponding lattice of closed sets. The next lemma lists some equivalent conditions on the matroid itself.

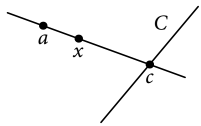
Lemma 3.6. *Let $\langle \Omega, \mathrm{cl} \rangle$ be a matroid. The following statements are equivalent:*

(1) *$\langle \Omega, \mathrm{cl} \rangle$ is modular.*

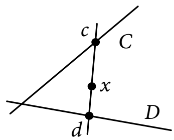
(2) *For all finite $X, Y \subseteq \Omega$, we have*

$$\dim_{\mathrm{cl}}(X \cap Y) + \dim_{\mathrm{cl}}(X \cup Y) = \dim_{\mathrm{cl}}(X) + \dim_{\mathrm{cl}}(Y)\,.$$

(3) *For all closed sets* $C \subseteq \Omega$ *and every pair of elements* $a, x \in \Omega$ *with* $x \in \mathrm{cl}(C \cup \{a\})$, *there exists an element* $c \in C$ *with* $x \in \mathrm{cl}(\{a, c\})$.



(4) *For all closed sets* $C, D \subseteq \Omega$ *and every element* $x \in \mathrm{cl}(C \cup D)$, *there exist elements* $c \in C$ *and* $d \in D$ *with* $x \in \mathrm{cl}(\{c, d\})$.



Proof. (1) ⇒ (2) We have $\dim_{\mathrm{cl}}(X) = \dim_{\mathrm{cl}}(\mathrm{cl}(X))$ and the latter dimension coincides with the height of $\mathrm{cl}(X)$ in the lattice $\langle \mathrm{fix}\,\mathrm{cl}, \subseteq \rangle$. Consequently, the equation follows from the modular law (Theorem B2.5.5).

(2) ⇒ (3) If $a \in C$, we can take $c := x$ and, if $x \in \mathrm{cl}(a)$, we can take an arbitrary $c \in C$. Hence, suppose that $a \notin C \cup \mathrm{cl}(a)$ and choose a finite set $C_0 \subseteq C$ with $x \in \mathrm{cl}(C_0 \cup \{a\})$. Then (2) implies that

$$\begin{aligned}\dim(C_0 \cap \mathrm{cl}(a, x)) &= \dim(C_0) + \dim(a, x) - \dim(C_0 \cup \{a, x\}) \\ &= \dim(C_0) + 2 - (\dim(C_0) + 1) = 1\,.\end{aligned}$$

Hence, there is some $c \in C_0 \cap \mathrm{cl}(a, x)$. By the exchange property it follows that $x \in \mathrm{cl}(a, c)$, as desired.

(3) ⇒ (4) Since cl has finite character, there are finite sets $C_0 \subseteq C$ and $D_0 \subseteq D$ such that $x \in \mathrm{cl}(C_0 \cup D_0)$. We prove the claim by induction on $|C_0|$. If $C_0 = \varnothing$ then $x \in \mathrm{cl}(D_0) \subseteq D$ and we are done. Suppose that $C_0 = A \cup \{a\}$. Since $x \in \mathrm{cl}(A \cup D_0 \cup \{a\})$, we can use (3) to find some $b \in \mathrm{cl}(A \cup D_0)$ with $x \in \mathrm{cl}(\{a, b\})$. By inductive hypothesis, there are $a' \in A$ and $d \in D_0$ such that $b \in \mathrm{cl}(\{a', d\})$. Hence, $x \in \mathrm{cl}(\{a, a', d\})$ and, applying (3) again, we can find some $c \in \mathrm{cl}(\{a, a'\}) \subseteq C$ with $x \in \mathrm{cl}(\{c, d\})$.

(4) ⇒ (1) Let $A, B, C \subseteq \Omega$ be closed sets with $A \subseteq B$. We have to show that $\mathrm{cl}(A \cup (B \cap C)) = B \cap \mathrm{cl}(A \cup C)$. According to Lemma B2.2.6, one inclusion holds in every lattice. For the other one, let $x \in B \cap \mathrm{cl}(A \cup C)$. By (4) there are elements $a \in A$ and $c \in C$ with $x \in B \cap \mathrm{cl}(\{a, c\})$. If $x \in \mathrm{cl}(a)$ then $x \in A$ and we are done. Hence, suppose that $x \notin \mathrm{cl}(a)$. By the exchange property, it then follows that $c \in \mathrm{cl}(\{a, x\}) \subseteq \mathrm{cl}(A \cup B) = B$. Hence, $c \in B \cap C$ and $x \in \mathrm{cl}(\{a, c\}) \subseteq \mathrm{cl}(A \cup (B \cap C))$. ☐

Disintegrated, projective, and affine geometries frequently appear in model theory. The next lemma lists some of their properties.

Lemma 3.7. *Disintegrated geometries and projective geometries are modular and homogeneous. Affine geometries are locally modular and homogeneous, but not modular if the dimension is at least* 3.

Proof. To show that a disintegrated geometry $\langle \Omega, \mathrm{cl} \rangle$ is modular, one only has to check that

$$X \subseteq Y \quad \text{implies} \quad X \cup (Y \cap Z) = Y \cap (X \cup Z)\,.$$

To show that it is homogeneous, let $U \subseteq \Omega$ and $a, b \in \Omega \setminus U$. The bijection $h : \Omega \to \Omega$ exchanging a and b and fixing every other element of Ω is continuous.

Suppose that $\langle \Omega, \mathrm{cl} \rangle$ is the projective geometry associated with a vector space $\mathfrak{V}$. Modularity follows from Lemma B6.4.5. For homogeneity, let $U \subseteq \Omega$ be finite and $a, b \notin \mathrm{cl}(U)$. Let $\langle V, \mathrm{cl}_\wedge \rangle$ be the corresponding linear matroid. For every element $x \in \Omega$ there is a non-zero vector $\hat{x} \in V$ such that $x = \mathrm{cl}_\wedge(\hat{x})$. Fix a basis B of $\hat{U} := \mathrm{cl}_\wedge(\{\, \hat{x} \mid x \in U \,\})$. Since $\hat{a}, \hat{b} \notin \hat{U}$, there exists a linear map $h : V \to V$ fixing B and interchanging $\hat{a}$ and $\hat{b}$. The function $\Omega \to \Omega$ induced by h is the desired continuous mapping.

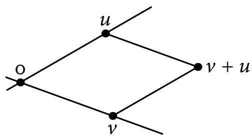
Suppose that $\langle \Omega, \mathrm{cl} \rangle$ is the affine geometry associated with a vector space $\mathfrak{V}$ and let $a \in \Omega$. Then $\langle \Omega, \mathrm{cl} \rangle_{(a)} \cong \langle \Omega, \mathrm{cl} \rangle_{(0)}$ and the latter geometry is isomorphic to the projective geometry associated with $\mathfrak{V}$. Since we have just seen that such geometries are modular, it follows that $\langle \Omega, \mathrm{cl} \rangle$ is locally modular.

To show that it is not modular let $u, v \in V$ be linearly independent vectors. Then $\mathrm{cl}(0) \subseteq \mathrm{cl}(0, u)$ but

$$\begin{aligned} \mathrm{cl}\big(\mathrm{cl}(0) \cup (\mathrm{cl}(0,u) \cap \mathrm{cl}(v, v+u))\big) &= \mathrm{cl}(\mathrm{cl}(0) \cup \varnothing) \\ &= \mathrm{cl}(0)\,, \end{aligned}$$

and

$$\begin{aligned} \mathrm{cl}(0,u) \cap \mathrm{cl}\big(\mathrm{cl}(0) \cup \mathrm{cl}(v, v+u)\big) &= \mathrm{cl}(0,u) \cap \mathrm{cl}(0,u,v) \\ &= \mathrm{cl}(0,u)\,. \end{aligned}$$



For homogeneity, let $U \subseteq \Omega$ be finite and $a, b \notin \mathrm{cl}(U)$ distinct elements. If $U = \varnothing$ and a and b are both non-zero, we can take some linear map $h : V \to V$ interchanging a and b. This map is continuous.

If $U = \varnothing$ and $a = 0$, we first apply a translation f that maps both a and b to non-zero vectors. Then we can use a linear map h as above. The composition $f^{-1} \circ h \circ f$ is the desired continuous map.

Note that there is one case where such a translation f does not exists. If $\mathfrak{V}$ has only two elements. Then $V = \{a, b\}$ and the function interchanging a and b is continuous.

It remains to consider the case that $U \neq \varnothing$. Fix some $x \in U$. By applying a suitable translation f, we can assume that $x = 0 \in U$. Hence, $\mathrm{cl}(U)$ is a linear subspace of $\mathfrak{V}$. Let B be a basis of $\mathrm{cl}(U)$ and let $h : V \to V$ be a linear map fixing B and interchanging a and b. Then $f^{-1} \circ h \circ f$ is the desired continuous map. □

Algebraically closed fields provide examples of geometries that are not locally modular.

Proposition 3.8. *Let $\mathfrak{K}$ be an algebraically closed field of infinite transcendence degree and let $\langle K, \mathrm{cl}\rangle$ be the matroid where* cl *maps a set $X \subseteq K$ to its algebraic closure.*

(a) $\langle K, \mathrm{cl}\rangle$ *is homogeneous.*

(b) *No localisation of* $\langle K, \mathrm{cl}\rangle$ *at a finite set is modular.*

Proof. (a) follows by Corollary B6.5.31.

(b) We consider the localisation $\langle K, \mathrm{cl}\rangle_{(U)}$ at a finite set $U \subseteq K$. Let $n := \dim_{\mathrm{cl}}(U)$. Since $\mathfrak{K}$ has infinite transcendence degree, there are elements a, b, c, d that are algebraically independent over U. Set $x := (a-c)/(b-d)$ and $y := a - bx$, and let

$$A := \mathrm{cl}(a, b, U) \quad \text{and} \quad B := \mathrm{cl}(x, y, U)\,.$$

Then $\mathrm{cl}(A \cup B) = \mathrm{cl}(a, b, x, U)$ has dimension $n+3$, while A and B both have dimension $n+2$. To show that $\langle K, \mathrm{cl}\rangle_{(U)}$ is not modular it is sufficient to prove that the dimension of $A \cap B$ is different from $n+1$.

In fact, we claim that $A \cap B = \mathrm{cl}(U)$ and, hence, the dimension is n. Clearly, we have $U \subseteq A \cap B$. Conversely, consider an element $z \in A \cap B$. By (a), there exists an automorphism $\pi \in \mathrm{Aut}\,\mathfrak{K}$ that fixes B pointwise and maps a tø c. It follows that $\pi(b) = \pi((a-y)/x) = (c-y)/x = d$. Consequently, $z \in B$ implies $\pi(z) = z$, and $z \in A = \mathrm{cl}(a, b, U)$ implies $z = \pi(z) \in \mathrm{cl}(c, d, U)$. Thus,

$$z \in \mathrm{cl}(a, b, U) \cap \mathrm{cl}(c, d, U) = \mathrm{cl}(U)\,. \qquad \square$$

We conclude this section with the following characterisation of homogeneous, locally finite geometries.

Theorem 3.9 (Cherlin, Mills, Zil'ber). *Let* $\langle \Omega, \mathrm{cl}\rangle$ *be a homogeneous, locally finite geometry of infinite dimension. Then exactly one of the following cases holds:*

(1) $\langle \Omega, \mathrm{cl}\rangle$ *is disintegrated.*

(2) $\langle \Omega, \mathrm{cl}\rangle$ *is isomorphic to a projective geometry over a finite field.*

(3) $\langle \Omega, \mathrm{cl}\rangle$ *is isomorphic to an affine geometry over a finite field.*

4. Strongly minimal sets

Having introduced geometries we are interested in first-order theories where the algebraic closure operator forms such a geometry.

Definition 4.1. Let $\mathfrak{M}$ be a structure and $S \subseteq M^n$ an infinite M-definable relation.

(a) We call S *minimal* if, for every M-definable subset $X \subseteq S$, either X, or $S \smallsetminus X$ is finite. A formula $\varphi(\bar{x}; \bar{c})$ with $\bar{c} \subseteq M$ is *minimal* if the relation $\varphi(\bar{x}; \bar{c})^{\mathfrak{M}}$ it defines is minimal.

(b) A relation S, or a formula $\varphi(\bar{x}; \bar{c})$, is *strongly minimal,* if it is minimal in every elementary extension of $\mathfrak{M}$.

Example. (a) Let $\mathfrak{E} = \langle E, \sim \rangle$ be a structure where $\sim$ is an equivalence relation with infinitely many classes each of which is infinite. For every $a \in E$, the formula $x \sim a$ is strongly minimal.

(b) Let $\mathfrak{K}$ be an algebraically closed field. Every definable set $X \subseteq K$ is a boolean combination of solution sets of polynomials. Hence, every such set is either finite or cofinite. Therefore, K is strongly minimal.

(c) In $\mathfrak{A} = \langle \omega, \leq \rangle$ the set ω is minimal, but not strongly minimal since, in every elementary extension $\mathfrak{B} \succ \mathfrak{A}$ we can pick an element $c \in B \smallsetminus \omega$ such that $(x \leq c)^{\mathfrak{B}}$ and $(x > c)^{\mathfrak{B}}$ are both infinite.

We are mainly interested in strongly minimal relations. As the next lemma shows, we can find such a relation by looking for a minimal relation in an $\aleph_0$-saturated structure.

Lemma 4.2. *Every minimal relation in an $\aleph_0$-saturated structure $\mathfrak{M}$ is strongly minimal.*

Proof. Let $\varphi(\bar{x}; \bar{c})$ be a minimal formula with parameters $\bar{c} \subseteq M$. To show that φ is strongly minimal we consider an elementary extension $\mathfrak{N} \succeq \mathfrak{M}$ and a formula $\psi(\bar{x}; \bar{d})$ with parameters $\bar{d} \subseteq N$. For a contradiction, suppose that both sets

$$\varphi(\bar{x}; \bar{c})^{\mathfrak{N}} \cap \psi(\bar{x}; \bar{d})^{\mathfrak{N}} \quad \text{and} \quad \varphi(\bar{x}; \bar{c})^{\mathfrak{N}} \smallsetminus \psi(\bar{x}; \bar{d})^{\mathfrak{N}}$$

are infinite.

Since $\mathfrak{M}$ is $\aleph_0$-saturated we can find a tuple $\bar{d}' \subseteq M$ with $\mathrm{tp}(\bar{d}'/\bar{c}) = \mathrm{tp}(\bar{d}/\bar{c})$. For every $n < \omega$, we have

$$\mathfrak{N} \vDash \exists^n \bar{x}[\varphi(\bar{x};\bar{c}) \wedge \psi(\bar{x};\bar{d})] \wedge \exists^n \bar{x}[\varphi(\bar{x};\bar{c}) \wedge \neg\psi(\bar{x};\bar{d})]$$

which implies that

$$\mathfrak{N} \vDash \exists^n \bar{x}[\varphi(\bar{x};\bar{c}) \wedge \psi(\bar{x};\bar{d}')] \wedge \exists^n \bar{x}[\varphi(\bar{x};\bar{c}) \wedge \neg\psi(\bar{x};\bar{d}')] .$$

It follows that all these formulae also hold in $\mathfrak{M}$. Consequently, both sets $\varphi(\bar{x};\bar{c})^{\mathfrak{M}} \cap \psi(\bar{x};\bar{d}')^{\mathfrak{M}}$ and $\varphi(\bar{x};\bar{c})^{\mathfrak{M}} \setminus \psi(\bar{x};\bar{d}')^{\mathfrak{M}}$ are infinite. A contradiction. ☐

The reason for studying strongly minimal sets is the fact that the algebraic closure operator has the exchange property for these sets.

Lemma 4.3. *Let $\mathfrak{M}$ be a structure and $S \subseteq M^n$ a minimal set. The restriction of* acl *to S forms a matroid.*

Proof. We have already seen in Lemma E2.1.2 that acl is a finitary closure operator. Hence, it remains to check that it has the exchange property.

Suppose that $\bar{a} \subseteq \mathrm{acl}(U \cup \bar{b}) \setminus \mathrm{acl}(U)$ for $\bar{a}, \bar{b} \in S$. We have to show that $\bar{b} \subseteq \mathrm{acl}(U \cup \bar{a})$. There exists a formula $\varphi(\bar{x};\bar{y})$ over U such that $\varphi^{\mathfrak{M}}(\bar{x};\bar{b})$ is a finite set containing $\bar{a}$. Set $m := |\varphi^{\mathfrak{M}}(\bar{x};\bar{b})|$ and let $\psi(\bar{y})$ be the formula stating that there are exactly m tuples $\bar{x} \in S$ such that $\mathfrak{M} \vDash \varphi(\bar{x};\bar{y})$. If $\psi^{\mathfrak{M}}(\bar{y})$ is finite, $\mathfrak{M} \vDash \psi(\bar{b})$ implies that $\bar{b} \subseteq \mathrm{acl}(U)$. Consequently, we have $\bar{a} \subseteq \mathrm{acl}(U)$. A contradiction.

Hence, the set $\psi^{\mathfrak{M}}(\bar{y})$ is infinite. If $(\varphi(\bar{a};\bar{y}) \wedge \psi(\bar{y}))^{\mathfrak{M}}$ is finite then $\bar{b} \subseteq \mathrm{acl}(U \cup \bar{a})$ and we are done. For a contradiction, suppose that this set if infinite. Since S is minimal it follows that the complement $S \cap \neg(\varphi(\bar{a};\bar{y}) \wedge \psi(\bar{y}))^{\mathfrak{M}}$ is finite. Let $k < \aleph_0$ be its cardinality and let $\vartheta(\bar{x})$ be the formula stating that there are exactly k elements $\bar{y} \in S$ that do not satisfy $\varphi(\bar{x};\bar{y}) \wedge \psi(\bar{y})$. If $\vartheta(\bar{x})^{\mathfrak{M}}$ is finite then $\bar{a} \subseteq \mathrm{acl}(U)$. A contradiction.

Hence, $\vartheta(\bar{x})^{\mathfrak{M}}$ is infinite and we can choose $m+1$ distinct elements $\bar{a}_0, \dots, \bar{a}_m \in \vartheta(\bar{x})^{\mathfrak{M}}$. The set

$$B := \bigcap_{i \leq m} [\varphi(\bar{a}_i; \bar{y}) \wedge \psi(\bar{y})]^{\mathfrak{M}}$$

is a finite intersection of cofinite sets and, therefore, cofinite itself. In particular, it is nonempty and we can find some element $\bar{b}^* \in B$. It follows that

$$\mathfrak{M} \vDash \varphi(\bar{a}_i; \bar{b}^*), \quad \text{for all } i \leq m\,.$$

Consequently, $|\varphi^{\mathfrak{M}}(\bar{x}; \bar{b}^*)| > m$. But this implies that $\mathfrak{M} \nvDash \psi(\bar{b}^*)$. A contradiction. □

The geometry of a strongly minimal relation is closely related to its logical properties. For instance, we shall show below that all independent sets are totally indiscernible with the same type. But first, let us collect some technical properties of strongly minimal relations.

Lemma 4.4. *Let $\varphi(\bar{x}; \bar{c})$ be a strongly minimal formula with parameters $\bar{c}$. Let $\bar{s}$ be the sorts of the variables $\bar{x}$.*

(a) *If $\bar{d}$ is a tuple with* $\mathrm{tp}(\bar{d}) = \mathrm{tp}(\bar{c})$ *then $\varphi(\bar{x}; \bar{d})$ is also strongly minimal.*

(b) *For every set $U \supseteq \bar{c}$, there exists a unique nonalgebraic type $\mathfrak{p} \in S^{\bar{s}}(U)$ with $\varphi \in \mathfrak{p}$.*

Proof. (a) For every formula $\psi(\bar{x}; \bar{a})$ with parameters $\bar{a} \subseteq \mathbb{M}$, we have to show that exactly one of

$$(\varphi(\bar{x}; \bar{d}) \wedge \psi(\bar{x}; \bar{a}))^{\mathbb{M}} \quad \text{and} \quad (\varphi(\bar{x}; \bar{d}) \wedge \neg\psi(\bar{x}; \bar{a}))^{\mathbb{M}}$$

is finite. Since $\mathrm{tp}(\bar{d}) = \mathrm{tp}(\bar{c})$ there is an automorphism π of $\mathbb{M}$ with $\pi(\bar{d}) = \bar{c}$. Let $\bar{b} := \pi(\bar{a})$. As $\varphi(\bar{x}; \bar{c})$ is strongly minimal, exactly one of

$$(\varphi(\bar{x}; \bar{c}) \wedge \psi(\bar{x}; \bar{b}))^{\mathbb{M}} \quad \text{and} \quad (\varphi(\bar{x}; \bar{c}) \wedge \neg\psi(\bar{x}; \bar{b}))^{\mathbb{M}}$$

is finite. Since

$$\pi\big[(\varphi(\bar{x};\bar{d}) \wedge \psi(\bar{x};\bar{a}))^{\mathbb{M}}\big] = (\varphi(\bar{x};\bar{c}) \wedge \psi(\bar{x};\bar{b}))^{\mathbb{M}},$$
$$\text{and}\quad \pi\big[(\varphi(\bar{x};\bar{d}) \wedge \neg\psi(\bar{x};\bar{a}))^{\mathbb{M}}\big] = (\varphi(\bar{x};\bar{c}) \wedge \neg\psi(\bar{x};\bar{b}))^{\mathbb{M}},$$

the claim follows.

(b) Let $\mathfrak{M}$ be an $\aleph_0$-saturated model containing U and set

$$\mathfrak{p} := \{\, \psi \mid \psi \text{ a formula over } U \text{ such that } (\varphi \wedge \psi)^{\mathfrak{M}} \text{ is infinite} \,\}.$$

Since φ is strongly minimal, it follows that

$$\psi \in \mathfrak{p} \quad \text{iff} \quad \neg\psi \notin \mathfrak{p}, \quad \text{for every formula } \psi \text{ over U}.$$

Hence, $\mathfrak{p}$ is a complete type over U containing φ. Clearly, $\mathfrak{p}$ is nonalgebraic since, if there were some algebraic formula $\psi \in \mathfrak{p}$, then $\varphi \wedge \psi$ were also algebraic, in contradiction to the definition of $\mathfrak{p}$.

Suppose that $\mathfrak{q} \in S^{\bar{s}}(U)$ is another nonalgebraic type containing φ. To show that $\mathfrak{q} \subseteq \mathfrak{p}$, consider $\psi \in \mathfrak{q}$. Then $\varphi \wedge \psi \in \mathfrak{q}$ and, by assumption, this formula is nonalgebraic. By definition of $\mathfrak{p}$ it follows that $\psi \in \mathfrak{p}$. □

Lemma 4.5. *Let $\varphi(\bar{x})$ be a strongly minimal formula over a set U of parameters. Let $\bar{s}$ be the sorts of the variables $\bar{x}$, and let $\mathfrak{p} \in S^{\bar{s}}(U)$ be the unique nonalgebraic type containing φ.*

(a) *$\mathfrak{p}$ is isolated if, and only if, $\varphi^{\mathbb{M}}$ contains only finitely many tuples in* $\mathrm{acl}(U)$.

(b) *Let $V \supseteq U$ and let $\mathfrak{q} \in S^{\bar{s}}(V)$ be the unique nonalgebraic extension of $\mathfrak{p}$. If $\mathfrak{p}$ is isolated, so is $\mathfrak{q}$.*

Proof. (a) Let $R := \{\, \bar{a} \in \varphi^{\mathbb{M}} \mid \bar{a} \subseteq \mathrm{acl}(U) \,\}$. For $(\Leftarrow)$, suppose that $R = \{\bar{a}_0, \dots, \bar{a}_{n-1}\}$ is finite. For each $i < n$, we fix an algebraic formula ψ_i over U such that $\mathbb{M} \vDash \psi_i(\bar{a}_i)$. It follows that $\psi := \bigvee_{i<n}[\psi_i \wedge \varphi]$ is a formula over U defining R. We claim that $\varphi \wedge \neg\psi$ isolates $\mathfrak{p}$.

Since $\mathfrak{p}$ is nonalgebraic, we have $\psi \notin \mathfrak{p}$. Therefore, $\varphi \wedge \neg\psi \in \mathfrak{p}$. Conversely, let $\mathfrak{q}$ be an arbitrary complete type over U containing $\varphi \wedge \neg\psi$. If

$\mathfrak{q}$ is nonalgebraic, it coincides with $\mathfrak{p}$, by Lemma 4.4 (b), and we are done. Therefore, we may assume that $\mathfrak{q}$ contains an algebraic formula ϑ. Then each of the finitely many realisations of $\mathfrak{q}$ is in $\operatorname{acl}(U)$. Consequently, $\mathfrak{q}^{\mathbb{M}} \subseteq R$, which implies that $\psi \in \mathfrak{q}$. A contradiction.

($\Rightarrow$) For a contradiction, suppose that there is some $\psi(\bar{x}) \in \mathfrak{p}$ isolating $\mathfrak{p}$, but R is infinite. Let Γ be the set of all algebraic formulae over U. As $\mathfrak{p}$ is the unique nonalgebraic type in $S^{\bar{s}}(U)$ containing φ, the set

$$\{\varphi \wedge \neg\psi\} \cup \{\, \neg\vartheta \mid \vartheta \in \Gamma\,\}$$

is inconsistent. Hence, there are finitely many formula $\vartheta_0, \dots, \vartheta_{n-1} \in \Gamma$ such that

$$T(U) \cup \{\varphi, \neg\vartheta_0, \dots, \neg\vartheta_{n-1}\} \vDash \psi \,.$$

Since R is infinite and all ϑ_i are algebraic, there is some element

$$\bar{a} \in R \setminus (\vartheta_0^{\mathbb{M}} \cup \dots \cup \vartheta_{n-1}^{\mathbb{M}}) \subseteq (\varphi \wedge \neg\vartheta_0 \wedge \dots \wedge \neg\vartheta_{n-1})^{\mathbb{M}} \subseteq \psi^{\mathbb{M}} \,.$$

But $\operatorname{tp}(\bar{a}/U) \neq \mathfrak{p}$ since the former type is algebraic, while the latter one is not. Consequently, ψ does not isolate $\mathfrak{p}$. A contradiction.

(b) follows immediately from (a). □

Proposition 4.6. *Let $\mathfrak{M}$ be a structure, $U \subseteq M$, and suppose that $S \subseteq M^k$ a U-definable minimal relation. If $\bar{a}, \bar{b} \in S^n$ are finite tuples each of which is independent over U, then*

$$\operatorname{tp}(\bar{a}/U) = \operatorname{tp}(\bar{b}/U) \,.$$

Proof. We prove the claim by induction on n. For $n = 0$ there is nothing to do. Suppose that we have already proved the claim for n-tuples and let $\bar{a}c \in S^{n+1}$ and $\bar{b}d \in S^{n+1}$ be both independent over U. By inductive hypothesis, we have $\operatorname{tp}(\bar{a}/U) = \operatorname{tp}(\bar{b}/U)$. Let $\psi(\bar{x}, y)$ be a formula over U such that

$$\mathfrak{M} \vDash \psi(\bar{a}, c) \,.$$

Since $c \notin \operatorname{acl}(U \cup \bar{a})$ it follows that the set $S \cap \psi(\bar{a}, y)^{\mathfrak{M}}$ is infinite and its complement $S \smallsetminus \psi(\bar{a}, y)^{\mathfrak{M}}$ is finite. Furthermore, $\operatorname{tp}(\bar{a}/U) = \operatorname{tp}(\bar{b}/U)$ implies that

$$\left|S \smallsetminus \psi(\bar{b}, y)^{\mathfrak{M}}\right| = \left|S \smallsetminus \psi(\bar{a}, y)^{\mathfrak{M}}\right| < \aleph_0 .$$

Hence, $d \notin \operatorname{acl}(U \cup \bar{b})$ implies that $\mathfrak{M} \vDash \psi(\bar{b}, d)$. □

Corollary 4.7. *Let $\mathfrak{M}$ be a structure, $U \subseteq M$ a set of parameters, and $S \subseteq M$ a U-definable minimal set. Every U-independent set $A \subseteq S$ is totally indiscernible over U.*

Proof. Let $\bar{a}, \bar{b} \in [A]^n$. Then $\bar{a}$ and $\bar{b}$ are U-independent and, therefore, they have the same type over U. □

We have seen that we can use geometric methods to study models containing minimal sets. Let us turn to prove the existence of minimal sets.

Lemma 4.8. *Let T be a $\aleph_0$-stable theory over a countable signature Σ, $\mathfrak{M} \vDash T$ infinite, $\vartheta(\bar{x})$ a formula over M, and let $\kappa \leq |\vartheta^{\mathfrak{M}}|$ be an infinite cardinal. There exists a formula $\varphi(\bar{x})$ over M such that $\varphi^{\mathfrak{M}} \subseteq \vartheta^{\mathfrak{M}}$, $\left|\varphi^{\mathfrak{M}}\right| \geq \kappa$ and, for every formula $\psi(\bar{x})$ over M, we either have*

$$\left|(\varphi \wedge \psi)^{\mathfrak{M}}\right| < \kappa \quad \text{or} \quad \left|(\varphi \wedge \neg\psi)^{\mathfrak{M}}\right| < \kappa .$$

Proof. For a contradiction, suppose that there is no such φ. We construct a family $(\varphi_w)_{w \in 2^{<\omega}}$ of formulae over M such that, for all $w \in 2^{<\omega}$, we have

$$\varphi_w^{\mathfrak{M}} \subseteq \vartheta^{\mathfrak{M}}, \quad \left|\varphi_w^{\mathfrak{M}}\right| \geq \kappa \quad \text{and} \quad \varphi_{w0}^{\mathfrak{M}} \cap \varphi_{w1}^{\mathfrak{M}} = \emptyset .$$

We start with $\varphi_{\langle\rangle} := \vartheta$. Then $\varphi_{\langle\rangle}^{\mathfrak{M}} = \vartheta^{\mathfrak{M}}$ and $\left|\varphi_{\langle\rangle}^{\mathfrak{M}}\right| \geq \kappa$. For the inductive step, suppose that we have already defined φ_w. By assumption, there is a formula ψ over M such that

$$\left|(\varphi_w \wedge \psi)^{\mathfrak{M}}\right| \geq \kappa \quad \text{and} \quad \left|(\varphi_w \wedge \neg\psi)^{\mathfrak{M}}\right| \geq \kappa .$$

We set $\varphi_{w0} := \varphi_w \wedge \psi$ and $\varphi_{w1} := \varphi_w \wedge \neg\psi$.

Having defined $(\varphi_w)_w$, let $U \subseteq M$ be the set of all parameters appearing in some φ_w. Then U is countable and the family $(\varphi_w)_{w \in 2^{<\omega}}$ is an embedding of $2^{<\omega}$ into $\mathrm{FO}^{\bar{s}}[\Sigma_U]/T$, where $\bar{s}$ are the sorts of $\bar{x}$. By Lemma B5.7.3, it follows that $|S^{\bar{s}}(U)| > \aleph_0$. A contradiction to $\aleph_0$-stability. ☐

Corollary 4.9. *Let T be a $\aleph_0$-stable theory over a countable signature Σ. Every infinite model of T contains a minimal relation.*

Proof. This follows from the preceding lemma for $\vartheta(x) :=$ true and $\kappa := \aleph_0$. ☐

We can use minimal sets to define isomorphisms between models.

Lemma 4.10. *Every elementary function $f_0 : A \to B$ can be extended to a elementary function $f : \mathrm{acl}(A) \to \mathrm{acl}(\mathrm{rng}\, f_0)$ that is bijective.*

Proof. W.l.o.g. we may assume that $B = \mathrm{rng}\, f_0$. Let F be the set of all elementary functions $g : C \to D$ such that $A \subseteq C \subseteq \mathrm{acl}(A)$ and $g \restriction A = f_0$. Then $\langle F, \subseteq \rangle$ is inductively ordered. Hence, it has a maximal element $f : C \to D$. We claim that f is the desired function.

First of all, every elementary function is injective. For surjectivity, suppose that $b \in \mathrm{acl}(B) \setminus D$. Since $b \in \mathrm{acl}(D)$, we can use Lemma E3.1.3 to find a formula $\varphi(x; \bar{d})$ with parameters $\bar{d} \subseteq D$ isolating $\mathrm{tp}(b/D)$. Since $\mathrm{tp}(b/D)$ is algebraic, φ must be an algebraic formula. Fixing $\bar{c} \subseteq C$ such that $f(\bar{c}) = \bar{d}$ it follows that

$$f[\varphi(x; \bar{c})^{\mathbb{M}}] \subseteq \varphi(x; \bar{d})^{\mathbb{M}} \quad \text{and} \quad \big|\varphi(x; \bar{c})^{\mathbb{M}}\big| = \big|\varphi(x; \bar{d})^{\mathbb{M}}\big|.$$

Consequently, there exists some element $a \in \varphi(x; \bar{c})^{\mathbb{M}} \setminus C$. Furthermore, $\varphi(x; \bar{c})$ isolates $\mathrm{tp}(a/C)$. Hence, $f[\mathrm{tp}(a/C)] = \mathrm{tp}(b/D)$ and we have $f \cup \{\langle a, b \rangle\} \in F$. This contradicts the maximality of f.

It remains to prove that $C = \mathrm{acl}(A)$. Suppose that there exists an element $a \in \mathrm{acl}(A) \setminus C$. Then $\mathrm{tp}(a/C)$ is isolated and, as above, we can find an element b such that $f \cup \{\langle a, b \rangle\} \in F$. Again a contradiction. ☐

Corollary 4.11. *Let T be a theory, $\varphi(x)$ a strongly minimal formula, and $\mathfrak{A}$ and $\mathfrak{B}$ models of T. If $\dim(\varphi^{\mathfrak{A}}) = \dim(\varphi^{\mathfrak{B}})$, there exists a bijective elementary map $f : \mathrm{acl}(\varphi^{\mathfrak{A}}) \to \mathrm{acl}(\varphi^{\mathfrak{B}})$.*

Proof. Fix bases I and J of, respectively, $\varphi^{\mathfrak{A}}$ and $\varphi^{\mathfrak{B}}$. By assumption, $|I| = |J|$. Let $f_0 : I \to J$ be a bijection. By Corollary 4.6, it follows that f_0 is elementary. Hence, we can use Lemma 4.10 to extend f_0 to an elementary map $f : \mathrm{acl}(I) \to \mathrm{acl}(J)$. Since $\mathrm{acl}(I) = \mathrm{acl}(\varphi^{\mathfrak{A}})$ and $\mathrm{acl}(J) = \mathrm{acl}(\varphi^{\mathfrak{B}})$, this is the desired map. □

We can apply the results on minimal sets to study theories where every model consists of a minimal set. In fact, it is sufficient that every model is generated by a minimal set.

Definition 4.12. Let T be a complete first-order theory.
(a) T is *strongly minimal* if the formula $x = x$ is strongly minimal.
(b) T is *almost strongly minimal* if there exists a strongly minimal formula $\varphi(x;\bar{c})$ with parameters $\bar{c}$ such that $\mathrm{tp}(\bar{c})$ is isolated and

$$\mathrm{acl}(\varphi^{\mathfrak{M}} \cup \bar{c}) = M\,, \quad \text{for every model } \mathfrak{M} \text{ of } T(\bar{c})\,.$$

Example. The theories DAG and ACF_p are strongly minimal.

Theorem 4.13. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be models of an almost strongly minimal theory T and let $\varphi(x;\bar{c})$ be the corresponding strongly minimal formula. Then*

$$\mathfrak{A} \cong \mathfrak{B} \quad \text{iff} \quad \dim(\varphi^{\mathfrak{A}}/\bar{c}) = \dim(\varphi^{\mathfrak{B}}/\bar{c})\,.$$

Proof. $(\Rightarrow)$ is trivial and $(\Leftarrow)$ follows from Corollary 4.11. □

Corollary 4.14. *Every almost strongly minimal theory T is κ-categorical, for all $\kappa > |T|$.*

Proof. Let $\varphi(x;\bar{c})$ be the strongly minimal formula associated with T and let $\mathfrak{A}$ and $\mathfrak{B}$ be models of T of the same size $|A| = |B| > |T|$. Since

$\mathrm{tp}(\bar c)$ is isolated, there are tuples $\bar a \subseteq A$ and $\bar b \subseteq B$ realising $\mathrm{tp}(\bar c)$. Fix bases $I \subseteq A$ and $J \subseteq B$ of $\varphi^{\mathfrak A}$ over $\bar a$ and of $\varphi^{\mathfrak B}$ over $\bar b$, respectively. Then

$$\dim(\varphi^{\mathfrak A}/\bar a) = |I| = |\mathrm{acl}(I)| = |A| \quad \text{and} \quad \dim(\varphi^{\mathfrak B}/\bar b) = \dots = |B|\,.$$

By Theorem 4.13, it follows that $\mathfrak A \cong \mathfrak B$. □

5. *Vaughtian pairs and the Theorem of Morley*

In this section we shall prove the Theorem of Morley which states that a countable first-order theory T that is κ-categorical, for *some* uncountable cardinal κ, is λ-categorical, for *all* uncountable cardinals λ. We have already seen in Theorem E6.3.16 that such a theory is necessarily $\aleph_0$-stable. It follows that every uncountable model is saturated. Note that, according to Lemma E1.2.17, we have $|\varphi^{\mathfrak M}| < \aleph_0$ or $|\varphi^{\mathfrak M}| = |M|$, for every saturated model $\mathfrak M$ of T and every formula φ. In fact, we will show below that a $\aleph_0$-stable theory T is uncountably categorical if, and only if, this property holds for *all* uncountable models $\mathfrak M$.

Definition 5.1. Let T be a first-order theory.

(a) A *Vaughtian pair* for T consists of two models $\mathfrak A \prec \mathfrak B$ of T such that, for some formula $\varphi(\bar x)$ over A, $\varphi^{\mathfrak A}$ is infinite and $\varphi^{\mathfrak A} = \varphi^{\mathfrak B}$.

(b) The *size* of a Vaughtian pair $\langle \mathfrak A, \mathfrak B\rangle$ is the tuple $\langle \kappa, \lambda\rangle$ where $\kappa := |A|$ and $\lambda := |B|$.

(c) If $\mathfrak A \preceq \mathfrak B$ are structures, we denote by $\langle \mathfrak B, A\rangle$ the expansion of $\mathfrak B$ by a new unary predicate P with value A.

Example. Let $\mathfrak A = \langle A, \sim\rangle$ where $\sim$ is an equivalence relation on A and let $\mathfrak B \succ \mathfrak A$. Then $\langle \mathfrak A, \mathfrak B\rangle$ is a Vaughtian pair if, and only if, there is some $a \in A$ whose equivalence class

$$[a]_\sim := \{\, b \in B \mid b \sim a \,\}$$

is infinite and contained in A.

In the first part of this section we will study constructions of Vaughtian pairs. The goal is Lemma 5.8 which states that a countable theory with a Vaughtian pair cannot be κ-categorical for an uncountable cardinal κ. In the second part of the section, we will then investigate minimal sets in theories without Vaughtian pairs.

We will use the following lemma to construct new Vaughtian pairs from a given one.

Lemma 5.2. *Suppose that $\mathfrak{A} \subseteq \mathfrak{B}$ and $\mathfrak{A}' \subseteq \mathfrak{B}'$ are structures such that $\langle \mathfrak{B}, A \rangle \equiv \langle \mathfrak{B}', A' \rangle$.*

(a) *$\mathfrak{A} \preceq \mathfrak{B}$ if, and only if, $\mathfrak{A}' \preceq \mathfrak{B}'$.*

(b) *Let $\varphi(\bar{x}, \bar{y})$ be a formula and $\bar{a} \subseteq A$ and $\bar{a}' \subseteq A'$ tuples such that $\langle \mathfrak{B}, A, \bar{a} \rangle \equiv \langle \mathfrak{B}', A', \bar{a}' \rangle$. Then $\varphi(\bar{x}, \bar{a})$ is a witness for $\langle \mathfrak{A}, \mathfrak{B} \rangle$ being Vaughtian if, and only if, $\varphi(\bar{x}, \bar{a}')$ is a witness for $\langle \mathfrak{A}', \mathfrak{B}' \rangle$ being Vaughtian.*

Proof. (a) By symmetry, it is sufficient to prove one direction. For every formula $\psi(\bar{x})$, $\mathfrak{A} \preceq \mathfrak{B}$ implies

$$\langle \mathfrak{B}, A \rangle \vDash (\forall \bar{x}. \textstyle\bigwedge_i Px_i)[\psi(\bar{x}) \leftrightarrow \psi^{(P)}(\bar{x})],$$

where $\psi^{(P)}$ is the relativisation of ψ to P. Hence, all these formulae also hold in $\langle \mathfrak{B}', A' \rangle$. This implies that $\mathfrak{A}' \preceq \mathfrak{B}'$.

(b) Suppose that $\varphi(\bar{x}, \bar{a})$ witnesses that $\langle \mathfrak{B}, A \rangle$ is Vaughtian. By (a) and the fact that

$$\langle \mathfrak{B}, A \rangle \vDash \exists x \neg Px,$$

it follows that $\mathfrak{A}' \prec \mathfrak{B}'$. Furthermore, for every $n < \omega$,

$$\langle \mathfrak{B}, A \rangle \vDash \exists^n \bar{x} \varphi(\bar{x}, \bar{a}) \wedge \forall \bar{x}[\varphi(\bar{x}, \bar{a}) \rightarrow \textstyle\bigwedge_i Px_i].$$

Hence, the tuple $\bar{a}'$ satisfies these formulae in $\langle \mathfrak{B}', A' \rangle$. Consequently, $\varphi(\bar{x}, \bar{a}')^{\mathfrak{A}'}$ is infinite and $\varphi(\bar{x}, \bar{a}')^{\mathfrak{A}'} = \varphi(\bar{x}, \bar{a}')^{\mathfrak{B}'}$. □

The aim of the following sequence of results is Proposition 5.7 below which states that, given an arbitrary Vaughtian pair, we can construct a pair of size $\langle \kappa, \aleph_0 \rangle$, for every infinite cardinal κ.

Lemma 5.3. *Let T be a complete first-order theory. If there is a Vaughtian pair for T, then there are Vaughtian pairs for T of size $\langle \kappa, \kappa \rangle$, for every $\kappa \geq |T|$.*

Proof. Let $\mathfrak{A} \prec \mathfrak{B}$ be a Vaughtian pair for T and let $\varphi(\bar{x})$ be the corresponding formula with parameters $\bar{a} \subseteq A$. Since $\varphi^{\mathfrak{A}}$ is infinite, we can use the Compactness Theorem to construct an elementary extension $\langle \mathfrak{B}_1, A_1 \rangle \succeq \langle \mathfrak{B}, A \rangle$ such that $|\varphi^{\mathfrak{A}_1}| \geq \kappa$. By the Theorem of Löwenheim and Skolem, we can choose an elementary substructure $\langle \mathfrak{B}_0, A_0 \rangle \preceq \langle \mathfrak{B}_1, A_1 \rangle$ with $|B_0| = \kappa$, $|A_0| = \kappa$, and $\bar{a} \subseteq A_0$. By Lemma 5.2, it follows that $\mathfrak{A}_0 \prec \mathfrak{B}_0$ is a Vaughtian pair. □

Proposition 5.4. *Let T be a countable complete first-order theory. For every pair $\mathfrak{A}_0 \preceq \mathfrak{B}_0$ of countable models of T there exist countable homogeneous models $\mathfrak{A} \preceq \mathfrak{B}$ of T such that $\langle \mathfrak{B}_0, A_0 \rangle \preceq \langle \mathfrak{B}, A \rangle$ and $\mathfrak{A}$ and $\mathfrak{B}$ realise the same types in $S^{<\omega}(T)$.*

Proof. We start by proving the following claims.

(a) For every finite subset $U \subseteq A_0$ and every type $\mathfrak{p} \in S^{<\omega}(U)$, there exists a countable extension $\langle \mathfrak{B}, A \rangle \succeq \langle \mathfrak{B}_0, A_0 \rangle$ such that $\mathfrak{p}$ is realised in $\mathfrak{A} := \mathfrak{B}|_A$.

(b) For every finite subset $U \subseteq B_0$ and every type $\mathfrak{p} \in S^{<\omega}(U)$, there exists a countable extension $\langle \mathfrak{B}, A \rangle \succeq \langle \mathfrak{B}_0, A_0 \rangle$ such that $\mathfrak{p}$ is realised in $\mathfrak{B}$.

(c) There exists a countable extension $\langle \mathfrak{B}, A \rangle \succeq \langle \mathfrak{B}_0, A_0 \rangle$ such that $\mathfrak{A} := \mathfrak{B}|_A$ realises every type over a finite subset $U \subseteq A_0$ that is realised in $\mathfrak{B}_0$.

(a) We set

$$\Phi := \Delta \cup \{ \varphi^{(P)} \mid \varphi \in \mathfrak{p} \},$$

where Δ is the elementary diagram of $\langle\mathfrak{B}_0, A_0\rangle$. To show that Φ is satisfiable, we consider finitely many formulae $\varphi_0(\bar{x}), \dots, \varphi_{n-1}(\bar{x}) \in \mathfrak{p}$. Since $\mathfrak{p}$ is a type and, hence, finitely satisfiable in every model of T, we have $\mathfrak{A}_0 \vDash \exists\bar{x} \bigwedge_{i<n} \varphi_i(\bar{x})$, which implies that

$$\langle\mathfrak{B}_0, A_0\rangle \vDash \exists\bar{x} \bigwedge_{i<n} \varphi_i^{(P)}(\bar{x}) .$$

Consequently, Φ is finitely satisfiable. Fix a countable model $\langle\mathfrak{B}, A, \bar{a}\rangle$ of Φ. Then $\langle\mathfrak{B}_0, A_0\rangle \preceq \langle\mathfrak{B}, A\rangle$ and $\bar{a} \subseteq A$ realises $\mathfrak{p}$.

(b) This claim follows immediately from compactness and the Theorem of Löwenheim and Skolem.

(c) Let $(\mathfrak{p}_\alpha)_{\alpha<\omega}$ be an enumeration of all types over a finite set $U \subseteq A_0$ that are realised in $\mathfrak{B}_0$. We can use (a) to construct an increasing chain $\langle\mathfrak{B}_\alpha, A_\alpha\rangle_{\alpha<\omega}$ of countable models starting with $\langle\mathfrak{B}_0, A_0\rangle$ such that $\mathfrak{A}_{\alpha+1} := \mathfrak{B}_{\alpha+1}|_{A_{\alpha+1}}$ realises $\mathfrak{p}_\alpha$. The union $\langle\mathfrak{B}, A\rangle := \bigcup_{\alpha<\omega} \langle\mathfrak{B}_\alpha, A_\alpha\rangle$ is the desired extension of $\langle\mathfrak{B}_0, A_0\rangle$.

To prove the proposition we construct a chain $\langle\mathfrak{B}_\alpha, A_\alpha\rangle_{\alpha<\omega}$ of countable models starting with $\langle\mathfrak{B}_0, A_0\rangle$ as follows.

(1) For indices of the form $\alpha = 3n$, we use (c) to find a countable extension $\langle\mathfrak{B}_{\alpha+1}, A_{\alpha+1}\rangle \succeq \langle\mathfrak{B}_\alpha, A_\alpha\rangle$ such that every type over a finite set $U \subseteq A_\alpha$ that is realised in $\mathfrak{B}_\alpha$ is realised in $\mathfrak{A}_{\alpha+1}$.

(2) For indices $\alpha = 3n + 1$, we iterate (a) to find a countable extension $\langle\mathfrak{B}_{\alpha+1}, A_{\alpha+1}\rangle \succeq \langle\mathfrak{B}_\alpha, A_\alpha\rangle$ such that, for all tuples $\bar{a}, \bar{b} \in A_\alpha^{<\omega}$ with $\mathrm{tp}(\bar{a}) = \mathrm{tp}(\bar{b})$ and every element $c \in A_\alpha$, there is an element $d \in A_{\alpha+1}$ such that $\mathrm{tp}(\bar{a}c) = \mathrm{tp}(\bar{b}d)$.

(3) For $\alpha = 3n + 2$, we use (b), amalgamation, and the Theorem of Löwenheim and Skolem to find an extension $\langle\mathfrak{B}_{\alpha+1}, A_{\alpha+1}\rangle \succeq \langle\mathfrak{B}_\alpha, A_\alpha\rangle$ such that, for all tuples $\bar{a}, \bar{b} \in B_\alpha^{<\omega}$ with $\mathrm{tp}(\bar{a}) = \mathrm{tp}(\bar{b})$ and every element $c \in B_\alpha$, there is an element $d \in B_{\alpha+1}$ such that $\mathrm{tp}(\bar{a}c) = \mathrm{tp}(\bar{b}d)$.

The limit $\langle\mathfrak{B}, A\rangle := \bigcup_{\alpha<\omega} \langle\mathfrak{B}_\alpha, A_\alpha\rangle$ is a countable elementary extension of $\langle\mathfrak{B}_0, A_0\rangle$. Furthermore, by (1), the structures $\mathfrak{A} := \mathfrak{B}|_A$ and $\mathfrak{B}$ realise the same types in $S^{<\omega}(T)$. Finally, (2) and (3) ensure that $\mathfrak{A}$ and $\mathfrak{B}$ are homogeneous. □

Proposition 5.5. *Let T be a countable complete first-order theory. If there is a Vaughtian pair for T, then there is a Vaughtian pair for T of size $\langle \aleph_0, \aleph_1 \rangle$.*

Proof. By Lemma 5.3 and Proposition 5.4, we can find a Vaughtian pair $\mathfrak{A} \prec \mathfrak{B}$ for T of size $\langle \aleph_0, \aleph_0 \rangle$ such that $\mathfrak{A}$ and $\mathfrak{B}$ are homogeneous and realise the same types. By Theorem E1.1.9, this implies that $\mathfrak{A} \cong \mathfrak{B}$. Let φ be a formula over A such that $\varphi^{\mathfrak{A}}$ is infinite and $\varphi^{\mathfrak{B}} = \varphi^{\mathfrak{A}}$.

We construct an elementary chain $(\mathfrak{M}_\alpha)_{\alpha<\aleph_1}$ of models of T such that, for every $\alpha < \aleph_1$, we have

$$\varphi^{\mathfrak{M}_\alpha} = \varphi^{\mathfrak{A}} \quad \text{and} \quad \langle \mathfrak{M}_{\alpha+1}, M_\alpha \rangle \cong \langle \mathfrak{B}, A \rangle .$$

Note that, in particular, every $\mathfrak{M}_\alpha$ is isomorphic to $\mathfrak{A}$.

We start with $\mathfrak{M}_0 := \mathfrak{B}$. For the successor step, suppose that we have already defined $\mathfrak{M}_\alpha \cong \mathfrak{A}$. We choose an elementary extension $\mathfrak{M}_{\alpha+1} \succeq \mathfrak{M}_\alpha$ such that $\langle \mathfrak{M}_{\alpha+1}, M_\alpha \rangle \cong \langle \mathfrak{B}, A \rangle$. Then $\varphi^{\mathfrak{M}_{\alpha+1}} = \varphi^{\mathfrak{M}_\alpha} = \varphi^{\mathfrak{A}}$.

For limit ordinals δ, we set $\mathfrak{M}_\delta := \bigcup_{\alpha<\delta} \mathfrak{M}_\alpha$. Then $\varphi^{\mathfrak{M}_\delta} = \bigcup_{\alpha<\delta} \varphi^{\mathfrak{M}_\alpha} = \varphi^{\mathfrak{A}}$. To show that $\mathfrak{M}_\delta \cong \mathfrak{A}$ it is sufficient to prove that $\mathfrak{M}_\delta$ is homogeneous and that it realises the same types as $\mathfrak{A}$. For homogeneity, suppose that $\bar{a}, \bar{b} \in M_\delta^{<\omega}$ and $c \in M_\delta$ are elements such that $\operatorname{tp}(\bar{a}) = \operatorname{tp}(\bar{b})$. Then there is some $\alpha < \delta$ such that $\bar{a}, \bar{b}, c \subseteq M_\alpha$. As $\mathfrak{M}_\alpha \cong \mathfrak{A}$ is homogeneous, there is some $d \in M_\alpha \subseteq M_\delta$ such that $\operatorname{tp}(\bar{a}c) = \operatorname{tp}(\bar{b}d)$.

Clearly, every type realised in $\mathfrak{A}$ is realised in $\mathfrak{M}_\delta \succeq \mathfrak{A}$. Conversely, let $\mathfrak{p} \in S^{<\omega}(T)$ be realised in $\mathfrak{M}_\delta$. Then there is some $\bar{a} \in M_\delta^{<\omega}$ with $\operatorname{tp}(\bar{a}) = \mathfrak{p}$. Let $\alpha < \delta$ be an index such that $\bar{a} \subseteq M_\alpha$. Then $\mathfrak{p}$ is realised in $\mathfrak{M}_\alpha \cong \mathfrak{A}$.

Having defined $(\mathfrak{M}_\alpha)_\alpha$ we set $\mathfrak{N} := \bigcup_{\alpha<\aleph_1} \mathfrak{M}_\alpha$. Then $|N| = \aleph_1$ and $\varphi^{\mathfrak{N}} = \varphi^{\mathfrak{A}}$. Hence, $\mathfrak{A} \prec \mathfrak{N}$ is the desired Vaughtian pair of size $\langle \aleph_0, \aleph_1 \rangle$. □

Lemma 5.6. *Let T be a complete $\aleph_0$-stable theory over a countable signature. Every uncountable model $\mathfrak{M}$ of T has a proper elementary extension $\mathfrak{N} \succ \mathfrak{M}$ such that every countable type $\mathfrak{p}$ realised in $\mathfrak{N}$ is already realised in $\mathfrak{M}$.*

Proof. By Lemma 4.8 there exists a formula $\varphi(\bar{x})$ over M such that $|\varphi^{\mathfrak{M}}| \geq \aleph_1$ and we have either

$$|(\varphi \wedge \psi)^{\mathfrak{M}}| \leq \aleph_0 \quad \text{or} \quad |(\varphi \wedge \neg\psi)^{\mathfrak{M}}| \leq \aleph_0 ,$$

for every formula $\psi(\bar{x})$ over M. Let $\bar{s}$ be the sorts of the variables $\bar{x}$ and set

$$\mathfrak{p} := \{\, \psi(\bar{x}) \in \mathrm{FO}^{\bar{s}}[\Sigma_M] \mid (\varphi \wedge \psi)^{\mathfrak{M}} \text{ is uncountable} \,\} .$$

Note that, for $\psi_0, \dots, \psi_{n-1} \in \mathfrak{p}$, we have

$$\Big|(\varphi \wedge \bigvee_{i<n} \neg\psi_i)^{\mathfrak{M}}\Big| = \big|(\varphi \wedge \neg\psi_0)^{\mathfrak{M}} \cup \dots \cup (\varphi \wedge \neg\psi_{n-1})^{\mathfrak{M}}\big| \leq \aleph_0 ,$$

which implies that $\bigwedge_{i<n} \psi_i \in \mathfrak{p}$. Hence, $(\bigwedge_i \psi_i)^{\mathfrak{M}} \neq \emptyset$ and $\mathfrak{p}$ is finitely satisfiable. Furthermore, by choice of φ, we have $\psi \in \mathfrak{p}$ or $\neg\psi \in \mathfrak{p}$, for every formula $\psi(\bar{x})$ over M. Therefore, $\mathfrak{p}$ is a complete type.

Let $\mathfrak{M}_+ \succeq \mathfrak{M}$ be an elementary extension containing a finite tuple $\bar{a} \in M_+^{\bar{s}}$ realising $\mathfrak{p}$. By Theorem E3.4.14, there exists a model $\mathfrak{M} \preceq \mathfrak{N} \preceq \mathfrak{M}_+$ that is atomic over $M \cup \bar{a}$.

To show that $\mathfrak{N}$ has the desired property, we consider a countable type $\Phi(\bar{y})$ over M that is realised by some finite tuple $\bar{b} \in N^{<\omega}$. Since $\mathfrak{N}$ is atomic over $M \cup \bar{a}$, there exists a formula $\chi(\bar{y}, \bar{a})$ over M isolating $\mathrm{tp}(\bar{b}/M)$. Then $\mathfrak{N} \vDash \chi(\bar{b}, \bar{a})$ implies

$$\exists \bar{y}\chi(\bar{y}, \bar{x}) \in \mathfrak{p}$$

and $\quad \forall \bar{y}[\chi(\bar{y}, \bar{x}) \rightarrow \vartheta(\bar{y})] \in \mathfrak{p}, \quad$ for all $\vartheta(\bar{y}) \in \mathrm{tp}(\bar{b}/M) \supseteq \Phi$.

Hence, the set

$$\Gamma := \{\exists \bar{y}\chi(\bar{y}, \bar{x})\} \cup \{\, \forall \bar{y}[\chi(\bar{y}, \bar{x}) \rightarrow \vartheta(\bar{y})] \mid \vartheta(\bar{x}) \in \Phi \,\}$$

is a countable subset of $\mathfrak{p}$. Furthermore, if a tuple $\bar{a}' \in M^{\bar{s}}$ realises Γ then we have

$$\mathfrak{M} \vDash \exists \bar{y}\chi(\bar{y}, \bar{a}')$$

and every $\bar{b}' \subseteq M$ with $\mathfrak{M} \vDash \chi(\bar{b}', \bar{a}')$ realises Φ. Let $\psi_0, \psi_1, \dots$ be an enumeration of Γ. By choice of $\mathfrak{p}$, we have

$$\left|\varphi^{\mathfrak{M}}\right| > \aleph_0 \quad \text{and} \quad \left|(\varphi \wedge \neg(\psi_0 \wedge \dots \wedge \psi_n))^{\mathfrak{M}}\right| \leq \aleph_0 , \quad \text{for all } n .$$

It follows that $(\varphi \wedge \neg \bigwedge \Gamma)^{\mathfrak{M}} = \bigcup_{n<\omega} (\varphi \wedge \neg \bigwedge_{i<n} \psi_i)^{\mathfrak{M}}$ is countable and

$$(\varphi \wedge \bigwedge \Gamma)^{\mathfrak{M}} = \varphi^{\mathfrak{M}} \setminus (\varphi \wedge \neg \bigwedge \Gamma)^{\mathfrak{M}}$$

is uncountable. Hence, there are uncountably many $\bar{a}' \in M^{\bar{s}}$ such that

$$\mathfrak{M} \vDash \varphi(\bar{a}') \wedge \bigwedge \Gamma(\bar{a}') .$$

As we have seen above, this implies that $\mathfrak{M}$ contains a realisation of Φ. □

Proposition 5.7. *Let T be an $\aleph_0$-stable, countable, complete first-order theory. If there is a Vaughtian pair for T, then there are Vaughtian pairs for T of size $\langle \aleph_0, \kappa \rangle$, for every uncountable cardinal κ.*

Proof. By Proposition 5.5, there is a Vaughtian pair $\mathfrak{A} \prec \mathfrak{B}$ for T of size $\langle \aleph_0, \aleph_1 \rangle$. Let φ be a formula over A such that $\varphi^{\mathfrak{A}}$ is infinite and $\varphi^{\mathfrak{B}} = \varphi^{\mathfrak{A}}$. Starting with $\mathfrak{M}_0 := \mathfrak{B}$, we construct a strictly increasing elementary chain $(\mathfrak{M}_\alpha)_{\alpha<\kappa}$ such that $\varphi^{\mathfrak{M}_\alpha} = \varphi^{\mathfrak{A}}$, for all α.

As usual, we take unions $\mathfrak{M}_\delta := \bigcup_{\alpha<\delta} \mathfrak{M}_\alpha$ for limit ordinals δ. For the successor step, suppose that $\mathfrak{M}_\alpha$ has already been defined. We apply Lemma 5.6 to find a proper elementary extension $\mathfrak{M}_{\alpha+1} \succ \mathfrak{M}_\alpha$ that realises the same countable types as $\mathfrak{M}_\alpha$. In particular, $\mathfrak{M}_{\alpha+1}$ does not realise the type

$$\{\varphi(x)\} \cup \left\{ x \neq c \;\middle|\; c \in \varphi^{\mathfrak{M}_\alpha} \right\} .$$

Therefore, $\varphi^{\mathfrak{M}_{\alpha+1}} = \varphi^{\mathfrak{M}_\alpha} = \varphi^{\mathfrak{A}}$.

Let $\mathfrak{N} := \bigcup_{\alpha<\kappa} \mathfrak{M}_\alpha$ be the union of the chain and choose an elementary substructure $\mathfrak{A} \prec \mathfrak{C} \preceq \mathfrak{N}$ of size $|C| = \kappa$. Then $\mathfrak{A} \prec \mathfrak{C}$ is the desired Vaughtian pair of size $\langle \aleph_0, \kappa \rangle$. □

We can use this proposition to show that uncountably categorical theories do not have Vaughtian pairs.

Lemma 5.8. *Let T be a countable complete first-order theory with infinite models. If T is κ-categorical, for some uncountable cardinal κ, then T has no Vaughtian pairs.*

Proof. For a contradiction, suppose that T is a κ-categorical theory with a Vaughtian pair. By Theorem E6.3.16, T is $\aleph_0$-stable. Hence, we can use Proposition 5.7 to find a Vaughtian pair $\mathfrak{A} < \mathfrak{B}$ of size $\langle \aleph_0, \kappa \rangle$. Let φ be a formula such that $\varphi^{\mathfrak{A}}$ is infinite and $\varphi^{\mathfrak{B}} = \varphi^{\mathfrak{A}}$. By Theorem E1.2.16, T has a saturated model $\mathfrak{C}$ of size κ. But $\mathfrak{B} \ncong \mathfrak{C}$ since we have $|\varphi^{\mathfrak{C}}| = \kappa$ by Lemma E1.2.17. This contradicts κ-categoricity. □

Next we study minimal formulae in theories without Vaughtian pairs. First, we show that such a theory is graduated which, according to Theorem D1.2.15, is equivalent to admitting elimination of the quantifier $\exists^{\aleph_0}$.

Lemma 5.9. *Suppose that T is a theory without Vaughtian pairs. Let $\mathfrak{M}$ be a model of T and $\varphi(\bar{x};\bar{y})$ a formula over M. There exists a number $n < \omega$, such that, for all $\bar{c} \subseteq M$,*

$$\left|\varphi(\bar{x};\bar{c})^{\mathfrak{M}}\right| > n \quad \textit{implies} \quad \left|\varphi(\bar{x};\bar{c})^{\mathfrak{M}}\right| \geq \aleph_0 .$$

Proof. Suppose that such a number n does not exist. Then we can find, for every $n < \omega$, parameters $\bar{c}_n \subseteq M$ with

$$n < \left|\varphi(\bar{x};\bar{c}_n)\right| < \aleph_0 .$$

Let P be a new unary predicate and let $\Phi(\bar{y})$ be the set of formulae containing the following statements:

- P induces a proper elementary substructure;
- $\bigwedge_i P y_i$;
- there are infinitely many tuples $\bar{x}$ such that $\varphi(\bar{x};\bar{y})$;
- $\forall \bar{x}[\varphi(\bar{x};\bar{y}) \to \bigwedge_i P x_i]$.

To see that $T \cup \Phi(\bar{y})$ is satisfiable, we fix an extension $\mathfrak{N} \succ \mathfrak{M}$. Since $\varphi(\bar{x}; \bar{c}_n)^{\mathfrak{N}}$ is finite, we have $\varphi(\bar{x}; \bar{c}_n)^{\mathfrak{N}} = \varphi(\bar{x}; \bar{c}_n)^{\mathfrak{M}}$. For every finite subset $\Phi_0 \subseteq \Phi$, we can therefore choose n large enough such that

$$\langle \mathfrak{N}, M \rangle \vDash T \cup \Phi_0(\bar{c}_n) .$$

Let $\langle \mathfrak{B}, A, \bar{c} \rangle$ be a model of $T \cup \Phi$. Then $\mathfrak{A} := \mathfrak{B}|_A \prec \mathfrak{B}$ are models of T and $\varphi(\bar{x}; \bar{c})^{\mathfrak{A}} = \varphi(\bar{x}; \bar{c})^{\mathfrak{B}}$ is infinite. Hence, $\mathfrak{A} \prec \mathfrak{B}$ is a Vaughtian pair. A contradiction. □

Corollary 5.10. *In a theory T without Vaughtian pairs, every minimal formula is strongly minimal.*

Proof. Let $\mathfrak{M}$ be a model of T and $\varphi(\bar{x})$ a minimal formula over M. For a contradiction, suppose that $\varphi(\bar{x})$ is not strongly minimal. Then we can find an extension $\mathfrak{N} \succ \mathfrak{M}$ and a formula $\psi(\bar{x}; \bar{c})$ with parameters $\bar{c} \subseteq N$ such that

$$\varphi(\bar{x})^{\mathfrak{N}} \cap \psi(\bar{x}; \bar{c})^{\mathfrak{N}} \quad \text{and} \quad \varphi(\bar{x})^{\mathfrak{N}} \setminus \psi(\bar{x}; \bar{c})^{\mathfrak{N}}$$

are both infinite. By Lemma 5.9 there exists a number $n < \omega$ such that, for all models $\mathfrak{A}$ and all $\bar{a} \subseteq A$,

$$\begin{aligned} & \left|\varphi(\bar{x})^{\mathfrak{A}} \cap \psi(\bar{x}; \bar{a})^{\mathfrak{A}}\right| > n \quad \text{implies} \quad \left|\varphi(\bar{x})^{\mathfrak{A}} \cap \psi(\bar{x}; \bar{a})^{\mathfrak{A}}\right| \geq \aleph_0 , \\ \text{and} \quad & \left|\varphi(\bar{x})^{\mathfrak{A}} \setminus \psi(\bar{x}; \bar{a})^{\mathfrak{A}}\right| > n \quad \text{implies} \quad \left|\varphi(\bar{x})^{\mathfrak{A}} \setminus \psi(\bar{x}; \bar{a})^{\mathfrak{A}}\right| \geq \aleph_0 . \end{aligned}$$

By minimality of φ, it follows that

$$\mathfrak{M} \vDash \forall \bar{y} \big[\left|\varphi(\bar{x})^{\mathfrak{M}} \cap \psi(\bar{x}; \bar{y})^{\mathfrak{M}}\right| \leq n \vee \left|\varphi(\bar{x})^{\mathfrak{M}} \setminus \psi(\bar{x}; \bar{y})^{\mathfrak{M}}\right| \leq n \big] .$$

Since $\mathfrak{M} \preceq \mathfrak{N}$, the same formula also holds in $\mathfrak{N}$. A contradiction. □

Corollary 5.11. *Let T be a countable, complete, $\aleph_0$-stable theory without Vaughtian pairs and let $\mathfrak{M}_0$ be the prime model of T. There exists a strongly minimal formula $\varphi(x)$ over M_0.*

Proof. We use Corollary 4.9 to find a minimal formula $\varphi(x)$ over M_0. By Corollary 5.10, this formula is strongly minimal. □

Lemma 5.12. *Let T be a theory without Vaughtian pairs, $\mathfrak{B}$ a model of T, and let $\varphi(\bar{x};\bar{c})$ be a strongly minimal formula with parameters $\bar{c} \subseteq B$.*

(a) *If $\mathfrak{A} \prec \mathfrak{B}$ is a proper elementary substructure with $\bar{c} \subseteq A$, then $\varphi^{\mathfrak{A}} \subset \varphi^{\mathfrak{B}}$.*

(b) $\dim(\varphi^{\mathfrak{B}}) = |B|$.

(c) *If T is $\aleph_0$-stable then $\mathfrak{B}$ is prime over $\varphi^{\mathfrak{B}} \cup \bar{c}$.*

Proof. (a) $\mathfrak{A} \prec \mathfrak{B}$ implies $\varphi^{\mathfrak{A}} \subseteq \varphi^{\mathfrak{B}}$. Furthermore, if $\varphi^{\mathfrak{A}} = \varphi^{\mathfrak{B}}$, then $\mathfrak{A} \prec \mathfrak{B}$ would be a Vaughtian pair.

(b) Let I be a basis of $\varphi^{\mathfrak{B}}$. If $|I| < |B|$ then we can use the Theorem of Löwenheim and Skolem to find an elementary substructure $\mathfrak{A} \prec \mathfrak{B}$ of size $|A| = |I|$ with $I \cup \bar{c} \subseteq A$. It follows that $\varphi^{\mathfrak{B}} \subseteq \operatorname{acl}(I) \subseteq A$. Hence, $\varphi^{\mathfrak{B}} = \varphi^{\mathfrak{A}}$ in contradiction to (a).

(c) Since T is $\aleph_0$-stable there exists, according to Theorem E3.4.14 a unique prime model $\mathfrak{M}$ over $\varphi^{\mathfrak{B}} \cup \bar{c}$. W.l.o.g. we may assume that $\mathfrak{M} \preceq \mathfrak{B}$. Since $\varphi^{\mathfrak{B}} \cup \bar{c} \subseteq M \subseteq B$ it follows by (a) that $M = B$, as desired. □

Lemma 5.13. *Let T be a countable, complete first-order theory with infinite models. Suppose that there exists a strongly minimal formula $\varphi(x;\bar{c})$ such that*

- $\operatorname{tp}(\bar{c})$ *is isolated,*
- *every model $\mathfrak{M}$ of $T(\bar{c})$ is prime over $\varphi^{\mathfrak{M}} \cup \bar{c}$,*
- *no model $\mathfrak{M}$ of $T(\bar{c})$ has a proper elementary substructure $\mathfrak{A} \prec \mathfrak{M}$ such that $\varphi^{\mathfrak{M}} \subseteq A$.*

Then

$$\dim(\varphi^{\mathfrak{A}}/\bar{c}) = \dim(\varphi^{\mathfrak{B}}/\bar{c}) \quad \textit{implies} \quad \mathfrak{A} \cong \mathfrak{B},$$

for all models $\mathfrak{A}, \mathfrak{B}$ of $T(\bar{c})$.

Proof. Set $S := \varphi(\bar{x};\bar{c})^{\mathfrak{A}}$ and $S' := \varphi(\bar{x};\bar{c})^{\mathfrak{B}}$. Since $\dim(S) = \dim(S')$ we can use Corollary 4.11 to find an elementary bijection $h_0 : S \to S'$. As $\mathfrak{A}$ and $\mathfrak{B}$ are models of $T(\bar{c})$, we can extend h_0 to an elementary map

$h_1 : S \cup \bar{c} \to S' \cup \bar{c}$. Because $\mathfrak{A}$ is prime over $S \cup \bar{c}$, we can extend this map h_1 to an elementary map $h : \mathfrak{A} \to \mathfrak{B}$. We claim that h is surjective and, therefore, the desired isomorphism.

For a contradiction, suppose otherwise. Then we obtain a proper elementary substructure $\mathfrak{B}_0 := f[\mathfrak{A}] \prec \mathfrak{B}$ with $S' \cup \bar{c} = \operatorname{rng} h_1 \subseteq B_0$. But $\mathfrak{B}$ is prime over $S' \cup \bar{c}$. A contradiction. □

Theorem 5.14 (Morley). *Let T be a countable, complete first-order theory with infinite models. The following statements are equivalent:*

(1) *T is κ-categorical, for some uncountable cardinal κ.*

(2) *T is κ-categorical, for every uncountable cardinal κ.*

(3) *T is $\aleph_0$-stable and it has no Vaughtian pairs.*

(4) *There exists a strongly minimal formula $\varphi(x; \bar{c})$ such that*

- $\operatorname{tp}(\bar{c})$ *is isolated,*
- *every model $\mathfrak{M}$ of $T(\bar{c})$ is prime over $\varphi^{\mathfrak{M}} \cup \bar{c}$,*
- *no model $\mathfrak{M}$ of $T(\bar{c})$ has a proper elementary substructure $\mathfrak{A} \prec \mathfrak{M}$ such that $\varphi^{\mathfrak{M}} \subseteq A$.*

Proof. (2) ⇒ (1) is trivial.

(1) ⇒ (3) follows by Theorem E6.3.16 and Lemma 5.8.

(3) ⇒ (4) Let T be an $\aleph_0$-stable theory without Vaughtian pairs. By Theorem E3.4.14, T has a prime model $\mathfrak{M}_0$. We can use Corollary 5.11 to find a strongly minimal formula $\varphi(x; \bar{c})$ with parameters $\bar{c} \subseteq M_0$. Since prime models are atomic, the type of $\bar{c} \subseteq M_0$ is isolated. The remaining two claims of (4) follow by Lemma 5.12 (a) and (c), respectively.

(4) ⇒ (2) Let κ be an uncountable cardinal. To show that T is κ-categorical, we consider two models $\mathfrak{A}$ and $\mathfrak{B}$ of size κ. Since $\operatorname{tp}(\bar{c})$ is isolated there are tuples $\bar{a} \subseteq A$ and $\bar{b} \subseteq B$ realising $\operatorname{tp}(\bar{c})$. Thus, $\langle \mathfrak{A}, \bar{a} \rangle$ and $\langle \mathfrak{B}, \bar{b} \rangle$ are models of $T(\bar{c})$. Set $S := \varphi(\bar{x}; \bar{a})^{\mathfrak{A}}$ and $S' := \varphi(\bar{x}; \bar{b})^{\mathfrak{B}}$.

Since $\mathfrak{A}$ and $\mathfrak{B}$ have no proper elementary substructures containing, respectively, $S \cup \bar{a}$ and $S' \cup \bar{b}$, it follows by the Theorem of Löwenheim and Skolem that

$$\dim(S) = |A| = |B| = \dim(S') .$$

Consequently, we can use Lemma 5.13 to show that $\mathfrak{A} \cong \mathfrak{B}$. $\square$

F2. *Ranks and forking*

1. *Morley rank and Δ-rank*

We have seen that each model of an uncountably categorical theory is governed by a strongly minimal set and that we can define a geometry on such a set. Unfortunately, for most theories we cannot find actual geometries. But there is a large class of theories where we have something slightly weaker. In this chapter we study the kind of combinatorial structure that will serve as our substitute for a geometry.

We start by defining certain ranks that provide a weak notion of dimension. Guided by the observation that, for a strongly-minimal formula φ over a model $\mathfrak{M}$, the Cantor-Bendixson rank of the set $\langle\varphi\rangle$ in $\mathfrak{S}^{\bar{s}}(M)$ is equal to 1, we take a look at the Cantor-Bendixson rank of type spaces. Let us first describe how to compute the Cantor-Bendixson rank in $\mathfrak{S}_\Delta(U)$ by using the equality of Cantor-Bendixson rank and partition rank.

Lemma 1.1. *Let Δ be a set of formulae, U a set of parameters, and let Δ_U^+ be the set of all finite boolean combinations of formulae of the form $\psi(\bar{x};\bar{c})$ with $\psi(\bar{x};\bar{y}) \in \Delta$ and $\bar{c} \subseteq U$.*

For an arbitrary formula φ over U and an ordinal $\alpha > 0$, we have

$$\mathrm{rk}_{\mathrm{CB}}\big(\langle\varphi\rangle_{\mathfrak{S}_\Delta(U)}\big) \geq \alpha$$

if, and only if, for all ordinals $\beta < \alpha$, there are formulae $\psi_i \in \Delta_U^+$, for $i < \omega$, such that

$$\begin{aligned} &\mathrm{rk}_{\mathrm{CB}}\big(\langle\varphi \wedge \psi_i\rangle_{\mathfrak{S}_\Delta(U)}\big) \geq \beta\,, && \textit{for every } i\,, \\ \textit{and}\quad &\psi_i^{\mathbb{M}} \cap \psi_k^{\mathbb{M}} = \varnothing\,, && \textit{for all } i \neq k\,. \end{aligned}$$

Proof. Note that, by definition of $\mathfrak{S}_\Delta(U)$ and Lemma C3.3.5,

$$\begin{aligned}\mathfrak{S}_\Delta(U) &= \mathfrak{S}_{\Delta_U^\neg}\big(\mathrm{FO}[\Sigma_U, X]/T(U)\big)\\ &\cong \mathfrak{S}_{\Delta_U^+}\big(\mathrm{FO}[\Sigma_U, X]/T(U)\big) = \mathfrak{S}_{\Delta^+}(U)\,,\end{aligned}$$

where Δ^+ is the set of all finite boolean combinations of formulae in Δ. Therefore, we may w.l.o.g. work in $\mathfrak{S}_{\Delta^+}(U)$. Set $C := \langle\varphi\rangle_{\mathfrak{S}_{\Delta^+}(U)}$ and let $\mathfrak{S}_C$ be the subspace of $\mathfrak{S}_{\Delta^+}(U)$ induced by C. According to Corollary B5.7.10, we have

$$\mathrm{rk}_{\mathrm{CB}}\big(\langle\varphi\rangle_{\mathfrak{S}_{\Delta^+}(U)}\big) = \mathrm{rk}_{\mathrm{P}}(C/\mathrm{clop}(\mathfrak{S}_C))\,.$$

Furthermore,

$$\mathrm{rk}_{\mathrm{P}}(C/\mathrm{clop}(\mathfrak{S}_C)) \geq \alpha$$

if, and only if, for all $\beta < \alpha$, there are clopen sets $D_i \in \mathrm{clop}(\mathfrak{S}_C)$, for $i < \omega$, such that

$$\mathrm{rk}_{\mathrm{P}}(D_i/\mathrm{clop}(\mathfrak{S}_C)) \geq \beta \quad \text{and} \quad D_i \cap D_k = \varnothing\,, \quad \text{for } i \neq k\,.$$

Hence, it is sufficient to show that this latter condition is equivalent to the existence of formulae $\psi_i \in \Delta_U^+$, for $i < \omega$, such that

$$\begin{aligned}&\mathrm{rk}_{\mathrm{CB}}\big(\langle\varphi \wedge \psi_i\rangle_{\mathfrak{S}_\Delta(U)}\big) \geq \beta\,, && \text{for every } i\,,\\ \text{and}\quad &\psi_i^{\mathbb{M}} \cap \psi_k^{\mathbb{M}} = \varnothing\,, && \text{for all } i \neq k\,.\end{aligned}$$

($\Leftarrow$) Given formulae ψ_i, we set $D_i := \langle\varphi \wedge \psi_i\rangle_{\mathfrak{S}_{\Delta^+}(U)}$. By Corollaries B5.7.10 and B5.7.13, it follows that

$$\begin{aligned}\mathrm{rk}_{\mathrm{CB}}\big(\langle\varphi \wedge \psi_i\rangle_{\mathfrak{S}_{\Delta^+}(U)}\big) &= \mathrm{rk}_{\mathrm{P}}(D_i/\mathrm{clop}(D_i))\\ &= \mathrm{rk}_{\mathrm{P}}(D_i/\mathrm{clop}(\mathfrak{S}_C)) \geq \beta\,,\end{aligned}$$

as desired.

(⇒) By Lemma B5.7.11, the clopen sets D_i are of the form

$$D_i = C \cap \langle \psi'_i \rangle_{\mathfrak{S}_{\Delta^+}(U)} = \langle \varphi \wedge \psi'_i \rangle_{\mathfrak{S}_{\Delta^+}(U)} ,$$

for formulae $\psi'_i \in \Delta^+_U$. Setting

$$\psi_i := \psi'_i \wedge \bigwedge_{k<i} \neg \psi'_k$$

we obtain formulae $\psi_i \in \Delta^+_U$ such that

$$\psi_i^{\mathbb{M}} \cap \psi_k^{\mathbb{M}} = \varnothing , \quad \text{for } i \neq k .$$

Furthermore, $D_i \cap D_k = \varnothing$, for $k < i$, implies that

$$D_i = D_i \setminus (D_0 \cup \dots \cup D_{i-1}) = \langle \varphi \wedge \psi_i \rangle_{\mathfrak{S}_{\Delta^+}(U)} .$$

The claim follows since, by Corollaries B5.7.10 and B5.7.13,

$$\begin{aligned} \operatorname{rk}_{\mathrm{CB}}\big(\langle \varphi \wedge \psi_i \rangle_{\mathfrak{S}_{\Delta^+}(U)}\big) &= \operatorname{rk}_{\mathrm{P}}(D_i/\operatorname{clop}(D_i)) \\ &= \operatorname{rk}_{\mathrm{P}}(D_i/\operatorname{clop}(\mathfrak{S}_C)) \geq \beta . \end{aligned}$$

□

When using the Cantor-Bendixson rank to define the dimension of a definable relation, we have first to choose a set Δ of formulae and a set U of parameters to know which type space $\mathfrak{S}_\Delta(U)$ to consider. Let us take a look at what happens to the Cantor-Bendixson rank when we change these two sets. First of all, the dependence is monotone: if we enlarge the set of formulae or the set of parameters, the rank either increases, or it stays the same.

Lemma 1.2. *Let Δ, Γ be sets of formulae, U, V sets of parameters, and Φ a set of formulae over U. Then*

$$\operatorname{rk}_{\mathrm{CB}}\big(\langle \Phi \rangle_{\mathfrak{S}_\Delta(U)}\big) \leq \operatorname{rk}_{\mathrm{CB}}\big(\langle \Phi \rangle_{\mathfrak{S}_{\Delta\cup\Gamma}(U\cup V)}\big) .$$

Proof. Let $\Delta^\neg_U$ be the sets of all formulae of the form $\psi(\bar{x};\bar{c})$ or $\neg\psi(\bar{x};\bar{c})$ with $\psi \in \Delta$ and $\bar{c} \subseteq U$, and let $\Delta\Gamma^\neg_{UV}$ be the corresponding set of formulae for $\Delta \cup \Gamma$ and $U \cup V$. The statement follows from Lemma B5.7.14 since

$$\mathfrak{S}(i)^{-1}\big[\langle \Phi \rangle_{\mathfrak{S}_\Delta(U)}\big] = \langle \Phi \rangle_{\mathfrak{S}_{\Delta\cup\Gamma}(U\cup V)} ,$$

where $i : \Delta^\neg_U \to \Delta\Gamma^\neg_{UV}$ is the inclusion map. □

If the set of parameters is an $\aleph_0$-saturated model, the Cantor-Bendixson rank does not change anymore.

Lemma 1.3. *Let Δ be a set of formulae and $\varphi(\bar{x};\bar{y})$ a single formula. If $\mathfrak{A}$ and $\mathfrak{B}$ are $\aleph_0$-saturated structures with $\langle\mathfrak{A},\bar{a}\rangle \equiv \langle\mathfrak{B},\bar{b}\rangle$, then*

$$\mathrm{rk}_{\mathrm{CB}}\big(\langle\varphi(\bar{x};\bar{a})\rangle_{\mathfrak{S}_\Delta(A)}\big) = \mathrm{rk}_{\mathrm{CB}}\big(\langle\varphi(\bar{x};\bar{b})\rangle_{\mathfrak{S}_\Delta(B)}\big)\,.$$

Proof. By symmetry it is sufficient to prove that

$$\mathrm{rk}_{\mathrm{CB}}\big(\langle\varphi(\bar{x};\bar{a})\rangle_{\mathfrak{S}_\Delta(A)}\big) \geq \alpha$$

implies

$$\mathrm{rk}_{\mathrm{CB}}\big(\langle\varphi(\bar{x};\bar{b})\rangle_{\mathfrak{S}_\Delta(B)}\big) \geq \alpha\,.$$

We proceed by induction on α. For $\alpha = 0$ there is nothing to do. Since the limit step follows immediately from the inductive hypothesis, we may therefore assume that $\alpha = \beta + 1$. If

$$\mathrm{rk}_{\mathrm{CB}}\big(\langle\varphi(\bar{x};\bar{a})\rangle_{\mathfrak{S}_\Delta(A)}\big) \geq \beta + 1\,,$$

we can use Lemma 1.1 to find formulae $\psi_n(\bar{x};\bar{c}^n) \in \Delta_A^+$, for $n < \omega$, with $\bar{c}^n \subseteq A$ such that

$$\mathrm{rk}_{\mathrm{CB}}\big(\langle\varphi(\bar{x};\bar{a}) \wedge \psi_n(\bar{x};\bar{c}^n)\rangle_{\mathfrak{S}_\Delta(A)}\big) \geq \beta\,,$$

and $\quad \mathfrak{A} \vDash \neg[\psi_m(\bar{x};\bar{c}^m) \wedge \psi_n(\bar{x};\bar{c}^n)]\,, \quad$ for $m \neq n$.

Since $\mathfrak{A} \sqsubseteq^{\aleph_0}_{\mathrm{FO}} \mathfrak{B}$, we can inductively find tuples $\bar{d}^n \subseteq B$, for $n < \omega$, such that

$$\langle\mathfrak{A}, \bar{a}\bar{c}^0 \dots \bar{c}^n\rangle \equiv \langle\mathfrak{B}, \bar{b}\bar{d}^0 \dots \bar{d}^n\rangle\,, \quad \text{for all } n < \omega\,.$$

This implies that

$$\mathfrak{B} \vDash \neg[\psi_m(\bar{x};\bar{d}^m) \wedge \psi_n(\bar{x};\bar{d}^n)]\,, \quad \text{for } m \neq n\,.$$

By inductive hypothesis, we furthermore have

$$\mathrm{rk}_{\mathrm{CB}}\big(\langle \varphi(\bar{x};\bar{b}) \wedge \psi_n(\bar{x};\bar{d}^n)\rangle_{\mathfrak{S}_\Delta(B)}\big) \geq \beta\,, \quad \text{for all } n\,.$$

Consequently, Lemma 1.1 implies that

$$\mathrm{rk}_{\mathrm{CB}}\big(\langle \varphi(\bar{x};\bar{b})\rangle_{\mathfrak{S}_\Delta(B)}\big) \geq \beta + 1\,. \qquad \square$$

It follows that there is a limit of the Cantor-Bendixson rank for increasing sets of parameters. This limit is called the *Δ-rank* of the theory.

Definition 1.4. (a) Let Δ be a set of formulae and $\varphi(\bar{x};\bar{c})$ an FO-formula with parameters $\bar{c} \subseteq \mathbb{M}$. The *$\Delta$-rank* of φ is

$$\mathrm{rk}_\Delta(\varphi(\bar{x};\bar{c})) := \mathrm{rk}_{\mathrm{CB}}\big(\langle \varphi(\bar{x};\bar{c})\rangle_{\mathfrak{S}_\Delta(M)}\big)\,,$$

where $\mathfrak{M} \leq \mathbb{M}$ is an arbitrary $\aleph_0$-saturated model with $\bar{c} \subseteq M$.

(b) Let $\bar{s}$ be a tuple of sorts and let $\varphi(\bar{x};\bar{c})$ be an FO-formula with parameters $\bar{c} \subseteq \mathbb{M}$. The *Morley rank* of φ is

$$\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi(\bar{x};\bar{c})) := \mathrm{rk}_\Delta(\varphi(\bar{x};\bar{c}))\,,$$

where Δ is the set of all first-order formulae $\psi(\bar{x};\bar{y})$ where the variables $\bar{x}$ have sorts $\bar{s}$.

(c) For a set of formulae $\Phi(\bar{x})$ (possibly with parameters) we define

$$\mathrm{rk}_\Delta(\Phi) := \min\{\, \mathrm{rk}_\Delta(\varphi) \mid \Phi \vDash \varphi \,\}\,,$$
$$\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\Phi) := \min\{\, \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi) \mid \Phi \vDash \varphi \,\}\,.$$

For $\bar{a} \in \mathbb{M}^{\bar{s}}$ and $U \subseteq \mathbb{M}$, we set

$$\mathrm{rk}_\Delta(\bar{a}/U) := \mathrm{rk}_\Delta(\mathrm{tp}(\bar{a}/U))\,,$$
$$\mathrm{rk}_{\mathrm{M}}(\bar{a}/U) := \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\mathrm{tp}(\bar{a}/U))\,.$$

Remark. (a) Note that, by Lemmas 1.2 and 1.3, the definitions of $\mathrm{rk}_\Delta(\varphi)$ and $\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi)$ do not depend on the choice of $\mathfrak{M}$. According to Theorem C3.4.5 (b), they also do not depend on what we consider the free variables of the formula φ. But note that, by Lemma 1.2, we have $\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi) \leq$

$\mathrm{rk}^{\bar{t}}_{\mathrm{M}}(\varphi)$, for $\bar{s} \subseteq \bar{t}$. This inequality can be strict. An example is given by the formula $x = x$ with respect to the theory of infinite structures with empty signature. Then $\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(x = x) = |\bar{s}|$.

(b) If $\mathfrak{p}$ is a complete type over an $\aleph_0$-saturated model $\mathfrak{M}$, it follows by Theorem B5.7.8 and Corollary B5.7.9 that

$$\mathrm{rk}_{\Delta}(\mathfrak{p}) = \mathrm{rk}_{\mathrm{CB}}(\mathfrak{p}/\mathfrak{S}_{\Delta}(M))\,.$$

Example. Consider the theory T of structures of the form $\langle A, \sim \rangle$, where $\sim$ is an equivalence relation on A with infinitely many classes, all of which are infinite. For $a \in \mathbb{M}$ and a model $\mathfrak{M} \prec \mathbb{M}$, we have

$$\mathrm{rk}_{\mathrm{M}}(a/M) = \begin{cases} 0 & \text{if } a \in M\,, \\ 1 & \text{if } a \notin M \text{ and } a \sim b \text{ for some } b \in M\,, \\ 2 & \text{otherwise}\,. \end{cases}$$

Exercise 1.1. Show that $\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi) = 1$, for every strongly minimal formula $\varphi(\bar{x})$.

Exercise 1.2. Let T be the theory of structures of the form $\langle A, \sim \rangle$, where $\sim$ is an equivalence relation on A with infinitely many classes, all of which are infinite. Determine the possible values of $\mathrm{rk}_{\mathrm{M}}(ab/M)$, for two elements $a, b \in \mathbb{M}$ and a model $\mathfrak{M} \prec \mathbb{M}$.

Let us collect some basic properties of the Δ-rank of a formula.

Lemma 1.5. *Let T be a theory and φ, ψ formulae.*

(a) $T \cup \{\varphi\} \vDash \psi$ *implies* $\mathrm{rk}_{\Delta}(\varphi) \leq \mathrm{rk}_{\Delta}(\psi)$.

(b) $\mathrm{rk}_{\Delta}(\varphi \vee \psi) = \max\{\mathrm{rk}_{\Delta}(\varphi), \mathrm{rk}_{\Delta}(\psi)\}$.

(c) *If Δ contains the formula $x = y$, then* $\mathrm{rk}_{\Delta}(\varphi) = 0$ *if, and only if, φ is algebraic and consistent with T.*

Proof. (a) follows from Lemma B2.5.10, (b) from Lemma B2.5.11, and (c) follows immediately from the definition. □

Exercise 1.3. Show that $\mathrm{rk}_\Delta(\varphi \wedge \psi) \leq \min\{\mathrm{rk}_\Delta(\varphi), \mathrm{rk}_\Delta(\psi)\}$, and that this inequality may be strict.

Lemma 1.6. *Let $\bar a, \bar b \subseteq \mathbb{M}$ be tuples and $U, V \subseteq \mathbb{M}$ sets of parameters.*

(a) $\mathrm{rk}_\Delta(\bar a/U) \leq \mathrm{rk}_{\Delta\cup\Gamma}(\bar a/U)$.

(b) $\mathrm{rk}_\Delta(\bar a/U) \geq \mathrm{rk}_\Delta(\bar a/U \cup V)$.

(c) *There exists a finite subset $U_0 \subseteq U$ with* $\mathrm{rk}_\Delta(\bar a/U) = \mathrm{rk}_\Delta(\bar a/U_0)$.

Proof. (a) follows immediately from Lemma 1.2.

(b) By definition of the Δ-rank of a type, we have

$$\begin{aligned}\mathrm{rk}_\Delta(\bar a/U) &= \min\{\, \mathrm{rk}_\Delta(\varphi) \mid \varphi \in \mathrm{tp}(\bar a/U)\,\} \\ &\geq \min\{\, \mathrm{rk}_\Delta(\varphi) \mid \varphi \in \mathrm{tp}(\bar a/U \cup V)\,\} \\ &= \mathrm{rk}_\Delta(\bar a/U \cup V)\,.\end{aligned}$$

(c) Fix a formula $\varphi \in \mathrm{tp}(\bar a/U)$ such that $\mathrm{rk}_\Delta(\varphi) = \mathrm{rk}_\Delta(\bar a/U)$. Let $U_0 \subseteq U$ be the finite set of parameters from φ. Then $\varphi \in \mathrm{tp}(\bar a/U_0)$ implies

$$\mathrm{rk}_\Delta(\bar a/U_0) \leq \mathrm{rk}_\Delta(\varphi) = \mathrm{rk}_\Delta(\bar a/U) \leq \mathrm{rk}_\Delta(\bar a/U_0)\,,$$

where the last inequality holds by (b). □

For theories where it is defined, the Morley rank is usually better behaved than the Δ-rank. Let us collect some of its properties, in particular with respect to strongly minimal sets. First of all note that, using the equivalence of the Morley rank of a formula and its partition rank, we can define a notion of degree.

Definition 1.7. The *Morley degree* $\deg^{\bar s}_{\mathrm{M}}(\varphi)$ of a formula φ is the maximal number $m < \omega$ such that there are formulae $\psi_0, \dots, \psi_{m-1}$ of rank $\mathrm{rk}^{\bar s}_{\mathrm{M}}(\psi_i) = \mathrm{rk}^{\bar s}_{\mathrm{M}}(\varphi)$ such that $\psi_i^{\mathbb{M}} \cap \psi_k^{\mathbb{M}} = \varnothing$, for $i \neq k$. If such a number m does not exist, we set $\deg^{\bar s}_{\mathrm{M}}(\varphi) := \infty$.

Remark. It follows by Lemma B2.5.16 that

$$\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi) < \infty \quad \text{implies} \quad \deg^{\bar{s}}_{\mathrm{M}}(\varphi) < \infty\,.$$

Exercise 1.4. Show that a formula $\varphi(\bar{x})$ is strongly minimal if, and only if, $\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi) = 1$ and $\deg^{\bar{s}}_{\mathrm{M}}(\varphi) = 1$.

For types there is a related notion of degree: the number of *free* extensions.

Definition 1.8. Let $\mathfrak{p} \subseteq \mathfrak{q}$ be (partial) types with free variables of sort $\bar{s}$. We say that $\mathfrak{q}$ is a *Morley-free extension* of $\mathfrak{p}$ if $\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\mathfrak{q}) = \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\mathfrak{p})$.

Lemma 1.9. *Let $\mathfrak{p}$ be a (partial) type over U and suppose that $U \subseteq V$.*

(a) *$\mathfrak{p}$ has a Morley-free extension $\mathfrak{q} \in S^{\bar{s}}(V)$.*

(b) *If $\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\mathfrak{p}) < \infty$, then $\mathfrak{p}$ has only finitely many Morley-free extensions in $S^{\bar{s}}(V)$.*

Proof. Choose an $\aleph_0$-saturated model $\mathfrak{M}$ containing V.

(a) First suppose that $\alpha := \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\mathfrak{p}) < \infty$. According to Lemma B5.5.15, the closed set $\langle \mathfrak{p} \rangle_{\mathfrak{S}^{\bar{s}}(M)}$ contains some type $\mathfrak{r}$ with

$$\mathrm{rk}_{\mathrm{CB}}(\mathfrak{r}/\mathfrak{S}^{\bar{s}}(M)) = \mathrm{rk}_{\mathrm{CB}}(\langle \mathfrak{p} \rangle_{\mathfrak{S}^{\bar{s}}(M)}) = \alpha\,.$$

Set $\mathfrak{q} := \mathfrak{r}|_V$. Then $\mathfrak{p} \subseteq \mathfrak{q} \subseteq \mathfrak{r}$ implies

$$\alpha = \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\mathfrak{p}) \geq \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\mathfrak{q}) \geq \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\mathfrak{r}) = \mathrm{rk}_{\mathrm{CB}}(\mathfrak{r}/\mathfrak{S}^{\bar{s}}(M)) = \alpha\,.$$

Consequently, $\mathfrak{q}$ is the desired extension of $\mathfrak{p}$.

It remains to consider the case where $\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\mathfrak{p}) = \infty$. Then

$$\mathrm{rk}_{\mathrm{CB}}(\langle \mathfrak{p} \rangle_{\mathfrak{S}^{\bar{s}}(M)}) = \infty$$

implies that there is some $\mathfrak{r} \in \langle \mathfrak{p} \rangle_{\mathfrak{S}^{\bar{s}}(M)}$ with $\mathrm{rk}_{\mathrm{CB}}(\mathfrak{r}/\mathfrak{S}^{\bar{s}}(M)) = \infty$. As above, it follows that $\mathfrak{q} := \mathfrak{r}|_V$ is the desired Morley-free extension of $\mathfrak{p}$ over V.

(b) Let $\alpha := \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\mathfrak{p})$. By (a), every type $\mathfrak{q} \in \langle \mathfrak{p} \rangle_{\mathfrak{S}^{\bar{s}}(V)}$ of rank α has an extension $\mathfrak{r} \in \langle \mathfrak{p} \rangle_{\mathfrak{S}^{\bar{s}}(M)}$ of the same rank. These extensions are obviously distinct, for different types $\mathfrak{q}$. The claim follows since, according to Lemma B5.5.15, the set $\langle \mathfrak{p} \rangle_{\mathfrak{S}^{\bar{s}}(M)}$ contains only finitely many types $\mathfrak{r}$ with $\mathrm{rk}_{\mathrm{CB}}(\mathfrak{r}/\mathfrak{S}^{\bar{s}}(M)) = \alpha$. □

Corollary 1.10. *For every formula $\varphi(\bar{x})$ over a set U, there exists some $\bar{a} \in \varphi^{\mathbb{M}}$ with $\mathrm{rk}_{\mathrm{M}}(\bar{a}/U) = \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi)$, where $\bar{s}$ are the sorts of $\bar{x}$.*

Proof. By Lemma 1.9, there exists a type $\mathfrak{q} \in S^{\bar{s}}(U)$ with $\{\varphi\} \subseteq \mathfrak{q}$ and $\mathrm{rk}_{\mathrm{M}}(\mathfrak{q}) = \mathrm{rk}_{\mathrm{M}}(\varphi)$. Every tuple $\bar{a}$ realising $\mathfrak{q}$ has the desired properties. □

The following lemmas show that the notion of Morley rank generalises the dimension of a strongly minimal set. We start by showing that the Morley rank increases with the length of a tuple and that elements in the algebraic closure do not increase the rank.

Lemma 1.11. *Let T be a first-order theory and let $\varphi(\bar{x}, \bar{y})$ be a formula with free variables $\bar{x}$ and $\bar{y}$ of sorts $\bar{s}$ and $\bar{t}$, respectively. Then*

$$\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\exists \bar{y} \varphi) \leq \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi)\,.$$

Proof. We prove by induction on α that

$$\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\exists \bar{y} \varphi) \geq \alpha \quad \text{implies} \quad \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi) \geq \alpha\,.$$

For $\alpha = 0$, it is sufficient to note that the consistency of $\exists \bar{y} \varphi$ implies the one of φ. Hence, suppose that $\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\exists \bar{y} \varphi) \geq \alpha$, for some $\alpha > 0$, and let $\beta < \alpha$. By Lemma 1.1, there are formulae $\psi_k(\bar{x})$, for $k < \omega$, such that

$$\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\exists \bar{y} \varphi \wedge \psi_k) \geq \beta \quad \text{and} \quad \psi_i^{\mathbb{M}} \cap \psi_k^{\mathbb{M}} = \varnothing\,, \quad \text{for all } i \neq k\,.$$

Note that, if $T \vDash \neg \exists \bar{y}\, \mathrm{true}$, then $\exists \bar{y} \varphi$ is inconsistent with T. Hence, $\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\exists \bar{y} \varphi) = -1 \leq \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi)$ and we are done. Consequently, we may assume that $T \vDash \exists \bar{y}\, \mathrm{true}$. We therefore have

$$\exists \bar{y} \varphi(\bar{x}, \bar{y}) \wedge \psi_k(\bar{x}) \equiv \exists \bar{y} [\varphi(\bar{x}, \bar{y}) \wedge \psi_k(\bar{x})] \quad \text{modulo } T.$$

It follows by inductive hypothesis that

$$\beta \le \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\exists \bar{y}\varphi \wedge \psi_k) = \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\exists \bar{y}(\varphi \wedge \psi_k)) \le \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi \wedge \psi_k)\,.$$

Since this holds for every β, it follows by Lemma 1.1 that $\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi) \ge \alpha$. □

Lemma 1.12. *Let $\bar{a} \in \mathbb{M}^{\bar{s}}$ and $\bar{b} \in \mathbb{M}^{\bar{t}}$ be finite tuples and $U \subseteq \mathbb{M}$ a set of parameters.*

(a) $\mathrm{rk}_{\mathrm{M}}(\bar{a}/U) \le \mathrm{rk}_{\mathrm{M}}(\bar{a}\bar{b}/U)$.

(b) $\mathrm{rk}_{\Delta}(\bar{a}/\operatorname{acl}(U)) = \mathrm{rk}_{\Delta}(\bar{a}/U)$.

(c) $\mathrm{rk}_{\mathrm{M}}(\bar{a}c/U) = \mathrm{rk}_{\mathrm{M}}(\bar{a}/U)$, *for all $c \in \operatorname{acl}(U \cup \bar{a})$.*

Proof. (a) Let $\alpha := \mathrm{rk}_{\mathrm{M}}(\bar{a}\bar{b}/U)$. By definition, there is a formula $\varphi(\bar{x}, \bar{y})$ over U such that $\mathbb{M} \vDash \varphi(\bar{a}, \bar{b})$ and $\mathrm{rk}^{\bar{s}\bar{t}}_{\mathrm{M}}(\varphi) = \mathrm{rk}_{\mathrm{M}}(\bar{a}\bar{b}/U)$. Then $\exists \bar{y}\varphi \in \operatorname{tp}(\bar{a}/U)$ implies, by Lemma 1.11, that

$$\mathrm{rk}_{\mathrm{M}}(\bar{a}/U) \le \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\exists \bar{y}\varphi) \le \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi) \le \mathrm{rk}^{\bar{s}\bar{t}}_{\mathrm{M}}(\varphi) = \mathrm{rk}_{\mathrm{M}}(\bar{a}\bar{b}/U)\,,$$

as desired.

(b) It follows by Lemma 1.6 that $\mathrm{rk}_{\mathrm{M}}(\bar{a}/\operatorname{acl}(U)) \le \mathrm{rk}_{\mathrm{M}}(\bar{a}/U)$. For a contradiction, suppose that this inequality is strict. Then there is some formula $\varphi(\bar{x}; \bar{c}) \in \operatorname{tp}(\bar{a}/\operatorname{acl}(U))$ such that $\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi(\bar{x}; \bar{c})) < \mathrm{rk}_{\mathrm{M}}(\bar{a}/U)$. Since $\bar{c}$ is algebraic over U, we know by Lemma E3.1.3 that $\operatorname{tp}(\bar{c}/U)$ is isolated. Let $\psi(\bar{y})$ be a formula over U isolating this type and set

$$\vartheta(\bar{x}) := \exists \bar{y}[\varphi(\bar{x}; \bar{y}) \wedge \psi(\bar{y})]\,.$$

Then $\vartheta(\bar{x}) \in \operatorname{tp}(\bar{a}/U)$ implies, by Lemmas 1.5 and 1.11, that

$$\mathrm{rk}_{\mathrm{M}}(\bar{a}/U) \le \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\vartheta) \le \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi \wedge \psi) \le \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi) < \mathrm{rk}_{\mathrm{M}}(\bar{a}/U)\,.$$

A contradiction.

(c) We have just seen in (a) that $\mathrm{rk}_{\mathrm{M}}(\bar{a}c/U) \ge \mathrm{rk}_{\mathrm{M}}(\bar{a}/U)$. For the converse inequality, we prove by induction on α that, for elements $c \in \operatorname{acl}(U \cup \bar{a})$,

$$\mathrm{rk}_{\mathrm{M}}(\bar{a}c/U) \ge \alpha \quad \text{implies} \quad \mathrm{rk}_{\mathrm{M}}(\bar{a}/U) \ge \alpha\,.$$

For $\alpha = 0$, note that $\mathrm{rk}_{\mathrm{M}}(\bar{a}/U) \geq 0$ since $\mathrm{tp}(\bar{a}/U)$ is satisfiable. For limit ordinals α, the claim follows immediately by the inductive hypothesis. For the successor step, let

$$\mathrm{rk}_{\mathrm{M}}(\bar{a}c/U) \geq \alpha + 1$$

and, for a contradiction, suppose that $\mathrm{rk}_{\mathrm{M}}(\bar{a}/U) \leq \alpha$. Fix a formula $\varphi(\bar{x}) \in \mathrm{tp}(\bar{a}/U)$ over U with minimal rank. Since $c \in \mathrm{acl}(\bar{a}/U)$, there is a formula $\chi(\bar{x}, y)$ over U such that $\chi(\bar{a}, y)^{\mathbb{M}}$ is a finite set containing c. Let $m := |\chi(\bar{a}, y)^{\mathbb{M}}|$ and set

$$\vartheta(\bar{x}, y) := \varphi(\bar{x}) \wedge \chi(\bar{x}, y) \wedge \neg\exists^{m+1} y \chi(\bar{x}, y) .$$

Since $\vartheta \in \mathrm{tp}(\bar{a}c/U)$ we have $\mathrm{rk}_{\mathrm{M}}^{\bar{s}u}(\vartheta) \geq \mathrm{rk}_{\mathrm{M}}(\bar{a}c/U) \geq \alpha + 1$, where u is the sort of c. By Lemma 1.1, there are formulae ψ_n, for $n < \omega$, such that $\mathrm{rk}_{\mathrm{M}}^{\bar{s}u}(\vartheta \wedge \psi_n) \geq \alpha$ and $\psi_i^{\mathbb{M}} \cap \psi_k^{\mathbb{M}} = \varnothing$, for $i \neq k$. Set

$$\eta_n := \exists y(\vartheta \wedge \psi_n) \quad \text{and} \quad \eta_I := \bigwedge_{i \in I} \eta_i , \quad \text{for } I \subseteq \omega .$$

First, let us show that $\mathrm{rk}_{\mathrm{M}}^{\bar{s}}(\eta_n) \geq \alpha$. By Lemma 1.10, there exists a tuple $\bar{b}d \in (\vartheta \wedge \psi_n)^{\mathbb{M}}$ such that $\mathrm{rk}_{\mathrm{M}}(\bar{b}d/U) = \mathrm{rk}_{\mathrm{M}}^{\bar{s}u}(\vartheta \wedge \psi_n)$. Then $d \in \mathrm{acl}(\bar{b})$ and, by inductive hypothesis,

$$\mathrm{rk}_{\mathrm{M}}(\bar{b}d/U) = \mathrm{rk}_{\mathrm{M}}^{\bar{s}u}(\vartheta \wedge \psi_n) \geq \alpha \quad \text{implies} \quad \mathrm{rk}_{\mathrm{M}}(\bar{b}/U) \geq \alpha .$$

Since $\eta_n \in \mathrm{tp}(\bar{b}/U)$, it follows that $\mathrm{rk}_{\mathrm{M}}^{\bar{s}}(\eta_n) \geq \alpha$.

Furthermore, for every set $I \subseteq \omega$ of size $|I| > m$, the formula η_I is unsatisfiable since $\mathbb{M} \vDash \eta_I(\bar{b})$ implies that there are elements $d_i \in \mathbb{M}$, for $i \in I$, such that $\mathbb{M} \vDash \vartheta_i(\bar{b}, d_i)$. But, since $|\vartheta(\bar{b}, y)|^{\mathbb{M}} \leq m$ there must be indices $i < k$ in I such that $d_i = d_k$. Hence, $\bar{b}d_i$ satisfies $\psi_i \wedge \psi_k$, which contradicts our choice of the formulae ψ_n, $n < \omega$.

In particular, $\mathrm{rk}_{\mathrm{M}}^{\bar{s}}(\eta_I) = -1 < \alpha$, for large enough sets I. The set

$$F := \left\{ I \subseteq \omega \;\middle|\; \mathrm{rk}_{\mathrm{M}}^{\bar{s}}(\eta_I) \geq \alpha \text{ and there is no } J \supset I \text{ with } \mathrm{rk}_{\mathrm{M}}^{\bar{s}}(\eta_J) \geq \alpha \right\}$$

is infinite, since every $I \in F$ is finite and, for each $n < \omega$, there is some $I \in F$ with $n \in I$. Fix countably many distinct sets $I_0, I_1, \dots \in F$ and set

$$\xi_n := \eta_{I_n} \wedge \bigwedge_{i<n} \neg\eta_{I_i} .$$

By definition of F, $i \neq k$ implies $I_i \nsubseteq I_k$. Therefore, $I_i \cup I_k \notin F$ and

$$\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\eta_{I_i} \wedge \eta_{I_k}) = \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\eta_{I_i \cup I_k}) < \alpha , \quad \text{for } i \neq k .$$

By Lemma 1.5, this implies that

$$\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\eta_{I_i} \wedge \bigvee_{k<i} \eta_{I_k}) = \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\bigvee_{k<i}(\eta_{I_i} \wedge \eta_{I_k})) < \alpha .$$

Since $\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\eta_{I_i}) = \alpha$, it therefore follows that

$$\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\xi_i) = \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\eta_{I_i} \wedge \neg \bigvee_{k<i} \eta_{I_k}) \geq \alpha .$$

Note that $\xi_i \models \exists y\vartheta \models \varphi$ implies $\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi \wedge \xi_i) \geq \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\xi_i) \geq \alpha$. As $\xi_i^{\mathbb{M}} \cap \xi_k^{\mathbb{M}} = \emptyset$, for $i \neq k$, it therefore follows by Lemma 1.1 that

$$\alpha < \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi) = \mathrm{rk}_{\mathrm{M}}(\bar{a}/U) \leq \alpha .$$

A contradiction. □

Corollary 1.13. *Let $\varphi(\bar{x})$ and $\psi(\bar{y})$ be formulae with parameters and let $\bar{s}$ and $\bar{t}$ by the sorts of, respectively, $\bar{x}$ and $\bar{y}$. If there exists a parameter-definable surjective function $f : \varphi^{\mathbb{M}} \to \psi^{\mathbb{M}}$ such that $f^{-1}(\bar{b})$ is finite, for every $\bar{b} \in \psi^{\mathbb{M}}$, then*

$$\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi) = \mathrm{rk}^{\bar{t}}_{\mathrm{M}}(\psi) .$$

Proof. Let $U \subseteq \mathbb{M}$ be a set of parameters such that φ and ψ are over U and f is definable over U. By assumption, every $\bar{a} \in \varphi^{\mathbb{M}}$ is algebraic over $U \cup \{f(\bar{a})\}$. Since $f(\bar{a})$ is algebraic over $U \cup \bar{a}$, it follows by Lemma 1.12 that

$$\mathrm{rk}_{\mathrm{M}}(\bar{a}/U) = \mathrm{rk}_{\mathrm{M}}(\bar{a}f(\bar{a})/U) = \mathrm{rk}_{\mathrm{M}}(f(\bar{a})/U) .$$

We can use Corollary 1.10 to find tuples $\bar{a} \in \varphi^{\mathbb{M}}$ and $\bar{b} \in \psi^{\mathbb{M}}$ with

$$\mathrm{rk}_{\mathrm{M}}(\bar{a}/U) = \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi) \quad \text{and} \quad \mathrm{rk}_{\mathrm{M}}(\bar{b}/U) = \mathrm{rk}^{\bar{t}}_{\mathrm{M}}(\psi)\,.$$

Then $\psi \in \mathrm{tp}(f(\bar{a})/U)$ implies

$$\mathrm{rk}^{\bar{t}}_{\mathrm{M}}(\psi) \geq \mathrm{rk}_{\mathrm{M}}(f(\bar{a})/U) = \mathrm{rk}_{\mathrm{M}}(\bar{a}/U) = \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi)\,.$$

Conversely, by surjectivity of f, there is some $\bar{c} \in f^{-1}(\bar{b})$. Therefore,

$$\mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\varphi) \geq \mathrm{rk}_{\mathrm{M}}(\bar{c}/U) = \mathrm{rk}_{\mathrm{M}}(\bar{b}/U) = \mathrm{rk}^{\bar{t}}_{\mathrm{M}}(\psi)\,. \qquad \square$$

Finally, we are able to show that, in a strongly minimal set, the Morley rank of a finite tuple coincides with its dimension.

Theorem 1.14. *Let $\varphi(x)$ be a strongly minimal formula over U.*

$$\mathrm{rk}_{\mathrm{M}}(\bar{a}/U) = \dim_{\mathrm{acl}}(\bar{a}/U)\,, \quad \textit{for all finite tuples } \bar{a} \subseteq \varphi^{\mathbb{M}}.$$

Proof. Let $\bar{a}_0 \subseteq \bar{a}$ be an acl-basis of $\bar{a}$ over U. Then $|\bar{a}_0| = \dim_{\mathrm{acl}}(\bar{a}/U)$ and it follows by Lemma 1.12 that

$$\mathrm{rk}_{\mathrm{M}}(\bar{a}/U) = \mathrm{rk}_{\mathrm{M}}(\bar{a}_0/U)\,.$$

Hence, it is sufficient to prove that $\mathrm{rk}_{\mathrm{M}}(\bar{a}_0/U) = |\bar{a}_0|$. W.l.o.g. we may assume that $\bar{a}_0 = \bar{a}$, i.e., $\bar{a}$ is independent over U. We prove the claim by induction on $m := |\bar{a}|$. Let $\bar{s}$ be the sorts of $\bar{a}$.

First, suppose that $m = 1$, i.e., $\bar{a} = a_0$ and $\bar{s} = s_0$. As $\mathrm{tp}(a_0/U)$ contains the strongly minimal formula $\varphi(x)$, we have $\mathrm{rk}_{\mathrm{M}}(a_0/U) \leq \mathrm{rk}^{s_0}_{\mathrm{M}}(\varphi) = 1$. Conversely, $a_0 \notin \mathrm{acl}(U)$ implies that $\mathrm{tp}(a_0/U)$ is non-algebraic. Hence, for every formula $\psi(x) \in \mathrm{tp}(a_0/U)$, the set $\psi^{\mathbb{M}}$ is infinite and, therefore, $\mathrm{rk}^{s_0}_{\mathrm{M}}(\psi) \geq 1$.

For the inductive step, suppose that $m > 1$. We start by showing that $\mathrm{rk}_{\mathrm{M}}(\bar{a}/U) \geq m$. Note that $|\mathrm{acl}(A)| \leq |T|$, for every countable set A, while $|\varphi^{\mathbb{M}}| = |\mathbb{M}| > |T|$. Therefore, $\dim_{\mathrm{acl}}(\varphi^{\mathbb{M}}) > \aleph_0$ and we can fix a countably infinite set $I = \{ b^n_i \mid n < \omega,\ i < m \} \subseteq \varphi^{\mathbb{M}}$ that is independent over U. Setting $\bar b^n := \langle b^n_0, \dots, \bar b^n_{m-1} \rangle$, it follows by Proposition F1.4.6 that

$$\operatorname{tp}(\bar b^n/U) = \operatorname{tp}(\bar a/U)\,, \quad \text{for every } n < \omega\,.$$

Let $I_0 := \{ b^n_0 \mid n < \omega \}$. Lemma F1.3.4 (a) implies that

$$\dim_{\mathrm{acl}}(\bar b^n/U \cup I_0) = \dim_{\mathrm{acl}}(\bar b^n/U \cup \{b^n_0\}) = m - 1\,.$$

By inductive hypothesis it therefore follows that

$$\operatorname{rk}_{\mathrm{M}}(\bar b^n/U \cup I_0) = m - 1\,.$$

Let $\vartheta(\bar x) \in \operatorname{tp}(\bar a/U)$ be a formula with $\operatorname{rk}^{\bar s}_{\mathrm{M}}(\vartheta) = \operatorname{rk}_{\mathrm{M}}(\bar a/U)$ and set $\psi_n(\bar x) := x_0 = b^n_0$. Then $\vartheta \wedge \psi_n \in \operatorname{tp}(\bar b^n/U \cup I_0)$ implies that

$$\operatorname{rk}^{\bar s}_{\mathrm{M}}(\vartheta \wedge \psi_n) \geq \operatorname{rk}_{\mathrm{M}}(\bar b^n/U \cup I_0) \geq m - 1\,.$$

Since $\psi^{\mathbb{M}}_i \cap \psi^{\mathbb{M}}_k = \emptyset$, for $i \neq k$, it follows by Lemma 1.1 that

$$\operatorname{rk}_{\mathrm{M}}(\bar a/U) = \operatorname{rk}^{\bar s}_{\mathrm{M}}(\vartheta) > \operatorname{rk}^{\bar s}_{\mathrm{M}}(\vartheta \wedge \psi_n) \geq m - 1\,.$$

It remains to prove that $\operatorname{rk}_{\mathrm{M}}(\bar a/U) \leq m$. Let $\mathfrak{M}$ be an $\aleph_0$-saturated model containing U. According to Proposition F1.4.6, every tuple $\bar c$ that is independent over M has the same type over U as $\bar a$. Replacing $\bar a$ by $\bar c$ we may therefore w.l.o.g. assume that $\bar a$ is independent over M. Fix a formula $\vartheta \in \operatorname{tp}(\bar a/U)$ such that $\operatorname{rk}^{\bar s}_{\mathrm{M}}(\vartheta) = \operatorname{rk}_{\mathrm{M}}(\bar a/U)$. For a contradiction, suppose that $\operatorname{rk}^{\bar s}_{\mathrm{M}}(\vartheta) > m$. Then, by Lemma 1.1, there are formulae ψ_i, $i < \omega$, such that $\operatorname{rk}^{\bar s}_{\mathrm{M}}(\vartheta \wedge \psi_i) \geq m$ and $\psi^{\mathbb{M}}_i \cap \psi^{\mathbb{M}}_k = \emptyset$, for $i \neq k$. By Lemma 1.3 and the definition of Morley rank, we can choose the formulae ψ_i over M. Since the sets $\psi^{\mathbb{M}}_i$ are disjoint, there is some index i such that $\bar a \notin \psi^{\mathbb{M}}_i$. Consequently, there exists a formula $\psi := \psi_i$ over M such that $\neg\psi \in \operatorname{tp}(\bar a/M)$ and $\operatorname{rk}^{\bar s}_{\mathrm{M}}(\psi) \geq \operatorname{rk}^{\bar s}_{\mathrm{M}}(\vartheta \wedge \psi) \geq m$.

By Corollary 1.10, there exists a tuple $\bar b \in \psi^{\mathbb{M}}$ with $\operatorname{rk}_{\mathrm{M}}(\bar b/M) = \operatorname{rk}^{\bar s}_{\mathrm{M}}(\psi)$. Since $\operatorname{tp}(\bar b/M) \neq \operatorname{tp}(\bar a/M)$, Proposition F1.4.6 implies that $\bar b$ is not independent over M. Let $\bar{b}_0 \subseteq \bar{b}$ be an acl-basis of $\bar{b}$ over M. By Lemma 1.12 and inductive hypothesis, it follows that

$$\begin{aligned} m \leq \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\psi) = \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\bar{b}/M) &= \mathrm{rk}^{\bar{s}}_{\mathrm{M}}(\bar{b}_0/M) \\ &= \dim_{\mathrm{acl}}(\bar{b}_0/M) = |\bar{b}_0| < m\,, \end{aligned}$$

a contradiction. □

2. Independence relations

Besides closure operators and dimensions, a matroid can also be characterised in terms of a so-called *independence relation.* This characterisation is the easiest to generalise to the geometry-like configurations appearing in model theory. In this section we introduce independence relations and show that they give an alternative characterisation of matroids. In the next section, we then present the generalisation used in model theory.

Definition 2.1. Let cl be a closure operator on the set Ω. The *independence relation* $\overset{\mathrm{cl}}{\surd}$ associated with cl is the ternary relation between sets $A, B, U \subseteq \Omega$ that is defined by

$$A \overset{\mathrm{cl}}{\surd}_U B \quad :\text{iff} \quad \begin{array}{l}\text{every set } I \subseteq B \text{ that is independent over } U \\ \text{is also independent over } U \cup A\,.\end{array}$$

Example. Let $\mathfrak{V}$ be a vector space, $A, B, U \subseteq V$ subspaces with $U \subseteq A, B$, and let cl be the closure operator mapping a set $X \subseteq V$ to the subspace $\langle\!\langle X \rangle\!\rangle_{\mathfrak{V}}$ spanned by X. Then

$$A \overset{\mathrm{cl}}{\surd}_U B \quad \text{iff} \quad A \cap B = U\,.$$

In the abstract, the properties of an independence relation $\overset{\mathrm{cl}}{\surd}$ are given by the following axioms.

Definition 2.2. Let Ω be a set and let $A \surd_U B$ be a ternary relation on subsets $A, B, U \subseteq \Omega$.

(a) $\surd$ is an *abstract independence relation* if it satisfies the following conditions:

(MON) *Monotonicity.* If $A_0 \subseteq A$ and $B_0 \subseteq B$ then

$$A \surd_U B \quad \text{implies} \quad A_0 \surd_U B_0 \,.$$

(NOR) *Normality.*

$$A \surd_U B \quad \text{implies} \quad A \cup U \surd_U B \cup U \,.$$

(LRF) *Left Reflexivity.*

$$A \surd_A B \,, \quad \text{for all } A, B \subseteq \Omega \,.$$

(LTR) *Left Transitivity.* If $A_0 \subseteq A_1 \subseteq A_2$ then

$$A_2 \surd_{A_1} B \quad \text{and} \quad A_1 \surd_{A_0} B \quad \text{implies} \quad A_2 \surd_{A_0} B \,.$$

(FIN) *Finite Character.*

$$A \surd_U B \quad \text{iff} \quad A_0 \surd_U B \quad \text{for all finite } A_0 \subseteq A \,.$$

(b) A *geometric independence relation* is an abstract independence relation $\surd$ that satisfies the following additional conditions:

(SYM) *Symmetry.*

$$A \surd_U B \quad \text{implies} \quad B \surd_U A \,.$$

(BMON) *Base Monotonicity.*

$$A \surd_U B \cup C \quad \text{implies} \quad A \surd_{U \cup C} B \cup C \,.$$

(SRB) *Strong Right Boundedness.* Let γ be an ordinal and let $(U_\alpha)_{\alpha \leq \gamma}$ be a strictly increasing chain of subsets $U_\alpha \subseteq \Omega$. If $A \not\sqrt{}_{U_\alpha} U_{\alpha+1}$, for all $\alpha < \gamma$, then $|\gamma| \leq |A|$.

(c) We call an abstract independence relation *symmetric, base monotone,* or *strongly right bounded* if it satisfies the corresponding axiom. Frequently, we will use the symbol $\perp$ to denote symmetric independence relations.

Example. (a) Let Ω be a set. For $A, B, U \subseteq \Omega$, we set

$$A \overset{\circ}{\sqrt{}}_U B \quad :\text{iff} \quad A \subseteq U\,.$$

$\overset{\circ}{\sqrt{}}$ is an abstract independence relation on Ω that satisfies (BMON) and (SRB), but not (SYM). Moreover, it is minimal in the sense that $\overset{\circ}{\sqrt{}} \subseteq \sqrt{}$, for every abstract independence relation $\sqrt{}$ on Ω.

(b) Let Ω be a set. For $A, B, U \subseteq \Omega$, define

$$A \perp^{\circ}_U B \quad :\text{iff} \quad A \cap B \subseteq U\,.$$

Then $\perp^{\circ}$ is a geometric independence relation. It is minimal in the sense that $\perp^{\circ} \subseteq \perp$, for every symmetric independence relation on Ω. Note that $\perp^{\circ} = \overset{\mathrm{cl}}{\sqrt{}}$, where $\mathrm{cl} : X \mapsto X$ is the trivial closure operator on Ω.

(c) Let $\mathfrak{G} = \langle V, E \rangle$ be an undirected graph. For $A, B, U \subseteq V$, we define

$$A \perp^{\mathrm{sep}}_U B \quad :\text{iff} \quad \text{every path connecting an element of } A \text{ to an element of } B \text{ contains an element of } U.$$

Then $\perp^{\mathrm{sep}}$ is an abstract independence relation that is symmetric and base monotone.

As most axioms are immediate we only check left transitivity. Suppose, for a contradiction, that $A_2 \perp^{\mathrm{sep}}_{A_1} B$ and $A_1 \perp^{\mathrm{sep}}_{A_0} B$, but $A_2 \not\perp^{\mathrm{sep}}_{A_0} B$. Then there exists a path π from some vertex $a_2 \in A_2$ to some $b \in B$ such that π does not contain an element of A_0. Since $A_2 \perp^{\mathrm{sep}}_{A_1} B$, this path contains a vertex $a_1 \in A_1$. Let π' be the subpath of π connecting a_1 to b.

Since $A_1 \downarrow^{\mathrm{sep}}_{A_0} B$, this subpath contains a vertex of A_0. Hence, so does π. A contradiction.

(d) Let $\mathfrak{X} = \langle X, d\rangle$ be a metric space. For $A, B, U \subseteq X$, we define

$$A \downarrow^{d}_{U} B \quad :\text{iff} \quad \begin{array}{l}\text{for all } a \in A \text{ and } b \in B \text{ there is some } c \in U \\ \text{such that } d(a,b) = d(a,c) + d(c,b)\,.\end{array}$$

Again, $\downarrow^{d}$ is a symmetric abstract independence relation.

Note that, for (undirected) trees, this definition generalises that in (c). Given a tree T, we define the distance between two vertices $u, v \in T$ as the length of the unique path between u and v. The independence relation $\downarrow^{d}$ corresponding to this metric coincides with $\downarrow^{\mathrm{sep}}$ from (c) since the equation $d(u,v) = d(u,w) + d(w,v)$ implies that w is a vertex on the path from u to v.

Exercise 2.1. Given an abstract independence relation $\surd$, we define the relation

$$A \overset{\mathrm{b}}{\surd}_{U} B \quad :\text{iff} \quad A \surd_{UB_0} B\,, \quad \text{for all } B_0 \subseteq B\,.$$

Prove that $\overset{\mathrm{b}}{\surd}$ is a base monotone abstract independence relation.

Let us collect some immediate consequences of the axioms of an abstract independence relation. In proofs we will usually use the axioms (MON), (NOR), and (LRF) tacitly, while all uses of other axioms will be explicit. The first two lemmas contain versions of the left transitivity axiom that are frequently more convenient to use. The third lemma presents an infinite version of left transitivity.

Lemma 2.3. *Let $\surd$ be an abstract independence relation.*

$$A \surd_{U \cup C} B \quad \textit{and} \quad C \surd_{U} B \quad \textit{implies} \quad A \cup C \surd_{U} B\,.$$

Proof. By (NOR), we have $A \cup U \cup C \surd_{U \cup C} B$ and $C \cup U \surd_{U} B$. By (LTR) it follows that $A \cup U \cup C \surd_{U} B$. □

Lemma 2.4. *Let $\surd$ be a base monotone abstract independence relation.*

$$A \surd_U B \cup C \quad \textit{and} \quad C \surd_U B \quad \textit{implies} \quad A \cup C \surd_U B\,.$$

Proof. By (BMON), $A \surd_U B \cup C$ implies $A \surd_{U \cup C} B \cup C$. Since $C \surd_U B$, it follows by Lemma 2.3 and monotonicity that $A \cup C \surd_U B$. □

Lemma 2.5. *Let $\surd$ be an abstract independence relation.*

(a) *If $(A_i)_{i \in I}$ is an increasing chain of sets with $A_i \surd_U B$, for all $i \in I$, then $\bigcup_{i \in I} A_i \surd_U B$.*

(b) *If γ is an ordinal and $(A_\alpha)_{\alpha < \gamma}$ an increasing chain of sets with $A_\alpha \surd_{U \cup \bigcup_{i<\alpha} A_i} B$, for all $\alpha < \gamma$, then $\bigcup_{\alpha<\gamma} A_\alpha \surd_U B$.*

Proof. (a) By (FIN) it is sufficient to show that $C \surd_U B$, for all finite $C \subseteq \bigcup_{i \in I} A_i$. Hence, let $C \subseteq \bigcup_{i \in I} A_i$ be finite. As $(A_i)_{i \in I}$ is increasing, there exists an index $i \in I$ such that $C \subseteq A_i$. Consequently, $A_i \surd_U B$ implies that $C \surd_U B$.

(b) We prove the claim by induction on γ. For $\gamma = 0$, we have $\varnothing \surd_U B$ by (LRF). For the inductive step, suppose that $\bigcup_{i<\alpha} A_i \surd_U B$, for all $\alpha < \gamma$. By (a) it follows that $\bigcup_{\alpha<\gamma} \bigcup_{i<\alpha} A_i \surd_U B$. If γ is a limit ordinal, then $\bigcup_{\alpha<\gamma} \bigcup_{i<\alpha} A_i = \bigcup_{\alpha<\gamma} A_\alpha$ and we are done. Hence, suppose that $\gamma = \beta + 1$. Then

$$A_\beta \surd_{U \cup \bigcup_{i<\beta} A_i} B \quad \text{and} \quad \bigcup_{i<\beta} A_i \surd_U B$$

implies, by Lemma 2.3, that $A_\beta \surd_U B$. □

We will show that geometric independence relations are precisely those associated with a matroid. The easy direction is to show that every matroid induces a geometric independence relation. As a first step, let us see which axioms hold if we do not assume the exchange property.

Lemma 2.6. *The independence relation* $\mathop{\smile\!\!\!\!\!\!|}^{\mathrm{cl}}$ *associated with a finitary closure operator* cl *on* Ω *is an abstract independence relation.*

Proof. We have to check five axioms.

(MON) Suppose that $A \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{cl}}_U B$ and let $A_0 \subseteq A$ and $B_0 \subseteq B$. To show that $A_0 \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{cl}}_U B_0$, consider a subset $I \subseteq B_0$ that is independent over U. Since $A \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{cl}}_U B$, I is also independent over $U \cup A$. In particular, it is independent over $U \cup A_0$.

(NOR) Suppose that $A \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{cl}}_U B$. To show that $A \cup U \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{cl}}_U B \cup U$, consider a set $I \subseteq B \cup U$ that is independent over U. Then $I \subseteq B$ and $A \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{cl}}_U B$ implies that I is independent over $U \cup A$.

(LRF) Trivially, if $I \subseteq B$ is independent over A, then it is independent over A.

(LTR) Suppose that $A_2 \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{cl}}_{A_1} B$ and $A_1 \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{cl}}_{A_0} B$, for $A_0 \subseteq A_1 \subseteq A_2$. If I is independent over A_0, it is independent over A_1 and, hence, also over A_2.

(FIN) Suppose that $A \mathop{\not\smile\!\!\!\!\!\!|}^{\mathrm{cl}}_U B$. We have to find a finite set $A_0 \subseteq A$ such that $A_0 \mathop{\not\smile\!\!\!\!\!\!|}^{\mathrm{cl}}_U B$. By assumption, there is a set $I \subseteq B$ that is independent over U, but not over $U \cup A$. Hence, there is some element $b \in I$ such that $b \in \mathrm{cl}(U \cup A \cup (I \setminus \{b\}))$. We choose a finite subset $A_0 \subseteq A$ such that $b \in \mathrm{cl}(U \cup A_0 \cup (I \setminus \{b\}))$. Since I is independent over U, but not over $U \cup A_0$, it follows that $A_0 \mathop{\not\smile\!\!\!\!\!\!|}^{\mathrm{cl}}_U B$. □

To show that, for a matroid $\langle \Omega, \mathrm{cl} \rangle$, the relation $\mathop{\smile\!\!\!\!\!\!|}^{\mathrm{cl}}$ is a geometric independence relation, we start with a technical lemma.

Lemma 2.7. *Let* $\langle \Omega, \mathrm{cl} \rangle$ *be a matroid and let* $I, J \subseteq \Omega$ *be sets that are both independent over* U. *If* I *is independent over* $U \cup J$, *then* J *is independent over* $U \cup I$.

Proof. Suppose that J is not independent over $U \cup I$. Then there is some $b \in J$ such that

$$b \in \mathrm{cl}(U \cup I \cup (J \setminus \{b\})) \setminus \mathrm{cl}(U \cup (J \setminus \{b\}))\,.$$

By the exchange property, there is some $a \in I$ such that

$$a \in \mathrm{cl}(U \cup (I \setminus \{a\}) \cup J)\,.$$

Consequently, I is not independent over $U \cup J$. □

Proposition 2.8. *The relation* $\overset{\mathrm{cl}}{\mathrel{\raise0.2ex\hbox{$\smile$}\kern-0.9em\lower0.25ex\hbox{$\mid$}}}$ *associated with a matroid* $\langle \Omega, \mathrm{cl} \rangle$ *is a geometric independence relation.*

Proof. We have already seen in Lemma 2.6 that $\overset{\mathrm{cl}}{\mathrel{\raise0.2ex\hbox{$\smile$}\kern-0.9em\lower0.25ex\hbox{$\mid$}}}$ is an abstract independence relation. Hence, it remains to check the following three axioms.

(SYM) Suppose that $A \overset{\mathrm{cl}}{\mathrel{\raise0.2ex\hbox{$\smile$}\kern-0.9em\lower0.25ex\hbox{$\mid$}}}_U B$. To show that $B \overset{\mathrm{cl}}{\mathrel{\raise0.2ex\hbox{$\smile$}\kern-0.9em\lower0.25ex\hbox{$\mid$}}}_U A$, consider a set $I \subseteq A$ that is independent over U. Let J be a basis of B over U. By assumption, J is independent over $U \cup A$. Hence, it follows by Lemma 2.7 that I is independent over $U \cup J$ and, therefore, over $U \cup B$.

(BMON) Since we have already shown (SYM), it is sufficient to prove that $A \cup C \overset{\mathrm{cl}}{\mathrel{\raise0.2ex\hbox{$\smile$}\kern-0.9em\lower0.25ex\hbox{$\mid$}}}_U B$ implies $A \cup C \overset{\mathrm{cl}}{\mathrel{\raise0.2ex\hbox{$\smile$}\kern-0.9em\lower0.25ex\hbox{$\mid$}}}_{U \cup C} B$. Thus, suppose that $A \cup C \overset{\mathrm{cl}}{\mathrel{\raise0.2ex\hbox{$\smile$}\kern-0.9em\lower0.25ex\hbox{$\mid$}}}_U B$. If $I \subseteq B$ is independent over $U \cup C$, it is also independent over U and, hence, over $U \cup A \cup C$.

(SRB) Let $(U_\alpha)_{\alpha \leq \gamma}$ be a strictly increasing sequence with $A \overset{\mathrm{cl}}{\not\mathrel{\raise0.2ex\hbox{$\smile$}\kern-0.9em\lower0.25ex\hbox{$\mid$}}}_{U_\alpha} U_{\alpha+1}$, for all $\alpha < \gamma$. By induction on α, we construct a decreasing chain $(I_\alpha)_{\alpha \leq \gamma}$ of subsets $I_\alpha \subseteq A$ such that I_α is a basis of A over U_α. We start with an arbitrary basis I_0 of A over U_0. For the inductive step, suppose that we have already defined I_β for all $\beta < \alpha$. For I_α we choose a maximal subset of $\bigcap_{\beta < \alpha} I_\beta$ that is independent over U_α.

Since $A \overset{\mathrm{cl}}{\not\mathrel{\raise0.2ex\hbox{$\smile$}\kern-0.9em\lower0.25ex\hbox{$\mid$}}}_{U_\alpha} U_{\alpha+1}$ we can find a set $J \subseteq U_{\alpha+1}$ that is independent over U_α, but not over $U_\alpha \cup A$. By Lemma 2.7 it follows that I_α is not independent over $U_\alpha \cup J \subseteq U_{\alpha+1}$. Therefore, each inclusion $I_{\alpha+1} \supset I_\alpha$ is strict. It follows that $|\gamma| \leq |I_0| \leq |A|$. □

Our next aim is to show that every geometric independence relation arises from a matroid. As motivation for the definition below, let us explain how one can recover the closure operation cl from the independence relation $\overset{\mathrm{cl}}{\mathrel{\raise0.2ex\hbox{$\smile$}\kern-0.9em\lower0.25ex\hbox{$\mid$}}}$ associated with it.

Lemma 2.9. *Let* $\overset{\text{cl}}{\surd}$ *be the independence relation associated with a closure operator* cl *on* Ω *and let* $a \in \Omega$ *and* $A, B, U \subseteq \Omega$.

(a) $a \in \text{cl}(U)$ iff $a \overset{\text{cl}}{\surd}_U a$

iff $a \overset{\text{cl}}{\surd}_{U \cup C} B$ *for all* $B, C \subseteq \Omega$.

(b) $A \subseteq \text{cl}(U \cup B)$ iff $B \overset{\text{cl}}{\surd}_U C \Rightarrow A \overset{\text{cl}}{\surd}_U C$ *for all* $C \subseteq \Omega$.

Proof. (a) First, suppose that $a \in \text{cl}(U)$. We claim that $a \overset{\text{cl}}{\surd}_{U \cup C} B$, for all $B, C \subseteq \Omega$. Fix B and C and let $I \subseteq B$ be independent over $U \cup C$. Then I is independent over $\text{cl}(U \cup C)$ and, therefore, over $U \cup \{a\} \subseteq \text{cl}(U \cup C)$.

If $a \overset{\text{cl}}{\surd}_{U \cup C} B$, for all B, C, then, trivially, $a \overset{\text{cl}}{\surd}_U a$.

Hence, it remains to show that $a \overset{\text{cl}}{\surd}_U a$ implies $a \in \text{cl}(U)$. Suppose that $a \overset{\text{cl}}{\surd}_U a$. Since the set $\{a\}$ is not independent over $U \cup \{a\}$, it follows that $\{a\}$ is not independent over U. Hence, $a \in \text{cl}(U)$.

(b) ($\Rightarrow$) Suppose that $A \subseteq \text{cl}(U \cup B)$ and $B \overset{\text{cl}}{\surd}_U C$. To show that $A \overset{\text{cl}}{\surd}_U C$, consider a set $I \subseteq C$ that is independent over U. Then I is also independent over $U \cup B$ and, hence, over $\text{cl}(U \cup B)$. In particular, I is independent over $U \cup A \subseteq \text{cl}(U \cup B)$.

($\Leftarrow$) Suppose that $A \nsubseteq \text{cl}(U \cup B)$ and fix an element $a \in A \setminus \text{cl}(U \cup B)$. Then $B \overset{\text{cl}}{\surd}_U a$ since $\varnothing$ and $\{a\}$ are both independent over U and independent over $U \cup B$. But $A \overset{\text{cl}}{\not\surd}_U a$ since $\{a\}$ is independent over U, but not over $U \cup A$. □

We use the characterisation in (a) to associate a closure operator with an arbitrary abstract independence relation $\surd$.

Definition 2.10. Let $\surd$ be an abstract independence relation on the set Ω. For $U \subseteq \Omega$, we define

$$\text{cl}_{\surd}(U) := \{ a \in \Omega \mid a \surd_{U \cup C} B \text{ for all } B, C \subseteq \Omega \}.$$

Let us start by proving that this definition results in a closure operator. The main technical argument is contained in the following lemma.

Lemma 2.11. *Let $\surd$ be an abstract independence relation on the set Ω.*

$$A \subseteq \mathrm{cl}_{\surd}(U) \quad \text{iff} \quad A \surd_{U \cup C} B \quad \text{for all } B, C \subseteq \Omega\,.$$

Proof. ($\Leftarrow$) Let $a \in A$. Then $a \surd_{U \cup C} B$, for all sets B, C. Consequently, $a \in \mathrm{cl}_{\surd}(U)$.

($\Rightarrow$) By (FIN), it is sufficient to prove the claim for finite sets A. We proceed by induction on $|A|$. For $A = \emptyset$ and arbitrary sets $B, C \subseteq \Omega$, $U \cup C \surd_{U \cup C} B$ implies that $\emptyset \surd_{U \cup C} B$, as desired.

Hence, suppose that $A = A_0 \cup \{a\}$ and that we have already shown that $A_0 \surd_{U \cup C} B$, for all sets B, C. Given $B, C \subseteq \Omega$, it follows that $A_0 \surd_{U \cup C \cup \{a\}} B$ and $a \surd_{U \cup C} B$ which, by Lemma 2.3, implies that $A_0 \cup \{a\} \surd_{U \cup C} B$. □

Corollary 2.12. *Let $\surd$ be an abstract independence relation on the set Ω.*

$$\mathrm{cl}_{\surd}(U) \surd_{U \cup C} B\,, \quad \text{for all } B, C, U \subseteq \Omega\,.$$

Proposition 2.13. *Let $\surd$ be an abstract independence relation on the set Ω. Then $\mathrm{cl}_{\surd}$ is a closure operator on Ω.*

Proof. To show that $U \subseteq \mathrm{cl}_{\surd}(U)$, consider $a \in U$ and $B, C \subseteq \Omega$. Then $U \cup C \surd_{U \cup C} B$ implies $a \surd_{U \cup C} B$. Hence, $a \in \mathrm{cl}_{\surd}(U)$.

For monotonicity, let $U \subseteq V$ and suppose that $a \surd_{U \cup C} B$, for all $B, C \subseteq \Omega$. Given $B, C \subseteq \Omega$, we have $a \surd_{U \cup V \cup C} B$. Hence, $\mathrm{cl}_{\surd}(U) \subseteq \mathrm{cl}_{\surd}(V)$.

To show that $\mathrm{cl}_{\surd}(\mathrm{cl}_{\surd}(U)) = \mathrm{cl}_{\surd}(U)$, fix an element $a \in \mathrm{cl}_{\surd}(\mathrm{cl}_{\surd}(U))$ and sets $B, C \subseteq \Omega$. Then

$$a \surd_{\mathrm{cl}_{\surd}(U) \cup \mathrm{cl}_{\surd}(U \cup C)} B\,.$$

Since we have already shown that $\mathrm{cl}_{\surd}$ is monotone, we have $\mathrm{cl}_{\surd}(U) \subseteq \mathrm{cl}_{\surd}(U \cup C)$ and it follows that $a \surd_{\mathrm{cl}_{\surd}(U \cup C)} B$. Furthermore, according to Corollary 2.12, $\mathrm{cl}_{\surd}(U \cup C) \surd_{U \cup C} B$. By Lemma 2.3 and monotonicity, it therefore follows that $a \surd_{U \cup C} B$. Hence, $a \in \mathrm{cl}_{\surd}(U)$. □

For symmetric independence relations we have the following desirable relationship to the associated closure operator.

Lemma 2.14. *Let* $\downarrow$ *be an abstract independence relation on the set* Ω *satisfying* (SYM) *and* (BMON).

$$A \downarrow_U B \quad \text{iff} \quad \mathrm{cl}_\downarrow(A) \downarrow_{\mathrm{cl}_\downarrow(U)} \mathrm{cl}_\downarrow(B)\,, \quad \textit{for all } A, B, U \subseteq \Omega\,.$$

Proof. ($\Leftarrow$) By Corollary 2.12, we have $\mathrm{cl}_\downarrow(U) \downarrow_U \mathrm{cl}_\downarrow(B)$. Therefore, $\mathrm{cl}_\downarrow(A) \downarrow_{\mathrm{cl}_\downarrow(U)} \mathrm{cl}_\downarrow(B)$ implies $\mathrm{cl}_\downarrow(A) \downarrow_U \mathrm{cl}_\downarrow(B)$, by Lemma 2.3. Hence, the claim follows by (MON).

($\Rightarrow$) Suppose that $A \downarrow_U B$. Then $A \cup U \downarrow_U B$. We have shown in Corollary 2.12 that $\mathrm{cl}_\downarrow(A \cup U) \downarrow_{A \cup U} B$. Using (LTR) we see that $\mathrm{cl}_\downarrow(A \cup U) \downarrow_U B$. By symmetry, it follows in exactly the same way that $\mathrm{cl}_\downarrow(A \cup U) \downarrow_U \mathrm{cl}_\downarrow(B \cup U)$. Hence, we can use (BMON) and (MON) to show that $\mathrm{cl}_\downarrow(A) \downarrow_{\mathrm{cl}_\downarrow(U)} \mathrm{cl}_\downarrow(B)$. □

If an abstract independence relation $\surd$ is induced by a closure operator, we obtain this operator back if we form $\mathrm{cl}_\surd$.

Lemma 2.15. $\mathrm{cl} = \mathrm{cl}_{\overset{\mathrm{cl}}{\surd}}$, *for every finitary closure operator* cl.

Proof. By definition of $\mathrm{cl}_{\overset{\mathrm{cl}}{\surd}}$ and Lemma 2.9,

$$\begin{aligned} a \in \mathrm{cl}_{\overset{\mathrm{cl}}{\surd}}(U) \quad &\text{iff} \quad a \overset{\mathrm{cl}}{\surd}_{U \cup C} B \text{ for all sets } B, C \\ &\text{iff} \quad a \in \mathrm{cl}(U)\,. \end{aligned}$$
□

Remark. Note that, in general, the dual statement does not hold: there are distinct independence relations inducing the same closure operator.

For a geometric independence relation $\downarrow$, we not only obtain a closure operator, but even a matroid. Again, we begin with two technical lemmas.

Lemma 2.16. *Let* $\downarrow$ *be a geometric independence relation. Then*

$$a \not\downarrow_U B \quad \text{iff} \quad a \in \mathrm{cl}_\downarrow(U \cup B) \smallsetminus \mathrm{cl}_\downarrow(U)\,.$$

Proof. (⇐) Suppose that $a \in \mathrm{cl}_{\downarrow}(U \cup B)$ and $a \downarrow_U B$. We have to show that $a \in \mathrm{cl}_{\downarrow}(U)$. Hence, let $C, D \subseteq \Omega$ be arbitrary sets. Then $a \downarrow_{U \cup B} C \cup D$ and $a \downarrow_U B$ implies, by Lemma 2.3 and symmetry, that $a \downarrow_U C \cup D$. Consequently, we have $a \downarrow_{U \cup C} D$ by (BMON).

(⇒) Suppose that $a \not\downarrow_U B$. Then $a \notin \mathrm{cl}_{\downarrow}(U)$. For a contradiction, assume that there are sets C, D such that $a \not\downarrow_{U \cup B \cup C} D$. Then (MON) implies

$$a \not\downarrow_U U \cup B \cup C \quad \text{and} \quad a \not\downarrow_{U \cup B \cup C} U \cup B \cup C \cup D\,.$$

By (SRB), it follows that $2 \leq |\{a\}| = 1$. A contradiction. □

Lemma 2.17. *Let $\downarrow$ be a geometric independence relation on Ω. For all $a \in \Omega$ and $B \subseteq \Omega$, there exists a finite set $B_0 \subseteq B$ such that $a \downarrow_{B_0} B$.*

Proof. We prove the claim by induction on $\kappa := |B|$. For $\kappa < \aleph_0$, we have $a \downarrow_B B$ by (LRF) and symmetry. Hence, suppose that $\kappa \geq \aleph_0$. Let $(b_\alpha)_{\alpha<\kappa}$ be an enumeration of B and set $B_\alpha := \{ b_i \mid i < \alpha \}$, for $\alpha \leq \kappa$. If $a \downarrow_\varnothing B$, we are done. Otherwise, let α be the minimal ordinal such that $a \not\downarrow_\varnothing B_\alpha$. By Lemma 2.16, it follows that $a \in \mathrm{cl}_{\downarrow}(B_\alpha)$. Consequently, $a \downarrow_{B_\alpha} B$. Note that $\alpha < \kappa$ since $a \downarrow_\varnothing B_\beta$ for all $\beta < \kappa$ would imply, by Lemma 2.5 and symmetry, that $a \downarrow_\varnothing B$. Hence $|B_\alpha| = |\alpha| < \kappa$, and we can apply the inductive hypothesis to find a finite set $U \subseteq B_\alpha$ with $a \downarrow_U B_\alpha$. Consequently, it follows by (LTR) and symmetry that $a \downarrow_U B$. □

Proposition 2.18. *If $\downarrow$ is a geometric independence relation on the set Ω, then $\langle \Omega, \mathrm{cl}_{\downarrow} \rangle$ is a matroid.*

Proof. We have already seen in Proposition 2.13 that $\mathrm{cl}_{\downarrow}$ is a closure operator. Hence, it remains to check that it has finite character and the exchange property.

For finite character, suppose that $a \in \mathrm{cl}_{\downarrow}(U)$. By Lemma 2.17 we can find a finite set $U_0 \subseteq U$ such that $a \downarrow_{U_0} U$. For all sets B, C it follows by $a \downarrow_U B \cup C$, Lemma 2.3, and (SYM) that $a \downarrow_{U_0} B \cup C$. Hence, (BMON) implies $a \downarrow_{U_0 \cup C} B$ and we have $a \in \mathrm{cl}_{\downarrow}(U_0)$.

It remains to check the exchange property. Suppose that

$$b \in \mathrm{cl}_{\downarrow}(U \cup \{a\}) \smallsetminus \mathrm{cl}_{\downarrow}(U)\,.$$

By Lemma 2.16, it follows that $b \not\downarrow_U a$. By symmetry, we have $a \not\downarrow_U b$ and we can use Lemma 2.16 again to conclude that

$$a \in \mathrm{cl}_{\downarrow}(U \cup \{b\}) \smallsetminus \mathrm{cl}_{\downarrow}(U)\,. \qquad \square$$

The next lemma, together with Lemma 2.15, shows that the operation $\mathrm{cl} \mapsto \overset{\mathrm{cl}}{\downarrow}$ is a bijective function from the class of all matroids to the class of all geometric independence relations. Its inverse is given by the function $\downarrow \mapsto \mathrm{cl}_{\downarrow}$.

Lemma 2.19. *If $\downarrow$ is a geometric independence relation then $\overset{\mathrm{cl}_\downarrow}{\downarrow} = \downarrow$.*

Proof. ($\supseteq$) Suppose that $A \overset{\mathrm{cl}_\downarrow}{\not\downarrow}_U B$. We have to show that $A \not\downarrow_U B$. By assumption, there exists a set $I \subseteq B$ that is $\mathrm{cl}_\downarrow$-independent over U, but not over $U \cup A$. Fix an element $b \in I$ such that $b \in \mathrm{cl}_\downarrow(U \cup A \cup I_0)$ where $I_0 := I \smallsetminus \{b\}$. Since $b \notin \mathrm{cl}_\downarrow(U \cup I_0)$, it follows by Lemma 2.16 that $b \not\downarrow_{U \cup I_0} A$. By monotonicity, this implies that $B \not\downarrow_{U \cup I_0} A$. Hence, we can use symmetry and (BMON) to deduce that $A \not\downarrow_U B$.

($\subseteq$) By (FIN) and symmetry, it is sufficient to show that $A \overset{\mathrm{cl}_\downarrow}{\downarrow}_U B$ implies $A \downarrow_U B$, for all finite sets A, B. Furthermore, we may assume by Lemmas 2.14 and 2.15 that A and B are $\mathrm{cl}_\downarrow$-independent over U. Hence, suppose that $A \overset{\mathrm{cl}_\downarrow}{\downarrow}_U B$ for finite sets A and B that are $\mathrm{cl}_\downarrow$-independent over U. We prove by induction on $|B|$ that $B \downarrow_U A$. If $B = \varnothing$, then $U \downarrow_U A$ implies $\varnothing \downarrow_U A$. Hence, suppose that $B = B_0 \cup \{b\}$ and that we have already shown that $B_0 \downarrow_U A$. Since B is $\mathrm{cl}_\downarrow$-independent over U, it is also $\mathrm{cl}_\downarrow$-independent over $U \cup A$. Hence, $b \notin \mathrm{cl}_\downarrow(U \cup A \cup B_0)$ and Lemma 2.16 implies that $b \downarrow_{U \cup B_0} A$. Together with $B_0 \downarrow_U A$ it follows by Lemma 2.3 that $B_0 \cup \{b\} \downarrow_U A$. $\square$

We conclude this section with a characterisation of modularity in terms of the independence relation $\overset{\mathrm{cl}}{\downarrow}$.

Proposition 2.20. *A matroid $\langle \Omega, \mathrm{cl}\rangle$ is modular if, and only if,*

$$A \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{cl}}_{\mathrm{cl}(A)\cap\mathrm{cl}(B)} B\,, \quad \text{for all } A, B \subseteq \Omega\,.$$

Proof. ($\Rightarrow$) Suppose that $\langle \Omega, \mathrm{cl}\rangle$ is modular and let $A, B \subseteq \Omega$. We have to show that $A \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{cl}}_{\mathrm{cl}(A)\cap\mathrm{cl}(B)} B$. By Lemmas 2.14 and 2.15, we may assume that A and B are closed sets. Hence, let A and B be closed and $I \subseteq B$ independent over $A \cap B$. Let $I_0 \subseteq I$ be a basis of I over A and set $C_0 := \mathrm{cl}(I_0)$ and $C := \mathrm{cl}(I)$. We have to show that $I_0 = I$. Note that

$$\mathrm{cl}(C_0 \cup A) = \mathrm{cl}(I_0 \cup A) = \mathrm{cl}(I \cup A) = \mathrm{cl}(C \cup A)\,.$$

By Lemma B2.2.9, it follows that

$$C = \mathrm{cl}(C_0 \cup (C \cap A)) = \mathrm{cl}(I_0 \cup (C \cap A))\,.$$

Hence, I_0 is a basis of C over $C \cap A$. Since $I \supseteq I_0$ is independent over $C \cap A$, it follows that $I = I_0$ and I is independent over A.

($\Leftarrow$) Suppose that $A \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{cl}}_{\mathrm{cl}(A)\cap\mathrm{cl}(B)} B$, for all $A, B \subseteq \Omega$. To show that $\langle \Omega, \mathrm{cl}\rangle$ is modular it is sufficient, by Lemma B2.2.9, to prove that

$$\mathrm{cl}(A \cup C) = \mathrm{cl}(B \cup C) \quad \text{implies} \quad \mathrm{cl}(A \cup (B \cap C)) = B\,,$$

for all closed sets $A, B, C \subseteq \Omega$ with $A \subseteq B$. Hence, fix closed sets $A, B, C \subseteq \Omega$ with $A \subseteq B$ and $\mathrm{cl}(A \cup C) = \mathrm{cl}(B \cup C)$. Choose a maximal set $I \subseteq A$ that is independent over C. Then $\mathrm{cl}(I \cup C) = \mathrm{cl}(A \cup C) = \mathrm{cl}(B \cup C)$ and I is a basis of $B \cup C$ over C. We claim that $B \subseteq \mathrm{cl}(I \cup (B \cap C))$. Suppose otherwise. Then there is some element $b \in B \setminus \mathrm{cl}(I \cup (B \cap C))$. Since $b \in B \subseteq \mathrm{cl}(I \cup C)$ and $b \notin \mathrm{cl}(I \cup (B \cap C))$, it follows that $I \cup \{b\}$ is independent over $B \cap C$, but not over C. Hence, $C \not\mathop{\smile\!\!\!\!\!\!|}^{\mathrm{cl}}_{B\cap C} B$. A contradiction.

We have shown that $B \subseteq \mathrm{cl}(I \cup (B \cap C))$. It follows that

$$B \subseteq \mathrm{cl}(I \cup (B \cap C)) \subseteq \mathrm{cl}(A \cup (B \cap C)) \subseteq B\,,$$

as desired. □

Corollary 2.21. *Let $\langle \Omega, \mathrm{cl} \rangle$ be a modular matroid. Then*

$$A \mathop{\sqrt{}}^{\mathrm{cl}}_{U} B \quad \text{iff} \quad \mathrm{cl}(A \cup U) \cap \mathrm{cl}(B \cup U) = \mathrm{cl}(U) .$$

Proof. ($\Leftarrow$) According to Proposition 2.20, we have

$$A \cup U \mathop{\sqrt{}}^{\mathrm{cl}}_{\mathrm{cl}(A \cup U) \cap \mathrm{cl}(B \cup U)} B \cup U .$$

If $\mathrm{cl}(A \cup U) \cap \mathrm{cl}(B \cup U) = \mathrm{cl}(U)$, then

$$A \cup U \mathop{\sqrt{}}^{\mathrm{cl}}_{\mathrm{cl}(U)} B \cup U \quad \text{implies} \quad A \mathop{\sqrt{}}^{\mathrm{cl}}_{U} B ,$$

by Lemma 2.14.

($\Rightarrow$) Suppose that $A \mathop{\sqrt{}}^{\mathrm{cl}}_{U} B$. By Lemma 2.14, it follows that

$$\mathrm{cl}(A \cup U) \mathop{\sqrt{}}^{\mathrm{cl}}_{\mathrm{cl}(U)} \mathrm{cl}(B \cup U) .$$

For a contradiction, suppose that there is some element

$$c \in (\mathrm{cl}(A \cup U) \cap \mathrm{cl}(B \cup U)) \smallsetminus \mathrm{cl}(U) .$$

Then $\{c\}$ is independent over $\mathrm{cl}(U)$, but not over $\mathrm{cl}(A \cup U)$. Hence, $\mathrm{cl}(A \cup U) \mathop{\not\sqrt{}}^{\mathrm{cl}}_{\mathrm{cl}(U)} \mathrm{cl}(B \cup U)$. A contradiction. □

3. *Preforking relations*

We would like to define an independence relation using Δ-rank or Morley rank as our notion of dimension. In general, the resulting relation will not be a geometric independence relation but something slightly weaker, called a *forking relation.* In this section, we introduce the abstract framework for forking relations and we will present several examples of such relations. To simplify notation, we will frequently omit union symbols and just write AB instead of $A \cup B$.

Definition 3.1. Let T be a complete first-order theory and suppose that $A \mathop{\surd}_U B$ is a ternary relation that is defined on the class of all small subsets $A, B, U \subseteq \mathbb{M}$.

(a) The relation $\surd$ is a *preforking relation* for T if it is an abstract independence relation that satisfies (BMON) and the following two axioms:

(INV) *Invariance.* $ABU \equiv_\varnothing A'B'U'$ implies that

$$A \mathop{\surd}_U B \quad \text{iff} \quad A' \mathop{\surd}_{U'} B' .$$

(DEF) *Definability.* If $A \mathop{\not\surd}_U B$, there are finite tuples $\bar{a} \subseteq A$ and $\bar{b} \subseteq B$ and a formula $\varphi(\bar{x}, \bar{x}') \in \mathrm{tp}(\bar{a}\bar{b}/U)$ such that

$$\bar{a}' \mathop{\not\surd}_U \bar{b}, \quad \text{for all } \bar{a}' \in \varphi(\bar{x}, \bar{b})^{\mathbb{M}} .$$

(b) The relation $\surd$ is a *forking relation* if it is a preforking relation that satisfies the following additional axiom:

(EXT) *Extension.* If $A \mathop{\surd}_U B_0$ and $B_0 \subseteq B_1$ then there is some A' with

$$A' \equiv_{UB_0} A \quad \text{and} \quad A' \mathop{\surd}_U B_1 .$$

We are mostly interested in symmetric forking relations since many properties of geometric independence relations can be generalised to them. Unfortunately, there are first-order theories were no nontrivial symmetric forking relations exist. On the other hand there are always several natural preforking relations and below we will see that every preforking relation can be used to define a corresponding forking relation, although not necessarily a symmetric one.

Remark. The intersection of an arbitrary family of preforking relations is again a preforking relation. It follows that the class of all preforking relations on a structure $\mathbb{M}$ forms a complete partial order.

Examples

Before proceeding let us collect several examples. We start with a trivial one.

Example. The trivial relation $\surd$ with $A \surd_U B$, for all sets A, B, U, is a symmetric forking relation.

Exercise 3.1. Prove that the relation

$$A \downarrow^{\circ}_{U} B \quad :\text{iff} \quad A \cap B \subseteq U$$

is a symmetric preforking relation.

More interesting are the following three examples. The second one has historically been used to develop stability theory.

Definition 3.2. For $\bar{a}, A, B, U \subseteq \mathbb{M}$, we define

$$\begin{aligned}
A \overset{\mathrm{at}}{\surd}_U B \quad &:\text{iff} \quad \text{for every finite } \bar{a} \subseteq A, \\
&\qquad\quad \mathrm{tp}(\bar{a}/UB) \text{ is isolated by a formula over } U. \\
\bar{a} \overset{\mathrm{df}}{\surd}_U B \quad &:\text{iff} \quad \mathrm{tp}(\bar{a}/UB) \text{ is definable over } U. \\
A \overset{\mathrm{s}}{\surd}_U B \quad &:\text{iff} \quad \bar{b} \equiv_U \bar{b}' \;\Rightarrow\; \bar{b} \equiv_{UA} \bar{b}', \quad \text{for all } \bar{b}, \bar{b}' \subseteq B.
\end{aligned}$$

If $\bar{a} \overset{\mathrm{s}}{\surd}_U B$, we say that the type $\mathrm{tp}(\bar{a}/UB)$ is *invariant* over U. Otherwise, it *splits* over U.

Lemma 3.3.

(a) $\overset{\mathrm{at}}{\surd} \subseteq \overset{\mathrm{df}}{\surd} \subseteq \overset{\mathrm{s}}{\surd}$

(b) $\overset{\mathrm{at}}{\surd}$ *is an abstract independence relation that satisfies* (INV) *and* (BMON).

(c) $\overset{\mathrm{df}}{\surd}$ *is an abstract independence relation that satisfies* (INV) *and* (BMON).

(d) $\overset{\mathrm{s}}{\surd}$ *is a preforking relation.*

Proof. (a) Suppose that $A \overset{\mathrm{at}}{\surd}_U B$ and let $\bar{a}$ be an enumeration of A. To show that $A \overset{\mathrm{df}}{\surd}_U B$, consider a formula $\varphi(\bar{x}; \bar{b}) \in \mathrm{tp}(\bar{a}/UB)$. Let $\bar{a}_0 \subseteq \bar{a}$

be the finite tuple of elements mentioned in φ. By assumption, there is a formula $\psi(\bar{x})$ over U isolating $\mathrm{tp}(\bar{a}_0/UB)$. It follows that

$$\delta(\bar{y}) := \forall\bar{x}[\psi(\bar{x}) \rightarrow \varphi(\bar{x};\bar{y})]$$

is a φ-definition of $\mathrm{tp}(\bar{a}/UB)$.

For the second inclusion, suppose that $A \overset{\mathrm{df}}{\mathop{\sqrt{}}}_U B$. Let $\bar{b}, \bar{b}' \subseteq B$ be tuples with $\bar{b} \equiv_{UA} \bar{b}'$. We have to show that $\bar{b} \equiv_U \bar{b}'$. Fix a formula $\varphi(\bar{x};\bar{a},\bar{c})$ with parameters $\bar{a} \subseteq A$ and $\bar{c} \subseteq U$ such that

$$\mathbb{M} \vDash \varphi(\bar{b};\bar{a},\bar{c}) \wedge \neg\varphi(\bar{b}';\bar{a},\bar{c}) .$$

By assumption, $\mathrm{tp}(\bar{a}/UB)$ has a φ-definition $\delta(\bar{x})$ over U. It follows that $\mathbb{M} \vDash \delta(\bar{b}) \wedge \neg\delta(\bar{b}')$. Consequently, $\bar{b} \not\equiv_U \bar{b}'$.

(b) (INV) and (FIN) follow immediately from the definition.

(MON) Suppose that $A \overset{\mathrm{at}}{\mathop{\sqrt{}}}_U B$ and let $A_0 \subseteq A$, $B_0 \subseteq B$. For $\bar{a} \subseteq A_0$ we know that $\mathrm{tp}(\bar{a}/UB)$ is isolated by a formula over U. Hence, so is $\mathrm{tp}(\bar{a}/UB_0)$.

(NOR) Suppose that $A \overset{\mathrm{at}}{\mathop{\sqrt{}}}_U B$. Let $\bar{a} \subseteq A \cup U$ be finite. Then $\bar{a} = \bar{a}' \cup \bar{c}$ for $\bar{a}' \subseteq A$ and $\bar{c} \subseteq U$. Furthermore, $\mathrm{tp}(\bar{a}'/UB)$ is isolated by a formula $\varphi(\bar{x})$ over U and $\mathrm{tp}(\bar{c}/UB)$ is isolated by the formula $\bar{x} = \bar{c}$. Consequently, $\mathrm{tp}(\bar{a}'\bar{c}/UB)$ is isolated by $\psi(\bar{x},\bar{x}') := \varphi(\bar{x}) \wedge \bar{x}' = \bar{c}$.

(LRF) If $\bar{a} \subseteq A$ is finite then $\mathrm{tp}(\bar{a}/AB)$ is isolated by the formula $\bar{x} = \bar{a}$. Hence, $A \overset{\mathrm{at}}{\mathop{\sqrt{}}}_A B$.

(LTR) Suppose that $A_2 \overset{\mathrm{at}}{\mathop{\sqrt{}}}_{A_1} B$ and $A_1 \overset{\mathrm{at}}{\mathop{\sqrt{}}}_{A_0} B$ for $A_0 \subseteq A_1 \subseteq A_2$. Let $\bar{a} \subseteq A_2$ be finite. Then $\mathrm{tp}(\bar{a}/A_1B)$ is isolated by a formula $\varphi(\bar{x};\bar{c})$ with parameters $\bar{c} \subseteq A_1$. Furthermore, $\mathrm{tp}(\bar{c}/A_0B)$ is isolated by a formula $\psi(\bar{x})$ over A_0. By Lemma E3.1.5, it follows that $\mathrm{tp}(\bar{a}\bar{c}/A_0B)$ is isolated by the formula $\varphi(\bar{x};\bar{z}) \wedge \psi(\bar{z})$. Therefore, $\mathrm{tp}(\bar{a}/A_0B)$ is isolated by the formula $\exists\bar{z}[\varphi(\bar{x};\bar{z}) \wedge \psi(\bar{z})]$.

(BMON) Suppose that $A \overset{\mathrm{at}}{\mathop{\sqrt{}}}_U BC$. For every $\bar{a} \subseteq A$, $\mathrm{tp}(\bar{a}/UBC)$ is isolated by a formula over U and, hence, by a formula over $U \cup C$.

(c) (INV) follows immediately from the definition.

(MON) Suppose that $\bar{a} \overset{\mathrm{df}}{\smile}_U B$. If $\bar{a}_0 \subseteq \bar{a}$ and $B_0 \subseteq B$ then

$$\mathrm{tp}(\bar{a}_0/UB_0) \subseteq \mathrm{tp}(\bar{a}/UB)$$

and every φ-definition of the latter type is also a φ-definition of the former one.

(NOR) Suppose that $\mathrm{tp}(\bar{a}/B\bar{c})$ is definable over $\bar{c}$. To find the desired $\varphi(\bar{x},\bar{x}';\bar{y})$-definition of $\mathrm{tp}(\bar{a}\bar{c}/B\bar{c})$ over $\bar{c}$, let $\psi(\bar{y},\bar{y}';\bar{c})$ be a $\varphi(\bar{x};\bar{y}',\bar{y})$-definition of $\mathrm{tp}(\bar{a}/B\bar{c})$ over $\bar{c}$. For $\bar{b} \subseteq B \cup \bar{c}$ it follows that

$$\mathbb{M} \vDash \varphi(\bar{a},\bar{c};\bar{b}) \quad \text{iff} \quad \mathbb{M} \vDash \psi(\bar{b},\bar{c};\bar{c})\,.$$

Hence, $\psi(\bar{y},\bar{c};\bar{c})$ is a φ-definition of $\mathrm{tp}(\bar{a}\bar{c}/B)$ over $\bar{c}$.

(LRF) Note that $\varphi(\bar{a};\bar{y})$ is a $\varphi(\bar{x};\bar{y})$-definition of $\mathrm{tp}(\bar{a}/B \cup \bar{a})$. Hence, $\mathrm{tp}(\bar{a}/B\bar{a})$ is definable over $\bar{a}$.

(LTR) Suppose that $\bar{a}_0\bar{a}_1\bar{a}_2 \overset{\mathrm{df}}{\smile}_{\bar{a}_0\bar{a}_1} B$ and $\bar{a}_0\bar{a}_1 \overset{\mathrm{df}}{\smile}_{\bar{a}_0} B$. For every formula $\varphi(\bar{x}_0,\bar{x}_1,\bar{x}_2;\bar{y})$, there exist

- a φ-definition $\psi(\bar{y};\bar{a}_0,\bar{a}_1)$ of $\mathrm{tp}(\bar{a}_0\bar{a}_1\bar{a}_2/B\bar{a}_0\bar{a}_1)$ over $\bar{a}_0\bar{a}_1$, and
- a $\psi(\bar{y};\bar{x}_0,\bar{x}_1)$-definition $\vartheta(\bar{y};\bar{a}_0)$ of $\mathrm{tp}(\bar{a}_0\bar{a}_1/B\bar{a}_0)$ over $\bar{a}_0$.

For $\bar{b} \subseteq B \cup \bar{a}_0$, we have

$$\begin{aligned}\mathbb{M} \vDash \varphi(\bar{a}_0,\bar{a}_1,\bar{a}_2;\bar{b}) \quad &\text{iff} \quad \mathbb{M} \vDash \psi(\bar{b};\bar{a}_0,\bar{a}_1)\\ &\text{iff} \quad \mathbb{M} \vDash \vartheta(\bar{b};\bar{a}_0)\,.\end{aligned}$$

Hence, ϑ is a φ-definition of $\mathrm{tp}(\bar{a}_0\bar{a}_1\bar{a}_2/B\bar{a}_0)$ over $\bar{a}_0$.

(BMON) Clearly, every φ-definition of $\mathrm{tp}(\bar{a}/UBC)$ over U is also a φ-definition of $\mathrm{tp}(\bar{a}/UBC)$ over $U \cup C$.

(FIN) Since each formula $\varphi(\bar{x}) \in \mathrm{tp}(\bar{a}/UB)$ contains only finitely many variables from $\bar{x}$, it follows that $\mathrm{tp}(\bar{a}/UB)$ is definable over U if, and only if, $\mathrm{tp}(\bar{a}_0/UB)$ is definable over U, for all finite $\bar{a}_0 \subseteq \bar{a}$.

(d) (INV) follows immediately from the definition.

(MON) Suppose that $A \sqrt[s]{}_U B$ and let $A_0 \subseteq A$ and $B_0 \subseteq B$. For $\bar{b}, \bar{b}' \subseteq B_0$ it follows that

$$\bar{b} \equiv_U \bar{b}' \quad \text{implies} \quad \bar{b} \equiv_{UA} \bar{b}' \quad \text{implies} \quad \bar{b} \equiv_{UA_0} \bar{b}' .$$

Hence, $A_0 \sqrt[s]{}_U B_0$.

(NOR) Suppose that $A \sqrt[s]{}_U B$. If $\bar{b}, \bar{b}' \subseteq B \cup U$ are tuples such that $\bar{b} \equiv_U \bar{b}'$, then there are tuples $\bar{b}_0, \bar{b}'_0 \subseteq B$ and $\bar{c} \subseteq U$ such that $\bar{b} = \bar{b}_0 \cup \bar{c}$ and $\bar{b}' = \bar{b}'_0 \cup \bar{c}$. It follows that

$$\begin{aligned} \bar{b} \equiv_U \bar{b}' \quad &\text{implies} \quad \bar{b}_0 \equiv_U \bar{b}'_0 \\ &\text{implies} \quad \bar{b}_0 \equiv_{UA} \bar{b}'_0 \\ &\text{implies} \quad \bar{b}_0\bar{c} \equiv_{UA} \bar{b}'_0\bar{c} \quad \text{implies} \quad \bar{b} \equiv_{UA} \bar{b}' . \end{aligned}$$

Consequently, $AU \sqrt[s]{}_U BU$.

(LRF) Since, trivially, $\bar{b} \equiv_A \bar{b}'$ implies $\bar{b} \equiv_A \bar{b}'$, we have $A \sqrt[s]{}_A B$, for all sets A and B.

(LTR) Suppose that $A_2 \sqrt[s]{}_{A_1} B$ and $A_1 \sqrt[s]{}_{A_0} B$, for $A_0 \subseteq A_1 \subseteq A_2$. For $\bar{b}, \bar{b}' \subseteq B$ it follows that

$$\bar{b} \equiv_{A_0} \bar{b} \quad \text{implies} \quad \bar{b} \equiv_{A_1} \bar{b} \quad \text{implies} \quad \bar{b} \equiv_{A_2} \bar{b} .$$

as desired.

(BMON) Suppose that $A \sqrt[s]{}_U BC$. Let $\bar{b}, \bar{b}' \subseteq B \cup C$ be tuples such that $\bar{b} \not\equiv_{UAC} \bar{b}'$. We claim that $\bar{b} \not\equiv_{UC} \bar{b}'$. There exists a formula $\varphi(\bar{x}; \bar{a}, \bar{c}, \bar{d})$ with parameters $\bar{a} \subseteq A$, $\bar{c} \subseteq C$, and $\bar{d} \subseteq U$ such that

$$\mathbb{M} \vDash \varphi(\bar{b}; \bar{a}, \bar{c}, \bar{d}) \wedge \neg\varphi(\bar{b}'; \bar{a}, \bar{c}, \bar{d}) .$$

Consequently, $\bar{b}\bar{c} \not\equiv_{UA} \bar{b}'\bar{c}$. Since $A \sqrt[s]{}_U BC$ it follows that $\bar{b}\bar{c} \not\equiv_U \bar{b}'\bar{c}$. As $\bar{c} \subseteq C$ this means that $\bar{b} \not\equiv_{UC} \bar{b}'$, as desired.

(DEF) Suppose that $A \not\sqrt[s]{}_U B$. Then there exist tuples $\bar{b}, \bar{b}' \subseteq B$ such that $\bar{b} \equiv_U \bar{b}'$ and $\bar{b} \not\equiv_{UA} \bar{b}'$. Fix a formula $\varphi(\bar{x}, \bar{y})$ over U and a tuple

$\bar{a} \subseteq A$ such that

$$\mathbb{M} \vDash \varphi(\bar{a}, \bar{b}) \wedge \neg\varphi(\bar{a}, \bar{b}') .$$

For every tuple $\bar{a}' \subseteq \mathbb{M}$ it follows that

$$\mathbb{M} \vDash \varphi(\bar{a}', \bar{b}) \wedge \neg\varphi(\bar{a}', \bar{b}') \quad \text{implies} \quad \bar{a}' \overset{\mathrm{s}}{\not\vee\!\!/}_U \bar{b}\bar{b}' . \qquad \square$$

Let us mention that, in general, $\overset{\mathrm{df}}{\vee\!\!/}$ and $\overset{\mathrm{at}}{\vee\!\!/}$ are no preforking relations.

Example. (a) The relation $\overset{\mathrm{df}}{\vee\!\!/}$ is not definable. As a counterexample, consider the theory T of dense linear orders. Note that T has quantifier elimination. Let $a \in \mathbb{R} \setminus \mathbb{Q}$ be an irrational number. Then $\mathrm{tp}(a/\mathbb{Q})$ is not definable over $\mathbb{Q}$. Consider a formula $\varphi(x; \bar{b}) \in \mathrm{tp}(a/\mathbb{Q})$ with rational parameters $b_0 < \dots < b_{n-1}$. By enlarging the tuple $\bar{b}$ we may assume that there is some index i such that $b_i < a < b_{i+1}$. It follows that $\langle \mathbb{R}, \leq \rangle \vDash \varphi(a'; \bar{b})$, for all $a' \in (b_i, b_{i+1})$. But for $a' \in (b_i, b_{i+1}) \cap \mathbb{Q}$ the type $\mathrm{tp}(a'/\mathbb{Q})$ is definable over $\mathbb{Q}$. This contradicts (DEF).

(b) The relation $\overset{\mathrm{at}}{\vee\!\!/}$ is not definable. As a counterexample, consider the theory T of the structure $\langle \mathbb{R}, s \rangle$ where $s(x) = x + 1$. Note that $\mathrm{tp}(a/b)$ is isolated if, and only if, $a = b + k$, for some $k \in \mathbb{Z}$. In particular $\mathrm{tp}(\tfrac{1}{2}/0)$ is not isolated. Using an Ehrenfeucht-Fraïssé argument, one can show that, for every formula $\varphi(x; y)$ with $\langle \mathbb{R}, s \rangle \vDash \varphi(\tfrac{1}{2}; 0)$, there exists a number $a \in \mathbb{R}$ such that $\langle \mathbb{R}, s \rangle \vDash \varphi(b; 0)$, for all $b \geq a$. But, for $b \in \mathbb{N}$, the type $\mathrm{tp}(b/0)$ is isolated by the formula $x = s^b(0)$.

Let us take a look at the closure operators associated with these relations. In each case, we obtain the definable closure.

Lemma 3.4. $\mathrm{cl}_{\overset{\mathrm{df}}{\vee\!\!/}} = \mathrm{cl}_{\overset{\mathrm{at}}{\vee\!\!/}} = \mathrm{cl}_{\overset{\mathrm{s}}{\vee\!\!/}} = \mathrm{dcl}$

Proof. Note that $\overset{\mathrm{at}}{\vee\!\!/} \subseteq \overset{\mathrm{df}}{\vee\!\!/} \subseteq \overset{\mathrm{s}}{\vee\!\!/}$ implies $\mathrm{cl}_{\overset{\mathrm{at}}{\vee\!\!/}} \subseteq \mathrm{cl}_{\overset{\mathrm{df}}{\vee\!\!/}} \subseteq \mathrm{cl}_{\overset{\mathrm{s}}{\vee\!\!/}}$. Hence, we only need to prove that $\mathrm{dcl} \subseteq \mathrm{cl}_{\overset{\mathrm{at}}{\vee\!\!/}}$ and $\mathrm{cl}_{\overset{\mathrm{s}}{\vee\!\!/}} \subseteq \mathrm{dcl}$.

For the first inclusion, note that every formula defining a over U isolates $\mathrm{tp}(a/UBC)$. Hence, $a \in \mathrm{dcl}(U)$ implies $a \overset{\mathrm{at}}{\vee\!\!/}_{UC} B$, for all B, C.

For the second inclusion, consider an element $a \notin \mathrm{dcl}(U)$. By Theorem E2.1.6, there exists an automorphism $\pi \in \mathbb{M}_U$ with $\pi(a) \neq a$. Setting $a' := \pi(a)$ it follows that $a \equiv_U a'$ and $a \not\equiv_{Ua} a'$. Hence, $a \not\sqrt{}^{\mathrm{s}}_U aa'$ and $a \notin \mathrm{cl}_{\sqrt{}^{\mathrm{s}}}(U)$. ☐

We conclude this section with the remark that, for forking relations, the definition of the closure operator $\mathrm{cl}_{\sqrt{}}$ can be simplified.

Lemma 3.5. *If $\sqrt{}$ is a forking relation, then*

$$a \sqrt{}_U a \quad \textit{implies} \quad a \sqrt{}_{UC} B \quad \textit{for all } B, C\,.$$

Proof. Suppose that $a \sqrt{}_U a$ and let B, C be arbitrary sets. By (EXT), there exists an element $a' \equiv_{Ua} a$ with $a' \sqrt{}_U BC$. It follows that $a' = a$. Therefore, (BMON) implies $a \sqrt{}_{UC} B$. ☐

Finitely satisfiable types

Let us take a look at some consequences of the definability axiom (DEF). First, note that, by invariance, we can extend every preforking relation from subsets of $\mathbb{M}$ to types.

Definition 3.6. Let $\sqrt{}$ be a preforking relation and $B, U \subseteq \mathbb{M}$.

(a) A partial type $\Phi(\bar{x})$ over B $\sqrt{}$*-forks* over U if

$$\bar{a} \not\sqrt{}_U B\,, \quad \text{for all } \bar{a} \in \Phi^{\mathbb{M}}.$$

Similarly, we say that a single formula $\varphi(\bar{x})$ over B $\sqrt{}$-forks over U, if the type $\{\varphi\}$ does.

(b) A type $\mathfrak{p}$ over B is $\sqrt{}$*-free* over U if it does not $\sqrt{}$-fork over U.

(c) For complete types $\mathfrak{p} \in S^{\bar{s}}(U)$ and $\mathfrak{q} \in S^{\bar{s}}(UB)$, we say that $\mathfrak{q}$ is a $\sqrt{}$*-free extension* of $\mathfrak{p}$ if

$$\mathfrak{p} \subseteq \mathfrak{q} \quad \text{and} \quad \mathfrak{q} \text{ is } \sqrt{}\text{-free over } U\,.$$

We denote this fact by $\mathfrak{p} \leq_{\sqrt{}} \mathfrak{q}$.

Remark. (a) By (INV), we have $\bar{a} \mathop{\sqrt{}}_U B$ if, and only if, $\mathrm{tp}(\bar{a}/UB)$ is $\sqrt{}$-free over U.

(b) By (DEF), a complete type $\mathfrak{p}$ $\sqrt{}$-forks over U if, and only if, some formula $\varphi(\bar{x}) \in \mathfrak{p}$ $\sqrt{}$-forks over U.

Lemma 3.7. *Let $\sqrt{}$ be a preforking relation. The set*

$$F^{\bar{s}}_{\sqrt{}}(A/U) := \{\, \mathfrak{p} \in S^{\bar{s}}(A) \mid \mathfrak{p} \text{ is } \sqrt{}\text{-free over } U \,\}$$

is a closed subset of $\mathfrak{S}^{\bar{s}}(A)$.

Proof. Let

$$\Phi := \{\, \neg\varphi \mid \varphi \text{ a formula over } A \text{ that } \sqrt{}\text{-forks over } U \,\}\,.$$

Then $\Phi \subseteq \mathfrak{p}$, for every $\mathfrak{p} \in F^{\bar{s}}_{\sqrt{}}(A/U)$, while (DEF) implies that $\Phi \nsubseteq \mathfrak{p}$, for every type $\mathfrak{p}$ that $\sqrt{}$-forks over U. Hence,

$$F^{\bar{s}}_{\sqrt{}}(A/U) := \langle \Phi \rangle_{\mathfrak{S}^{\bar{s}}(A)}\,. \qquad \square$$

Let us treat in more detail one important forking relation that is connected with the definability axiom. It is based on the notion of a finitely satisfiable type.

Definition 3.8. A type $\mathfrak{p}$ is *finitely satisfiable* in a set U if, for every formula $\varphi(\bar{x};\bar{c}) \in \mathfrak{p}$, there is some tuple $\bar{a} \subseteq U$ with $\mathbb{M} \vDash \varphi(\bar{a};\bar{c})$. We write

$$\bar{a} \mathop{\overset{\mathrm{u}}{\sqrt{}}}_U B \quad :\text{iff} \quad \mathrm{tp}(\bar{a}/U \cup B) \text{ is finitely satisfiable in } U\,.$$

Example. Let T be the theory of dense linear orders. For a single element $a \in \mathbb{M}$ and sets $U, B \subseteq \mathbb{M}$, we have $a \mathop{\overset{\mathrm{u}}{\sqrt{}}}_U B$ if, and only if, at least one of the following conditions is satisfied:

- $a \in U$, or
- $\Uparrow a \cap U \neq \varnothing$ and, for every $b \in \Uparrow a \cap (U \cup B)$, there is some $c \in \Uparrow a \cap U$ with $c \leq b$, or

- $\Downarrow a \cap U \neq \varnothing$ and, for every $b \in \Downarrow a \cap (U \cup B)$, there is some $c \in \Downarrow a \cap U$ with $c \geq b$.

We shall prove that $\overset{\mathfrak{u}}{\surd}$ is the least preforking relation and that it is, in fact, a forking relation. Before doing so, let us give an alternative characterisation of finitely satisfiable types in terms of ultrafilters. (The letter 'u' in $\overset{\mathfrak{u}}{\surd}$ stands for 'ultrafilter'.)

Definition 3.9. Let T be a theory over the signature Σ, let $U, B \subseteq \mathbb{M}$ be sets, and $\mathfrak{u}$ an ultrafilter over $U^{\bar{s}}$, for some tuple $\bar{s}$ of sorts. The *average type* of $\mathfrak{u}$ over B is the set

$$\mathrm{Av}(\mathfrak{u}/B) := \left\{ \varphi(\bar{x}) \in \mathrm{FO}^{\bar{s}}[\Sigma_B] \;\middle|\; U^{\bar{s}} \cap \varphi(\bar{x})^{\mathbb{M}} \in \mathfrak{u} \right\}.$$

Lemma 3.10. *Let T be a complete first-order theory and $\mathfrak{u}$ an ultrafilter over $U^{\bar{s}}$. Then $\mathrm{Av}(\mathfrak{u}/B)$ is a complete type over B that is finitely satisfiable in U.*

Proof. We start by showing that $\mathrm{Av}(\mathfrak{u}/B)$ is a type. For a contradiction, suppose that $T \cup \mathrm{Av}(\mathfrak{u}/B)$ is unsatisfiable. Then there exist a finite subset $\Phi \subseteq \mathrm{Av}(\mathfrak{u}/B)$ such that $T \vDash \neg \bigwedge \Phi$. By definition of $\mathrm{Av}(\mathfrak{u}/B)$,

$$U^{\bar{s}} \cap \varphi^{\mathbb{M}} \in \mathfrak{u}, \quad \text{for all } \varphi \in \Phi.$$

As ultrafilters are closed under finite intersections, it follows that

$$U^{\bar{s}} \cap (\bigwedge \Phi)^{\mathbb{M}} \in \mathfrak{u}.$$

In particular, $(\bigwedge \Phi)^{\mathbb{M}} \neq \varnothing$. Hence, $T \vDash \exists \bar{x} \bigwedge \Phi$. A contradiction.

Moreover, $\mathrm{Av}(\mathfrak{u}/B)$ is complete since, for every formula $\varphi(\bar{x})$ over B,

$$\begin{aligned}
\varphi(\bar{x}) \in \mathrm{Av}(\mathfrak{u}/B) \quad &\text{iff} \quad U^{\bar{s}} \cap \varphi(\bar{x})^{\mathbb{M}} \in \mathfrak{u} \\
&\text{iff} \quad U^{\bar{s}} \setminus \varphi(\bar{x})^{\mathbb{M}} \notin \mathfrak{u} \\
&\text{iff} \quad U^{\bar{s}} \cap \neg\varphi(\bar{x})^{\mathbb{M}} \notin \mathfrak{u} \\
&\text{iff} \quad \neg\varphi(\bar{x}) \notin \mathrm{Av}(\mathfrak{u}/B).
\end{aligned}$$

Finally, to show that $\mathrm{Av}(\mathfrak{u}/B)$ is finitely satisfiable in U, note that $\varphi(\bar{x}) \in \mathrm{Av}(\mathfrak{u}/B)$ implies $U^{\bar{s}} \cap \varphi(\bar{x})^{\mathbb{M}} \in \mathfrak{u}$. In particular, this set is not empty. Hence, there is some $\bar{a} \in U^{\bar{s}}$ satisfying $\varphi(\bar{x})$. □

Lemma 3.11. *A type $\mathfrak{p} \in S^{\bar{s}}(B)$ is finitely satisfiable in U if, and only if, $\mathfrak{p} = \mathrm{Av}(\mathfrak{u}/B)$, for some ultrafilter $\mathfrak{u}$ over $U^{\bar{s}}$.*

Proof. ($\Leftarrow$) follows by Lemma 3.10. For ($\Rightarrow$), suppose that $\mathfrak{p}$ is finitely satisfiable in U. We start by showing that the set

$$\mathfrak{u}_0 := \left\{ U^{\bar{s}} \cap \varphi(\bar{x})^{\mathbb{M}} \;\middle|\; \varphi(\bar{x}) \in \mathfrak{p} \right\}.$$

has the finite intersection property. Let

$$U^{\bar{s}} \cap \varphi_0(\bar{x})^{\mathbb{M}}, \dots, U^{\bar{s}} \cap \varphi_n(\bar{x})^{\mathbb{M}} \in \mathfrak{u}_0\,, \quad \text{for } \varphi_0, \dots, \varphi_n \in \mathfrak{p}\,.$$

Since $\mathfrak{p}$ is closed under conjunction, it follows that $\varphi_0 \wedge \dots \wedge \varphi_n \in \mathfrak{p}$. As $\mathfrak{p}$ is finitely satisfiable in U,

$$\begin{aligned}&(U^{\bar{s}} \cap \varphi_0(\bar{x}))^{\mathbb{M}} \cap \dots \cap (U^{\bar{s}} \cap \varphi_n(\bar{x}))^{\mathbb{M}} \\ &\quad = U^{\bar{s}} \cap (\varphi_0(\bar{x}) \wedge \dots \wedge \varphi_n(\bar{x}))^{\mathbb{M}} \neq \varnothing\,,\end{aligned}$$

as desired.

By Corollary B2.4.10, there exists an ultrafilter $\mathfrak{u} \supseteq \mathfrak{u}_0$ over $U^{\bar{s}}$. Since, for every formula φ over B,

$$U^{\bar{s}} \cap \varphi(\bar{x})^{\mathbb{M}} \in \mathfrak{u} \quad \text{iff} \quad U^{\bar{s}} \cap \varphi(\bar{x})^{\mathbb{M}} \in \mathfrak{u}_0\,,$$

it follows that

$$\begin{aligned}\mathrm{Av}(\mathfrak{u}/B) &= \left\{ \varphi(\bar{x}) \;\middle|\; U^{\bar{s}} \cap \varphi(\bar{x})^{\mathbb{M}} \in \mathfrak{u} \right\} \\ &= \left\{ \varphi(\bar{x}) \;\middle|\; U^{\bar{s}} \cap \varphi(\bar{x})^{\mathbb{M}} \in \mathfrak{u}_0 \right\} = \{ \varphi(\bar{x}) \mid \varphi \in \mathfrak{p} \} = \mathfrak{p}\,,\end{aligned}$$

as desired. □

Using this characterisation of finite satisfiable types, we can prove that $\overset{\mathfrak{u}}{\surd}$ is a forking relation.

Proposition 3.12. $\overset{\mathrm{u}}{\surd}$ *is a forking relation.*

Proof. (INV) follows immediately from the definition.

(MON) If $\mathrm{tp}(\bar a_0 \bar a_1 / UB)$ is finitely satisfiable in U and $B_0 \subseteq B$, then $\mathrm{tp}(\bar a_0 / UB_0)$ is finitely satisfiable in U.

(NOR) If $\mathrm{tp}(\bar a / \bar c B)$ is finitely satisfiable in $\bar c$ then so is $\mathrm{tp}(\bar a \bar c / \bar c B)$.

(LRF) Clearly, $\mathrm{tp}(\bar a / B \bar a)$ is finitely satisfiable in $\bar a$.

(LTR) Suppose that $\mathrm{tp}(\bar a_0 \bar a_1 \bar a_2 / \bar a_0 \bar a_1 B)$ is finitely satisfiable in $\bar a_0 \bar a_1$ and $\mathrm{tp}(\bar a_0 \bar a_1 / \bar a_0 B)$ is finitely satisfiable in $\bar a_0$. If $\mathbb{M} \vDash \varphi(\bar a_0, \bar a_1, \bar a_2, \bar b)$, for $\bar b \subseteq \bar a_0 B$, there exists a tuple $\bar a_2' \subseteq \bar a_0 \bar a_1$ such that $\mathbb{M} \vDash \varphi(\bar a_0, \bar a_1, \bar a_2', \bar b)$. Suppose that $\bar a_2' = \bar a_0' \bar a_1'$ with $\bar a_0' \subseteq \bar a_0$ and $\bar a_1' \subseteq \bar a_1$. Then there are tuples $\bar c_1, \bar c_1' \subseteq \bar a_0$ with $\mathbb{M} \vDash \varphi(\bar a_0, \bar c_1, \bar a_0' \bar c_1', \bar b)$. Hence, $\mathrm{tp}(\bar a_0 \bar a_1 \bar a_2 / \bar a_0 B)$ is finitely satisfiable in $\bar a_0$.

(BMON) Obviously, if $\mathrm{tp}(\bar a / UBC)$ is finitely satisfiable in U, it is also finitely satisfiable in $U \cup C$.

(DEF) Suppose that $\mathrm{tp}(\bar a / UB)$ is not finitely satisfiable in U. Then there is some formula $\varphi(\bar x; \bar b) \in \mathrm{tp}(\bar a / UB)$ such that $\mathbb{M} \nvDash \varphi(\bar a'; \bar b)$, for all $\bar a' \subseteq U$. It follows that $\mathrm{tp}(\bar a' / U \bar b)$ is not finitely satisfiable in U, for every tuple $\bar a'$ that satisfies $\varphi(\bar x; \bar b)$.

(EXT) Suppose that the type $\mathfrak{p} := \mathrm{tp}(\bar a / UB_0)$ is finitely satisfiable in U and let $B_1 \supseteq B_0$. According to Lemma 3.11 there exists an ultrafilter $\mathfrak{u}$ such that $\mathfrak{p} = \mathrm{Av}(\mathfrak{u} / UB_0)$. Let $\bar a'$ be a realisation of $\mathrm{Av}(\mathfrak{u} / UB_1)$. Then $\mathrm{tp}(\bar a' / UB_0) = \mathrm{Av}(\mathfrak{u} / UB_0) = \mathfrak{p}$ and $\mathrm{tp}(\bar a' / UB_1) = \mathrm{Av}(\mathfrak{u} / UB_1)$ is finitely satisfiable in U. ☐

Our next aim is to show that $\overset{\mathrm{u}}{\surd}$ is the least preforking relation.

Theorem 3.13 (Adler). $\overset{\mathrm{u}}{\surd} \subseteq \surd$, *for every preforking relation* $\surd$.

Proof. For a contradiction, suppose that $A \overset{\mathrm{u}}{\surd}_U B$ but $A \not\surd_U B$. By (DEF), there are a formula $\varphi(\bar x, \bar y)$ over U and tuples $\bar a \subseteq A$ and $\bar b \subseteq B$ such that $\mathbb{M} \vDash \varphi(\bar a, \bar b)$ and

$$\bar a' \not\surd_U \bar b\,, \quad \text{for all } \bar a' \in \varphi(\bar x, \bar b)^{\mathbb{M}}\,.$$

Since $\mathrm{tp}(\bar a/BU)$ is finitely satisfiable in U, there is some tuple $\bar c \subseteq U$ with $\mathbb{M} \vDash \varphi(\bar c, \bar b)$. Consequently, $\bar c \not\mathop{\downarrow}_U \bar b$ which, by (MON), implies that $U \not\mathop{\downarrow}_U B$. A contradiction to (LRF). ☐

As a corollary we obtain the following result which, in the terminology introduced below, states that the relation $\mathop{\downarrow}^{\mathrm{u}}$ is *left local.* Below we will extend this result to all preforking relations.

Lemma 3.14. *Let T be a complete first-order theory. For all $\bar a, B \subseteq \mathbb{M}$, there is a set $U \subseteq \bar a$ of size $|U| \leq |T| \oplus |B|$ such that* $\mathrm{tp}(\bar a/UB)$ *is finitely satisfiable in U.*

Proof. We construct an increasing sequence $U_0 \subseteq U_1 \subseteq \dots$ of sets $U_n \subseteq \bar a$ with $|U_n| \leq |T| \oplus |B|$ as follows. We start with $U_0 := \varnothing$. For the inductive step suppose that we have already constructed $U_n \subseteq \bar a$. For every formula $\varphi(\bar x; \bar b) \in \mathrm{tp}(\bar a/BU_n)$, let $\bar c_\varphi \subseteq \bar a$ be the elements of $\bar a$ that are mentioned in $\varphi(\bar x)$. Note that $\bar c_\varphi$ is finite. Let U_{n+1} be the set obtained from U_n by adding all these tuples $\bar c_\varphi$. Then $|U_{n+1}| \leq |T| \oplus |B| \oplus |U_n| \leq |T| \oplus |B|$.

Setting $U := \bigcup_{n<\omega} U_n$ it follows that $\mathrm{tp}(\bar a/UB)$ is finitely satisfiable in U. Furthermore, $|U| \leq |T| \oplus |B|$. ☐

Let us conclude this section with a remark about sets where $\mathop{\downarrow}^{\mathrm{u}}$ is right reflexive.

Lemma 3.15. *Let T be a complete first-order theory. A subset $M \subseteq \mathbb{M}$ is the universe of a model of T if, and only if, $A \mathop{\downarrow}^{\mathrm{u}}_M M$, for all sets A.*

Proof. ($\Rightarrow$) Let $\mathfrak{M} \preceq \mathbb{M}$ be a model of T and $\bar a \subseteq \mathbb{M}$ a tuple. To show that $\bar a \mathop{\downarrow}^{\mathrm{u}}_M M$, consider a formula $\varphi(\bar x) \in \mathrm{tp}(\bar a/M)$. Then $\mathbb{M} \vDash \exists \bar x \varphi$ implies $\mathfrak{M} \vDash \exists \bar x \varphi$. Hence, there is some $\bar c \subseteq M$ with $\mathfrak{M} \vDash \varphi(\bar c)$.

($\Leftarrow$) Suppose that $A \mathop{\downarrow}^{\mathrm{u}}_M M$ for all sets A. We prove that M satisfies the Tarski-Vaught Test. Let $\varphi(x)$ be a formula over M such that $\mathbb{M} \vDash \exists x \varphi(x)$. We fix an element $a \in \mathbb{M}$ with $\mathbb{M} \vDash \varphi(a)$. Since $a \mathop{\downarrow}^{\mathrm{u}}_M M$, there is some element $c \in M$ with $\mathbb{M} \vDash \varphi(c)$. By Theorem C2.2.5, it follows that $M \preceq \mathbb{M}$. Consequently, M is a model of T. ☐

Local character and forking sequences

In the remainder of this section we study preforking relations with a property called *local character*. In the next section, we will prove that having local character is equivalent to being symmetric.

Definition 3.16. A ternary relation $\surd$ has *local character* if it satisfies the following two axioms:

(LLOC) *Left Locality.* There exists some cardinal κ such that, for all sets A and B, there is a subset $A_0 \subseteq A$ of size $|A_0| < \kappa \oplus |B|^+$ with $A \surd_{A_0} B$.

(RLOC) *Right Locality.* There exists a cardinal κ such that, for all sets A and B, there is a subset $B_0 \subseteq B$ of size $|B_0| < \kappa \oplus |A|^+$ with $A \surd_{B_0} B$.

If $\surd$ is right local, we denote by $\operatorname{loc}(\surd)$ the least cardinal κ such that $\surd$ satisfies the condition in (RLOC). Similarly, $\operatorname{loc}_0(\surd)$ the least cardinal κ such that $\surd$ satisfies the above condition for *finite* sets A. If $\surd$ is not right local, we set $\operatorname{loc}(\surd) := \infty$ and $\operatorname{loc}_0(\surd) := \infty$.

We start by proving that every preforking relation is left local.

Proposition 3.17. *Let T be a complete first-order theory and let $\surd$ be a preforking relation. For all sets $A, B \subseteq \mathbb{M}$, there exists a subset $A_0 \subseteq A$ of size $|A_0| \leq |T| \oplus |B|$ such that*

$$A \surd_{A_0} B\,.$$

Proof. Let A and B be sets. By Lemma 3.14, there is a set $A_0 \subseteq A$ of size $|A_0| \leq |T| \oplus |B|$ such that $A \overset{\mathrm{u}}{\surd}_{A_0} B$. By Theorem 3.13, this implies that $A \surd_{A_0} B$. □

Corollary 3.18. *Let T be a complete first-order theory and let $\downarrow$ a symmetric preforking relation. Then* $\operatorname{loc}(\downarrow) \leq |T|^+$.

The two parameters $\operatorname{loc}_0(\surd)$ and $\operatorname{loc}(\surd)$ are nearly the same. They can only differ if the first one is a singular cardinal.

Definition 3.19. For a cardinal κ, we denote by κ^{reg} the minimal regular cardinal with $\kappa^{\text{reg}} \geq \kappa$, that is,

$$\kappa^{\text{reg}} := \begin{cases} \kappa & \text{if } \kappa \text{ is regular,} \\ \kappa^+ & \text{if } \kappa \text{ is singular.} \end{cases}$$

Lemma 3.20. *Let $\surd$ be an abstract independence relation that satisfies* (BMON) *and* (RLOC). *Then*

$$\text{loc}_0(\surd) \leq \text{loc}(\surd) \leq \text{loc}_0(\surd)^{\text{reg}}.$$

Proof. The lower bound follows immediately from the definitions. For the upper bound, let $\kappa := \text{loc}_0(\surd)^{\text{reg}}$ and consider sets $A, B \subseteq \mathbb{M}$. We have to find a set $U \subseteq A$ of size $|U| < \kappa \oplus |A|^+$ with $A \surd_U B$.

For every finite set $A_0 \subseteq A$, we choose a set $U(A_0) \subseteq B$ of size $|U(A_0)| < \text{loc}_0(\surd) \leq \kappa$ such that

$$A_0 \surd_{U(A_0)} B\,.$$

Setting $U := \bigcup \{\, U(A_0) \mid A_0 \subseteq A \text{ finite} \,\}$ it follows by (BMON) that

$$A_0 \surd_U B\,, \quad \text{for all finite } A_0 \subseteq A\,.$$

By (FIN), this implies $A \surd_U B$. Since the cardinal $\kappa \oplus |A|^+$ is regular, we furthermore have

$$|U| \leq \sum_{A_0 \subseteq A \text{ finite}} |U(A_0)| < \kappa \oplus |A|^+\,.$$

□

We can characterise preforking relations with local character in terms of so-called *forking chains.*

Definition 3.21. Let $\surd$ be a preforking relation.

(a) Let $A, U \subseteq \mathbb{M}$ be sets. A sequence of finite sets $(B_\alpha)_{\alpha<\gamma}$ is a $\surd$-*forking chain* for A over U if

$$A \not\surd_{UB[<\alpha]} B_\alpha\,, \quad \text{for every } \alpha < \gamma\,,$$

where we have set $B[<\alpha] := \bigcup_{\beta<\alpha} B_\beta$. The ordinal γ is the *length* of the chain.

(b) We denote by $\mathrm{fc}(\surd)$ the least cardinal κ such that no finite set A has a $\surd$-forking chain over $\varnothing$ of length κ. If such a cardinal does not exist, we set $\mathrm{fc}(\surd) := \infty$.

In the theorem below we show that the cardinal $\mathrm{fc}(\surd)$ is closely related to the parameter $\mathrm{loc}(\surd)$. As we will apply these results in a later chapter to relations that are not preforking relations, we state them in a slightly more general setting.

Definition 3.22. A ternary relation $\surd$ has *strong finite character* if it satisfies the following axiom:

(SFIN) *Strong Finite Character.*

$$A \surd_U B \quad \text{iff} \quad A_0 \surd_U B_0 \quad \text{for all finite } A_0 \subseteq A \text{ and } B_0 \subseteq B\,.$$

Remark. Note that every preforking relation has strong finite character since (SFIN) follows from (FIN) and (DEF).

The following lemma contains the key argument of the translation between $\mathrm{fc}(\surd)$ and $\mathrm{loc}(\surd)$.

Lemma 3.23. *Let $\surd$ be an abstract independence relation that satisfies* (BMON) *and* (SFIN), *let κ be an infinite cardinal and $A \subseteq \mathbb{M}$.*

(a) *If there exists some set B such that $A \not\surd_U B$, for all $U \subseteq B$ of size $|U| < \kappa$, then there is a $\surd$-forking chain for A over $\varnothing$ of length κ.*

(b) *If κ is regular and every set B has a subset $U \subseteq B$ of size $|U| < \kappa$ with $A \surd_U B$, then there is no $\surd$-forking chain for A over $\varnothing$ of length κ.*

Proof. (a) We construct the desired $\surd$-forking chain $(B_\alpha)_{\alpha<\kappa}$ by induction on α. Suppose that we have already defined B_α, for all $\alpha < \beta$. Then

$$\begin{aligned} &\big|B[<\beta]\big| < \aleph_0 \leq \kappa\,, && \text{for } \beta < \omega\,, \\ \text{and}\quad &\big|B[<\beta]\big| \leq \aleph_0 \otimes |\beta| < \kappa\,, && \text{for } \omega \leq \beta < \kappa\,. \end{aligned}$$

In both cases it follows that $A \not\mathrel{\sqrt{}}_{B[<\beta]} B$. Hence, we can use (SFIN) to find a finite set $B_\beta \subseteq B$ with $A \not\mathrel{\sqrt{}}_{B[<\beta]} B_\beta$.

(b) Let $(B_\alpha)_{\alpha<\kappa}$ a sequence of finite sets of length κ. By assumption, there exists a set $U \subseteq B[<\kappa]$ of size $|U| < \kappa$ such that

$$A \sqrt{}_U B[<\kappa] .$$

As κ is regular, there is some index $\alpha < \kappa$ with $U \subseteq B[<\alpha]$. By (BMON) and (MON) it follows that

$$A \sqrt{}_{B[<\alpha]} B_\alpha .$$

Consequently, $(B_\alpha)_{\alpha<\kappa}$ is no $\sqrt{}$-forking chain for A over $\varnothing$. □

Proposition 3.24. *Let* $\sqrt{}$ *be an abstract independence relation satisfying* (BMON) *and* (SFIN). *Then*

$$\mathrm{loc}_0(\sqrt{}) \leq \mathrm{fc}(\sqrt{}) \leq \mathrm{loc}_0(\sqrt{})^{\mathrm{reg}}.$$

Proof. For the lower bound, consider a finite set A and an arbitrary set B. If there were no set $U \subseteq B$ of size $|U| < \mathrm{fc}(\sqrt{})$ with $A \sqrt{}_U B$, we could use Lemma 3.23 (a) to construct a $\sqrt{}$-forking chain for A over $\varnothing$ of length $\mathrm{fc}(\sqrt{})$. A contradiction.

For the upper bound, consider a finite set A. Then Lemma 3.23 (b) implies that there is no $\sqrt{}$-forking chain for A over $\varnothing$ of length $\mathrm{loc}_0(\sqrt{})^{\mathrm{reg}}$. □

Theorem 3.25. *For a preforking relation* $\sqrt{}$, *the following statements are equivalent:*

(1) $\sqrt{}$ *has local character.*

(2) $\sqrt{}$ *is right local.*

(3) *For every set A, there exists a cardinal* κ *such that there is no* $\sqrt{}$-*forking chain for A over* $\varnothing$ *of length* κ.

(4) *There exists a cardinal κ such that, for every finite set A, there is no $\surd$-forking chain for A over $\varnothing$ of length κ.*

Proof. (1) ⇔ (2) ⇔ (4) follow by Propositions 3.17 and 3.24, respectively.

(2) ⇒ (3) Given a set A, it follows by Lemma 3.23 (b) that there is no $\surd$-forking chain for A over $\varnothing$ of length $\kappa := \mathrm{loc}(\surd)^+ \oplus |A|^+$.

(3) ⇒ (4) For every type $\mathfrak{p} \in S^{<\omega}(\varnothing)$, fix a tuple $\bar{a}_\mathfrak{p}$ realising $\mathfrak{p}$. By (3), there are cardinals $\kappa_\mathfrak{p}$ such that there are no $\surd$-forking chains for $\bar{a}_\mathfrak{p}$ over $\varnothing$ of length $\kappa_\mathfrak{p}$. We claim that the cardinal

$$\kappa := \sup \{ \kappa_\mathfrak{p} \mid \mathfrak{p} \in S^{<\omega}(\varnothing) \}$$

has the desired properties. Let $\bar{a}$ be a finite tuple and $(B_\alpha)_{\alpha<\kappa}$ a sequence of finite sets of length κ. Then $\bar{a} \equiv_\varnothing \bar{a}_\mathfrak{p}$, for $\mathfrak{p} := \mathrm{tp}(\bar{a})$, and there exists an automorphism π with $\pi(\bar{a}) = \bar{a}_\mathfrak{p}$. Since $\kappa \geq \kappa_\mathfrak{p}$, there is an index $\alpha < \kappa$ such that

$$\bar{a}_\mathfrak{p} \ \surd_{\pi[B[<\alpha]]} \ \pi[B_\alpha] \,.$$

By invariance, it follows that $\bar{a} \ \surd_{B[<\alpha]} \ B_\alpha$. Hence, $(B_\alpha)_{\alpha<\kappa}$ is not a $\surd$-forking chain for $\bar{a}$ over $\varnothing$. □

4. Forking relations

In this section we consider the special properties of forking relations that follow form the extension axiom. We start by presenting a canonical way to turn every preforking relation into a forking relation.

Definition 4.1. Let $\surd$ be a preforking relation. We define a relation $\overset{*}{\surd}$ by

$$A \ \overset{*}{\surd}_U \ B \quad :\text{iff} \quad \text{for every set } C \subseteq \mathbb{M} \text{ there is some set } A' \subseteq \mathbb{M} \text{ with } A' \equiv_{UB} A \text{ and } A' \ \surd_U \ BC \,.$$

Remark. Note that $\overset{*}{\surd} \subseteq \surd$. Furthermore, by Proposition 4.5 below it will follow that $\overset{*}{\surd} = \surd$ if, and only if, $\surd$ is a forking relation. Consequently, the operation $\surd \mapsto \overset{*}{\surd}$ is a so-called *kernel operator,* the dual of a closure operator:

$$\overset{*}{\surd} \subseteq \surd\,, \qquad \overset{**}{\surd} = \overset{*}{\surd}\,, \quad \text{and} \quad \surd_0 \subseteq \surd_1 \;\Rightarrow\; \overset{*}{\surd}_0 \subseteq \overset{*}{\surd}_1\,.$$

Before proving that $\overset{*}{\surd}$ is a forking relation, we present two alternative definitions. The first one characterises such relations in terms of global types.

Definition 4.2. A *global type* is a complete type over $\mathbb{M}$.

Proposition 4.3. *Let $\surd$ be a preforking relation and $\bar a, U, B \subseteq \mathbb{M}$. Then*

$$\bar a \overset{*}{\surd}_U B \quad \text{iff} \quad \mathrm{tp}(\bar a/UB) \textit{ can be extended to a global type that is } \surd\textit{-free over } U\,.$$

Proof. ($\Leftarrow$) Let $\mathfrak{p} \supseteq \mathrm{tp}(\bar a/UB)$ be a global type that is $\surd$-free over U. To show that $\bar a \overset{*}{\surd}_U B$, consider a set $C \subseteq \mathbb{M}$. Choosing some tuple $\bar a'$ realising $\mathfrak{p} \upharpoonright UBC$, we have $\bar a' \equiv_{UB} \bar a$ and $\bar a' \surd_U BC$.

($\Rightarrow$) Suppose that $\bar a \overset{*}{\surd}_U B$ and set

$$\Phi(\bar x) := \mathrm{tp}(\bar a/UB) \cup \{\, \neg\varphi(\bar x) \mid \varphi \text{ a formula over } \mathbb{M} \text{ that } \surd\text{-forks over } U \,\}\,.$$

We start by proving that Φ is satisfiable. Let $\Phi_0 \subseteq \Phi$ be finite. Then

$$\Phi_0 \equiv \{\psi(\bar x), \neg\varphi_0(\bar x; \bar c_0), \dots, \neg\varphi_n(\bar x; \bar c_n)\}\,,$$

for some $\psi \in \mathrm{tp}(\bar a/UB)$ and formulae $\varphi_i(\bar x; \bar c_i)$ that $\surd$-fork over U. Since $\bar a \overset{*}{\surd}_U B$, there exists a tuple $\bar a' \equiv_{UB} \bar a$ such that $\bar a' \surd_U B\bar c_0 \dots \bar c_n$. Then $\bar a'$ satisfies Φ_0.

Hence, Φ is satisfiable and there exists a global type $\mathfrak{p} \supseteq \Phi$. We claim that $\mathfrak{p}$ is $\surd$-free over U. For a contradiction, suppose that $\mathfrak{p} \vDash \varphi(\bar x)$, for some formula φ that $\surd$-forks over U. Then $\neg\varphi \in \Phi \subseteq \mathfrak{p}$. A contradiction. □

The second characterisation considers forking relations in terms of types and formulae. The key here is that the formulae ψ_i below might have parameters that do not appear in Φ.

Lemma 4.4. *Let $\surd$ be a preforking relation. A partial type Φ $\overset{*}{\surd}$-forks over U if, and only if, for some $n < \omega$, there are formulae $\psi_0, \dots, \psi_{n-1}$ with parameters such that*

$$\Phi(\bar{x}) \vDash \bigvee_{i<n} \psi_i(\bar{x}) \quad \text{and each } \psi_i \ \surd\text{-forks over } U\,.$$

Proof. ($\Leftarrow$) Fix a tuple $\bar{a} \in \Phi^{\mathbb{M}}$ and let B be a set such that Φ is a partial type over B. For a contradiction, suppose that $\bar{a} \ \overset{*}{\surd}_U\ B$. We choose a set C containing the parameters of every formula ψ_i. By definition of $\overset{*}{\surd}$, there is some tuple $\bar{a}' \equiv_{UB} \bar{a}$ such that $\bar{a}' \ \surd_U\ BC$. Since $\Phi \vDash \bigvee_i \psi_i$, we have $\mathbb{M} \vDash \psi_i(\bar{a}')$, for some $i < n$. As ψ_i $\surd$-forks over U, it follows that $\bar{a}' \not\surd_U BC$. A contradiction.

($\Rightarrow$) Suppose that Φ $\overset{*}{\surd}$-forks over U and let B be some set such that Φ is a partial type over B. By definition of $\overset{*}{\surd}$, there exists, for every tuple $\bar{a} \in \Phi^{\mathbb{M}}$, some set $C_{\bar{a}}$ such that

$$\bar{a}' \not\surd_U BC_{\bar{a}}\,, \quad \text{for all } \bar{a}' \equiv_{UB} \bar{a}\,.$$

By (DEF), we can find a formula $\psi_{\bar{a}}(\bar{x}, \bar{b}_{\bar{a}}, \bar{c}_{\bar{a}})$ with parameters $\bar{b}_{\bar{a}} \subseteq B$ and $\bar{c}_{\bar{a}} \subseteq C_{\bar{a}}$ such that

$$\mathbb{M} \vDash \psi_{\bar{a}}(\bar{a}, \bar{b}_{\bar{a}}, \bar{c}_{\bar{a}}) \quad \text{and} \quad \psi_{\bar{a}}(\bar{x}, \bar{b}_{\bar{a}}, \bar{c}_{\bar{a}})\ \surd\text{-forks over } U\,.$$

Consequently, the set

$$\Phi(\bar{x}) \cup \big\{\, \neg\psi_{\bar{a}}(\bar{x}, \bar{b}_{\bar{a}}, \bar{c}_{\bar{a}}) \,\big|\, \bar{a} \in \Phi^{\mathbb{M}} \,\big\}$$

is inconsistent. By compactness, we can therefore find finitely many tuples $\bar{a}_0, \dots, \bar{a}_{n-1}$ such that

$$\Phi(\bar{x}) \vDash \bigvee_{i<n} \psi_{\bar{a}_i}(\bar{x}, \bar{b}_{\bar{a}_i}, \bar{c}_{\bar{a}_i})$$

and each formula $\psi_{\bar{a}_i}(\bar{x}, \bar{b}_{\bar{a}_i}, \bar{c}_{\bar{a}_i})$ $\surd$-forks over U. □

Next we prove that the operations $\sqrt{} \mapsto \sqrt[*]{}$ turns every preforking relation into a forking relation.

Proposition 4.5. *If* $\sqrt{}$ *is a preforking relation then* $\sqrt[*]{}$ *is a forking relation.*

Proof. (INV) follows easily from the definition.

(MON) Suppose that $A_0 A_1 \sqrt[*]{}_U\, B$ and let $B_0 \subseteq B$. To show that $A_0 \sqrt[*]{}_U\, B_0$ let $C \subseteq \mathbb{M}$. By definition of $\sqrt[*]{}$, there are sets A'_0 and A'_1 with $A'_0 A'_1 \equiv_{UB} A_0 A_1$ and $A'_0 A'_1 \sqrt{}_U\, BC$. This implies that $A'_0 \equiv_{UB_0} A_0$ and $A'_0 \sqrt{}_U\, BC$.

(NOR) Suppose that $A \sqrt[*]{}_U\, B$. To show that $AU \sqrt[*]{}_U\, BU$, let $C \subseteq \mathbb{M}$. There is some set A' such that $A' \equiv_{UB} A$ and $A' \sqrt{}_U\, BCU$. It follows by (NOR) that $A'U \sqrt{}_U\, BCU$. Since $A'U \equiv_{UB} AU$ the claim follows.

(LRF) For all sets $A, B, C \subseteq \mathbb{M}$, we have $A \sqrt{}_A\, BC$. Hence, $A \sqrt[*]{}_A\, B$.

(LTR) Suppose that $A_2 \sqrt[*]{}_{A_1}\, B$ and $A_1 \sqrt[*]{}_{A_0}\, B$ for $A_0 \subseteq A_1 \subseteq A_2$. To show that $A_2 \sqrt[*]{}_{A_0}\, B$ let $C \subseteq \mathbb{M}$. There exists a set A'_1 with $A'_1 \equiv_{A_0 B} A_1$ and $A'_1 \sqrt{}_{A_0}\, BC$. Let A'_2 be some set such that $A'_1 A'_2 \equiv_{A_0 B} A_1 A_2$. By (INV) it follows that $A'_2 \sqrt[*]{}_{A'_1}\, B$. Therefore, there exists a set A''_2 with $A''_2 \equiv_{A'_1 B} A'_2$ and $A''_2 \sqrt{}_{A'_1}\, BC$. By (LTR) it follows that $A''_2 \sqrt{}_{A_0}\, BC$, as desired.

(BMON) Suppose that $A \sqrt[*]{}_U\, BC$. To show that $A \sqrt[*]{}_{UC}\, BC$, let $D \subseteq \mathbb{M}$. There is a set A' with $A' \equiv_{UBC} A$ such that $A' \sqrt{}_U\, BCD$. By (BMON) it follows that $A' \sqrt{}_{UC}\, BCD$.

(EXT) Suppose that $A \sqrt[*]{}_U\, B$ and let $\bar{a}$ be an enumeration of A. By Proposition 4.3, there exists some global type $\mathfrak{p} \supseteq \operatorname{tp}(\bar{a}/UB)$ that is $\sqrt{}$-free over U. Given a set $C \subseteq \mathbb{M}$, we choose some tuple $\bar{a}'$ realising $\mathfrak{p} \upharpoonright UBC$. Then $\bar{a}' \equiv_{UB} \bar{a}$ and $\operatorname{tp}(\bar{a}'/UBC) = \mathfrak{p} \upharpoonright UBC$ has the global extension $\mathfrak{p}$, which is $\sqrt{}$-free over U. Hence, Proposition 4.3 implies that $\bar{a}' \sqrt[*]{}_U\, BC$.

(DEF) Suppose that $\bar{a} \not\sqrt[*]{}_U\, B$. Then there is a set $C \subseteq \mathbb{M}$ such that $\bar{a}' \not\sqrt{}_U\, BC$ for all tuples $\bar{a}' \equiv_{UB} \bar{a}$. Let Φ be the set of all formulae

$\varphi(\bar{x}) \in \operatorname{tp}(\bar{a}/UBC)$ that $\surd$-fork over U. Since $\surd$ is definable, it follows by choice of C that the set

$$\operatorname{tp}(\bar{a}/UB) \cup \{\, \neg\varphi \mid \varphi \in \Phi \,\}$$

is inconsistent. Hence, there is some formula $\psi(\bar{x};\bar{b}) \in \operatorname{tp}(\bar{a}/UB)$ such that

$$\psi(\bar{x};\bar{b}) \vDash \bigvee \Phi\,.$$

We claim that $\mathbb{M} \vDash \psi(\bar{a}';\bar{b})$ implies $\bar{a}' \not\overset{*}{\surd}_U \bar{b}$. Suppose otherwise. Then there exists a tuple $\bar{a}''$ such that $\bar{a}'' \equiv_{U\bar{b}} \bar{a}'$ and $\bar{a}'' \surd_U BC$. But there is some formula $\varphi \in \Phi$ with $\mathbb{M} \vDash \varphi(\bar{a}'')$. By definition of Φ this implies that $\bar{a}'' \not\surd_U BC$. A contradiction. □

Lemma 4.6. $\mathrm{cl}_{\surd} = \mathrm{cl}_{\overset{*}{\surd}}$, *for every preforking relation* $\surd$.

Proof. Note that $\overset{*}{\surd} \subseteq \surd$ implies $\mathrm{cl}_{\overset{*}{\surd}} \subseteq \mathrm{cl}_{\surd}$. Conversely, suppose that $a \notin \mathrm{cl}_{\overset{*}{\surd}}(U)$. Then there are sets B and C such that $a \not\overset{*}{\surd}_{UC} B$. Hence, we can find a set D such that $a' \not\surd_{UC} BD$, for all $a' \equiv_{UCB} a$. In particular, we have $a \not\surd_{UC} BD$, which implies that $a \notin \mathrm{cl}_{\surd}(U)$. □

Exercise 4.1. Let $\surd$ be a preforking relation. Prove that, if $\overset{*}{\surd}$ is right local, then so is $\surd$.

To check whether a forking relation is contained in another one, we can frequently use the following lemma.

Lemma 4.7. *Let* $\overset{0}{\surd}$ *be a relation satisfying* (ext) *and let* $\overset{1}{\surd}$ *be a relation satisfying* (inv) *and* (mon). *If, for all sets B and U, there exists some set C such that*

$$A \overset{0}{\surd}_U BC \quad \text{implies} \quad A \overset{1}{\surd}_U BC\,, \quad \text{for all sets } A\,,$$

then $\overset{0}{\surd} \subseteq \overset{1}{\surd}$.

Proof. Suppose that $A \sqrt[0]{}_U B$. By assumption, we can find a set C such that

$$A \sqrt[0]{}_U BC \quad \text{implies} \quad A \sqrt[1]{}_U BC\,, \quad \text{for all sets } A\,.$$

By (EXT), there is some set $A' \equiv_{UB} A$ such that $A' \sqrt[0]{}_U BC$. By choice of C, it follows that $A' \sqrt[1]{}_U BC$. Consequently, (MON) and (INV) imply that $A \sqrt[1]{}_U B$. □

Morley sequences

The aim of this section is to introduce the notion of a basis for an arbitrary forking relation. Since, in general, forking relations are not symmetric, these bases are ordered. To simplify notation we write $\bar{a}[<k]$, for a sequence $(\bar{a}_i)_{i\in I}$, to denote the set $\bigcup_{i<k} \bar{a}_i$.

Definition 4.8. Let $\sqrt{}$ be a preforking relation and $\mathfrak{p} \in S^{\bar{s}}(U \cup B)$ a type.

(a) A $\sqrt{}$-*Morley sequence* for $\mathfrak{p}$ over U is an indiscernible sequence $(\bar{a}_i)_{i\in I}$ over $U \cup B$ such that every $\bar{a}_i$ realises $\mathfrak{p}$ and

$$\bar{a}_i \sqrt{}_U \bar{a}[<i]\,, \quad \text{for all } i \in I\,.$$

We call $(\bar{a}_i)_{i\in I}$ a $\sqrt{}$-*Morley sequence over U* if it is a $\sqrt{}$-Morley sequence for $\operatorname{tp}(\bar{a}_i/U)$ over U.

(b) A *reverse* $\sqrt{}$-*Morley sequence* for $\mathfrak{p}$ over U is an indiscernible sequence $(\bar{a}_i)_{i\in I}$ over $U \cup B$ such that every $\bar{a}_i$ realises $\mathfrak{p}$ and

$$\bar{a}[<i] \sqrt{}_U \bar{a}_i\,, \quad \text{for all } i \in I\,.$$

Remark. If $(\bar{a}_i)_{i\in I}$ is a $\sqrt{}$-Morley sequence for $\mathfrak{p}$ over U, then it follows by (FIN), Lemma 2.4, and induction, that

$$\bar{a}[I_1] \sqrt{}_U \bar{a}[I_0]\,, \quad \text{for all } I_0, I_1 \subseteq I \text{ with } I_0 < I_1\,.$$

For symmetric preforking relations, we obtain the following stronger result.

Lemma 4.9. *Let $\downfree$ be a symmetric preforking relation and $(\bar{a}_i)_{i \in I}$ a sequence such that*

$$\bar{a}_i \downfree_U \bar{a}[<i] , \quad \textit{for all } i \in I .$$

Then

$$\bar{a}[K] \downfree_U \bar{a}[L] , \quad \textit{for all disjoint } K, L \subseteq I .$$

Proof. By (FIN), it is sufficient to prove the claim for finite sets K and L. We do so by induction on $|K \cup L|$. If both sets are empty, the claim follows by (NOR). Otherwise, let $k := \max(K \cup L)$. By (SYM), we may assume without loss of generality that $k \in K$. Set $K_0 := K \setminus \{k\}$. By inductive hypothesis, we have

$$\bar{a}[K_0] \downfree_U \bar{a}[L] .$$

Furthermore,

$$\bar{a}_k \downfree_U \bar{a}[<k] \quad \text{implies} \quad \bar{a}_k \downfree_U \bar{a}[K_0]\bar{a}[L] .$$

Consequently, it follows by Lemma 2.4 that

$$\bar{a}_k\bar{a}[K_0] \downfree_U \bar{a}[L] . \qquad \square$$

We can use the extension axiom to construct Morley sequences.

Proposition 4.10. *Let $\surd$ be a forking relation. If $\bar{a} \surd_U B$ then there is a $\surd$-Morley sequence $(\bar{a}_n)_{n<\omega}$ for* $\mathrm{tp}(\bar{a}/UB)$ *over U.*

Proof. Set $\lambda := |T| \oplus |U| \oplus |B| \oplus |\bar{a}| \oplus \aleph_0$ and let $\kappa > \beth_{2^\lambda}$. First, we construct a sequence $(\bar{c}_\alpha)_{\alpha<\kappa}$ of tuples realising $\mathrm{tp}(\bar{a}/UB)$ such that

$$\bar{c}_\alpha \surd_U B\bar{c}[<\alpha] , \quad \text{for all } \alpha < \kappa .$$

By induction, suppose that we have already defined $\bar{c}_\beta$, for all $\beta < \alpha$. Since $\bar{a} \surd_U B$, we can use (EXT) to find a tuple $\bar{c}_\alpha \equiv_{UB} \bar{a}$ such that $\bar{c}_\alpha \surd_U B\bar{c}[<\alpha]$.

Having constructed $(\bar{c}_\alpha)_{\alpha<\kappa}$, we use Theorem E5.3.7 to find an indiscernible sequence $(\bar{a}_n)_{n<\omega}$ over $U \cup B$ such that, for every $n < \omega$, there are indices $\alpha_0 < \cdots < \alpha_{n-1} < \kappa$ with

$$\bar{a}_0 \dots \bar{a}_{n-1} \equiv_{UB} \bar{c}_{\alpha_0} \dots \bar{c}_{\alpha_{n-1}} .$$

By (INV) and (MON) it follows that $\bar{a}_n \sqrt{}_U B\bar{a}[<n]$. Hence, $(\bar{a}_n)_{n<\omega}$ is the desired $\sqrt{}$-Morley sequence. □

Corollary 4.11. *Let $\downarrow$ be a symmetric forking relation. For every tuple $\bar{a}$, every set U, and every linear order I, there exists a $\downarrow$-Morley sequence $(\bar{a}_i)_{i\in I}$ for $\mathrm{tp}(\bar{a}/U)$ over U.*

Proof. As $\downarrow$ is symmetric, we have $\bar{a} \downarrow_U U$. Therefore, we can use Proposition 4.10 to find a $\downarrow$-Morley sequence $(\bar{c}_n)_{n<\omega}$ for $\mathrm{tp}(\bar{a}/U)$ over U. By compactness and (FIN), it follows that there also exists a $\downarrow$-Morley sequence $(\bar{a}_i)_{i\in I}$ for $\mathrm{tp}(\bar{a}/U)$ over U that is indexed by I. □

Lemma 4.12. *Let $\sqrt{}$ be a forking relation and let $\mathfrak{p}$ be a type over $U \cup B$. If there exists a $\sqrt{}$-Morley sequence $(\bar{c}_n)_{n<\omega}$ for $\mathfrak{p}$ over U, then there exists a reverse $\sqrt{}$-Morley sequence $(\bar{a}_n)_{n<\omega}$ for $\mathfrak{p}$ over U.*

Proof. Let $(\bar{c}_n)_{n<\omega}$ be a $\sqrt{}$-Morley sequence for $\mathfrak{p}$ over U. By compactness, there exists a sequence $(\bar{a}_n)_{n<\omega}$ such that

$$\bar{a}_0 \dots \bar{a}_n \equiv_{UB} \bar{c}_n \dots \bar{c}_0 , \quad \text{for all } n < \omega .$$

By definition of a Morley sequence we have

$$\bar{c}_n \sqrt{}_U \bar{c}_0 \dots \bar{c}_{n-1} .$$

Hence (INV) implies that

$$\bar{a}_i \sqrt{}_U \bar{a}_{i+1} \dots \bar{a}_n , \quad \text{for all } i < n < \omega .$$

Repeatedly applying Lemma 2.4 it follows that

$$\bar{a}_0 \dots \bar{a}_{n-1} \sqrt{}_U \bar{a}_n , \quad \text{for every } n < \omega .$$ □

The following lemma can be used in some cases to construct a reverse $\surd$-Morley sequence out of an indiscernible sequence.

Lemma 4.13. *Let $\surd$ be a preforking relation and let I, J be linear orders such that I has no maximal element. If $(\bar{a}_i)_{i \in I+J}$ is indiscernible over U then $(\bar{a}_j)_{j \in J}$ is a reverse $\surd$-Morley sequence over $U \cup \bar{a}[I]$.*

Proof. Clearly, $(\bar{a}_j)_{j \in J}$ is indiscernible over $U \cup \bar{a}[I]$. To show that it is a reverse $\surd$-Morley sequence over $U \cup \bar{a}[I]$, it is sufficient, by (FIN), to prove that

$$\bar{a}_{j_0} \dots \bar{a}_{j_{k-1}} \surd_{U\bar{a}[I]} \bar{a}_{j_k}, \quad \text{for all } j_0 < \dots < j_k \text{ in } J,\ k < \omega.$$

Hence, consider indices $j_0 < \dots < j_k$ in J. By indiscernibility and the fact that I has no maximal element, we can find, for every finite set $I_0 \subseteq I$, indices $i_0 < \dots < i_{k-1}$ in I such that

$$\bar{a}_{j_0} \dots \bar{a}_{j_{k-1}} \bar{a}_{j_k} \equiv_{U\bar{a}[I_0]} \bar{a}_{i_0} \dots \bar{a}_{i_{k-1}} \bar{a}_{j_k}.$$

It follows that $\mathrm{tp}(\bar{a}_{j_0} \dots \bar{a}_{j_{k-1}} / U \cup \bar{a}[I] \cup \bar{a}_{j_k})$ is finitely satisfiable in $U \cup \bar{a}[I]$. Consequently,

$$\bar{a}_{j_0} \dots \bar{a}_{j_{k-1}} \overset{\mathrm{u}}{\surd}_{U \cup \bar{a}[I]} \bar{a}_{j_k} \quad \text{implies} \quad \bar{a}_{j_0} \dots \bar{a}_{j_{k-1}} \surd_{U \cup \bar{a}[I]} \bar{a}_{j_k},$$

as desired. ☐

For preforking relations that are contained in the splitting relation $\overset{\mathrm{s}}{\surd}$, we no not need to check for indiscernibility when proving that a given sequence is a Morley sequence.

Lemma 4.14. *Let $\alpha = (\bar{a}_i)_{i \in I}$ and $\beta = (\bar{b}_i)_{i \in I}$ be two sequences and $U \subseteq \mathbb{M}$ a set of parameters.*

(a) *If $\bar{b}_i \equiv_{U\bar{a}[<i]} \bar{a}_i$ and $\bar{b}_i \overset{\mathrm{s}}{\surd}_U \bar{a}[<i]\bar{b}[<i]$, for all $i \in I$, then $\alpha \equiv_U \beta$.*

(b) *If $\bar{a}_j \equiv_{U\bar{a}[<i]} \bar{a}_i$ and $\bar{a}_i \overset{\mathrm{s}}{\surd}_U \bar{a}[<i]$, for all $i \leq j$ in I, then α is indiscernible over U.*

Proof. (a) We prove by induction on $n < \omega$ that

$$\bar{a}[\bar{\imath}] \equiv_U \bar{b}[\bar{\imath}], \quad \text{for all } \bar{\imath} \in [I]^n .$$

For $n = 0$, the claim is trivial. For the inductive step, suppose that we have already proved it for n and consider a tuple of indices $\bar{\imath} \in [I]^{n+1}$. Setting $\bar{\imath}' := i_0 \dots i_{n-1}$ we have

$$\bar{a}[\bar{\imath}'] \equiv_U \bar{b}[\bar{\imath}'] \quad \text{and} \quad \bar{b}_{i_n} \sqrt[s]{}_U \, \bar{a}[\bar{\imath}']\bar{b}[\bar{\imath}'] ,$$

which implies that $\bar{a}[\bar{\imath}'] \equiv_{U\bar{b}_{i_n}} \bar{b}[\bar{\imath}']$. Since $\bar{b}_{i_n} \equiv_{U\bar{a}[<i_n]} \bar{a}_{i_n}$, it follows that

$$\bar{a}[\bar{\imath}']\bar{a}_{i_n} \equiv_U \bar{a}[\bar{\imath}']\bar{b}_{i_n} \equiv_U \bar{b}[\bar{\imath}']\bar{b}_{i_n} .$$

(b) We have to prove that

$$\bar{a}[\bar{\imath}] \equiv_U \bar{a}[\bar{\jmath}], \quad \text{for all } \bar{\imath}, \bar{\jmath} \in [I]^n, \ n < \omega .$$

Hence, let $\bar{\imath}, \bar{\jmath} \in [I]^n$. First, we consider the case where $i_s \leq j_s$, for all $s < n$. Then we have

$$\bar{a}_{j_s} \equiv_{U\bar{a}_{i_0}\dots\bar{a}_{i_{s-1}}} \bar{a}_{i_s} \quad \text{and} \quad \bar{a}_{j_s} \sqrt[s]{}_U \, \bar{a}_{i_0} \dots \bar{a}_{i_{s-1}} \bar{a}_{j_0} \dots \bar{a}_{j_{s-1}} ,$$

for all $s < n$. Consequently, it follows by (a) that $\bar{a}[\bar{\imath}] \equiv_U \bar{a}[\bar{\jmath}]$.

For the general case, let $\bar{\imath}, \bar{\jmath} \in [I]^n$ be arbitrary. We set

$$k_s := \max\{i_s, j_j\}, \quad \text{for } s < n .$$

Then $\bar{k} \in [I]^n$ and it follows by the special case considered above that $\bar{a}[\bar{\imath}] \equiv_U \bar{a}[\bar{k}] \equiv_U \bar{a}[\bar{\jmath}]$. □

As an application of Morley sequences we show that, for forking relations, right locality and symmetry are equivalent. One direction is based on the following two lemmas.

Lemma 4.15. *Let $\surd$-be a right local forking relation, $B, U \subseteq \mathbb{M}$ sets, and let $\kappa \geq \operatorname{loc}(\surd) \oplus |B|^+$ be a regular cardinal. For every reverse $\surd$-Morley sequence $(\bar{a}_i)_{i<\kappa}$ over U, there exists an index $\alpha < \kappa$ such that*

$$B\bar{a}[<\beta] \surd_U \bar{a}_\beta\,, \quad \textit{for all } \alpha \leq \beta < \kappa.$$

Proof. By (RLOC), there exists a set $U_0 \subseteq U \cup \bar{a}[<\kappa]$ of size

$$|U_0| < \operatorname{loc}(\surd) \oplus |B|^+ \leq \kappa$$

such that

$$B \surd_{U_0} U\bar{a}[<\kappa]\,.$$

Set $I := \{\, i < \kappa \mid \bar{a}_i \cap U_0 \neq \varnothing \,\}$. Then $|I| < \kappa$ and, by regularity of κ, there exists an index $\alpha < \kappa$ that is larger than every element of I. For $\alpha \leq \beta < \kappa$, it follows by (BMON) and monotonicity that $B \surd_{U\bar{a}[<\beta]} \bar{a}_\beta$. Since $(\bar{a}_i)_{i<\kappa}$ is a reverse $\surd$-Morley sequence, we furthermore have $\bar{a}[<\beta] \surd_U \bar{a}_\beta$. By Lemma 2.3, it follows that $B\bar{a}[<\beta] \surd_U \bar{a}_\beta$. □

Lemma 4.16. *Let $\surd$ be a right local preforking relation. If there exists a reverse $\surd$-Morley sequence $(\bar{a}_n)_{n<\omega}$ for $\operatorname{tp}(\bar{a}/BU)$ over U then $B \surd_U \bar{a}$.*

Proof. Set $\kappa := |B|^+ \oplus \operatorname{loc}(\surd)^+$ and let $(\bar{a}_n)_{n<\omega}$ be a reverse $\surd$-Morley sequence. By compactness, we can extend $(\bar{a}_n)_{n<\omega}$ to an indiscernible sequence $(\bar{a}_i)_{i<\kappa}$ over $B \cup U$ of length κ. By (FIN) and (INV) it follows that

$$\bar{a}[<\alpha] \surd_U \bar{a}_\alpha\,, \quad \text{for all } \alpha < \kappa\,.$$

Hence, $(\bar{a}_i)_{i<\kappa}$ is a reverse $\surd$-Morley sequence. By Lemma 4.15, there is some index $\alpha < \kappa$ with $B \surd_U \bar{a}_\alpha$. As $\bar{a}_\alpha \equiv_{UB} \bar{a}$, we can use (INV) to conclude that $B \surd_U \bar{a}$. □

Theorem 4.17 (Adler). *A forking relation $\surd$ is right local if, and only if, it is symmetric.*

Proof. ($\Leftarrow$) follows by Corollary 3.18.

($\Rightarrow$) If $\bar{a} \surd_U B$, we can use Proposition 4.10 and Lemma 4.12 to construct a reverse $\surd$-Morley sequence of $\operatorname{tp}(\bar{a}/UB)$ over U. Therefore, it follows by Lemma 4.16 that $B \surd_U \bar{a}$. □

F3. *Simple theories*

1. *Dividing and forking*

In this section we introduce the central forking relation of model theory, which is simply called *forking.*

Definition 1.1. Let T be a first-order theory, U a set of parameters, and $k < \omega$.

(a) We say that a set Φ of formulae over U is k*-inconsistent* (with respect to T) if $T(U) \cup \Phi_0$ is inconsistent, for every subset $\Phi_0 \subseteq \Phi$ of size $|\Phi_0| \geq k$.

(b) A formula $\varphi(\bar{x};\bar{c})$ with parameters $\bar{c}$ k*-divides* over U if there exists a sequence $(\bar{c}_n)_{n<\omega}$ such that

- $\bar{c}_n \equiv_U \bar{c}$, for all $n < \omega$, and
- the set $\{\, \varphi(\bar{x};\bar{c}_n) \mid n < \omega \,\}$ is k-inconsistent.

We say that $\varphi(\bar{x};\bar{c})$ *divides* over U if it k-divides over U, for some $k < \omega$.

(c) A set Φ of formulae *divides* over U if $T(\mathbb{M}) \cup \Phi \vDash \varphi$, for some formula φ that divides over U. We define

$$\bar{a} \mathop{\sqrt[\mathrm{d}]{}}_U B \quad :\text{iff} \quad \mathrm{tp}(\bar{a}/UB) \text{ does not divide over } U.$$

(d) A set Φ of formulae *forks* over U if there are finitely many formulae $\varphi_0, \dots, \varphi_{n-1}$ such that

$$T(\mathbb{M}) \cup \Phi \vDash \varphi_0 \vee \dots \vee \varphi_{n-1}$$

and each φ_i divides over U. We define

$$\bar{a} \mathop{\sqrt[\mathrm{f}]{}}_U B \quad :\text{iff} \quad \mathrm{tp}(\bar{a}/UB) \text{ does not fork over } U.$$

Example. (a) Consider the structure $\langle \mathbb{Q}, < \rangle$ and let $b < c$ be rational numbers. The formula $\varphi(x; b, c) := b < x \wedge x < c$ divides over the set $U := \{ a \in \mathbb{Q} \mid a < b \}$ since we can choose numbers b_n and c_n such that $b \leq b_0 < c_0 < b_1 < c_1 < \dots$. Then $b_n c_n \equiv_U bc$ and the set $\{ b_n < x \wedge x < c_n \mid n < \omega \}$ is 2-inconsistent.

(b) We consider the tree $\langle A^{<\omega}, \preceq \rangle$ where A is an infinite set. Fix a vertex $u_0 \in A^{<\omega}$, an element $a \in A$, and set $u := u_0 a$. The formula $\varphi(x; u) := u \preceq x$ divides over the set $U := \{ v \in A^{<\omega} \mid u_0 \npreceq v \}$ since, fixing distinct elements $b_n \in A$, for $n < \omega$, we can set $c_n := ub_n$. Then $c_n \equiv_U u$ and $\{ c_n \preceq x \mid n < \omega \}$ is 2-inconsistent.

Remark. Note that, if a formula φ divides over U and $\psi \vDash \varphi$, then ψ also divides over U. It follows that a formula φ divides over U if, and only if, the set $\{\varphi\}$ divides over U. Furthermore, if a set Φ divides over U, then there exists a finite subset $\Phi_0 \subseteq \Phi$ such that the formula $\bigwedge \Phi_0$ divides over U. In particular, a complete type $\mathfrak{p}$ divides over U if, and only if, some formula $\varphi \in \mathfrak{p}$ divides over U. The same holds for forking.

Below we will prove that $\sqrt[\mathrm{d}]{}$ is a preforking relation and $\sqrt[\mathrm{f}]{}$ the associated forking relation. Before doing so, let us give an alternative characterisation of dividing in terms of indiscernible sequences.

Lemma 1.2. *Let $\varphi(\bar{x}; \bar{y})$ be a formula and $\bar{c}, U \subseteq \mathbb{M}$. The following statements are equivalent:*

(1) *$\varphi(\bar{x}; \bar{c})$ divides over U.*

(2) *There exists an indiscernible sequence $(\bar{c}_n)_{n<\omega}$ over U such that $\bar{c}_0 = \bar{c}$ and the set $\{ \varphi(\bar{x}; \bar{c}_n) \mid n < \omega \}$ is k-inconsistent, for some $k < \omega$.*

(3) *There exists an indiscernible sequence $(\bar{c}_n)_{n<\omega}$ over U such that $\bar{c}_0 = \bar{c}$ and the set*

$$T(\bigcup_{n<\omega} \bar{c}_n) \cup \{ \varphi(\bar{x}; \bar{c}_n) \mid n < \omega \}$$

is inconsistent.

Proof. (2) ⇒ (1) and (2) ⇒ (3) are trivial.

(3) ⇒ (2) Let $(\bar{c}_n)_{n<\omega}$ be an indiscernible sequence over U with $\bar{c}_0 = \bar{c}$ such that

$$T(\bigcup_{n<\omega} \bar{c}_n) \cup \{\, \varphi(\bar{x};\bar{c}_n) \mid n < \omega \,\}$$

is inconsistent. Then there exists a finite subset $I \subseteq \omega$ such that

$$T(\bigcup_{n\in I} \bar{c}_n) \cup \{\, \varphi(\bar{x};\bar{c}_n) \mid n \in I \,\}$$

is inconsistent. Let $n_0 < \dots < n_{k-1}$ be an enumeration of I. For every k-tuple of indices $i_0 < \dots < i_{k-1}$, $\bar{c}[\bar{\imath}] \equiv_U \bar{c}[\bar{n}]$ implies that

$$T(\bar{c}_{i_0} \dots \bar{c}_{i_{k-1}}) \cup \{\varphi(\bar{x};\bar{c}_{i_0}), \dots, \varphi(\bar{x};\bar{c}_{i_{k-1}})\}$$

is inconsistent. Hence, $\{\, \varphi(\bar{x};\bar{c}_n) \mid n < \omega \,\}$ is k-inconsistent.

(1) ⇒ (2) Suppose that $\varphi(\bar{x};\bar{c})$ divides over U. Then there exists a sequence $(\bar{c}_n)_{n<\omega}$ such that $\bar{c}_n \equiv_U \bar{c}$ and $\{\, \varphi(\bar{x};\bar{c}_n) \mid n < \omega \,\}$ is k-inconsistent, for some k. By Proposition E5.3.6, there exists an indiscernible sequence $(\bar{d}_n)_{n<\omega}$ over U with

$$\mathrm{Av}((c_n)_n/U) \subseteq \mathrm{Av}((d_n)_n/U)\,.$$

In particular, $\mathrm{tp}(\bar{c}/U) \subseteq \mathrm{Av}((d_n)_n/U)$ and

$$\neg\exists\bar{z}[\varphi(\bar{z};\bar{x}_0) \wedge \dots \wedge \varphi(\bar{z};\bar{x}_{k-1})] \in \mathrm{Av}((d_n)_n/U)\,.$$

Consequently, $\bar{d}_0 \equiv_U \bar{c}$ and the set $\{\, \varphi(\bar{x};\bar{d}_n) \mid n < \omega \,\}$ is k-inconsistent. Fixing an automorphism $\pi \in \mathrm{Aut}\,\mathbb{M}_U$ with $\pi(\bar{d}_0) = \bar{c}$, we obtain a sequence $(\pi(\bar{d}_n))_{n<\omega}$ with the desired properties. □

Exercise 1.1. Prove that a formula $\varphi(\bar{x};\bar{c})$ divides over a set U if, and only if, it divides over some model $M \supseteq U$. (*Hint.* Use Lemma E5.3.11.)

Lemma 1.3. *The following statements are equivalent:*

(1) $\bar{a} \not\mathop{\smile}\limits^{\mathrm{d}}_U \bar{b}$

(2) *For every infinite indiscernible sequence $(\bar b_i)_{i \in I}$ over U with $\bar b = \bar b_i$, for some i, there exists a tuple $\bar a' \equiv_{U\bar b} \bar a$ such that $(\bar b_i)_{i \in I}$ is indiscernible over $U \cup \bar a'$.*

(3) *For every indiscernible sequence $(\bar b_n)_{n<\omega}$ over U with $\bar b = \bar b_0$, there is some $\bar a' \equiv_{U\bar b} \bar a$ such that*

$$\bar b_m \equiv_{U\bar a'} \bar b_n\,, \quad \text{for all } m, n < \omega\,.$$

Proof. (2) ⇒ (3) is trivial.

(3) ⇒ (1) Suppose that $\bar a \not\mathrel{\smile\!\!\!\!|}^{\mathrm{d}}_U \bar b$. By Lemma 1.2, we can find a formula $\varphi(\bar x; \bar c) \in \mathrm{tp}(\bar a/U\bar b)$ and an indiscernible sequence $(\bar c_n)_{n<\omega}$ over U such that $\bar c_n \equiv_U \bar c$ and $\{\, \varphi(\bar x; \bar c_n) \mid n < \omega \,\}$ is k-inconsistent, for some $k < \omega$. By adding and permuting free variables of φ, we may assume that $\bar c_n = \bar b_n \bar d$ where $\bar d \subseteq U$ and $\bar b_n \equiv_U \bar b$, for all n. Finally, applying an automorphism of $\mathbb{M}$, we may assume that $\bar b_0 = \bar b$.

To show that (3) fails, consider a tuple $\bar a' \equiv_{U\bar b} \bar a$. Then

$$\mathbb{M} \vDash \varphi(\bar a'; \bar b_0 \bar d)\,,$$

but the k-inconsistency of $\{\, \varphi(\bar x; \bar b_n \bar d) \mid n < \omega \,\}$ implies that there is some $n < k$ with

$$\mathbb{M} \nvDash \varphi(\bar a'; \bar b_n \bar d)\,.$$

Consequently, $\bar b_n \not\equiv_{U\bar a'} \bar b_0$.

(1) ⇒ (3) Consider an indiscernible sequence $(\bar b_n)_{n<\omega}$ over U with $\bar b = \bar b_0$ and suppose that there is no such tuple $\bar a'$. Then the set

$$\mathrm{tp}(\bar a/U\bar b) \cup \{\, \varphi(\bar x; \bar b_i) \leftrightarrow \varphi(\bar x; \bar b_j) \mid i, j < \omega \text{ and } \varphi(\bar x; \bar y) \text{ a formula over } U \,\}$$

is inconsistent. This set is equivalent to the union

$$\bigcup_{n<\omega} \mathfrak{p}(\bar x, \bar b_n)\,, \quad \text{where} \quad \mathfrak{p}(\bar x, \bar x') := \mathrm{tp}(\bar a \bar b/U)\,.$$

By compactness, we can therefore find a finite subset $\Phi \subseteq \mathfrak{p}$ and indices $n_0 < \dots < n_{k-1} < \omega$ such that

$$T \cup \Phi(\bar{x}, \bar{b}_{n_0}) \cup \dots \cup \Phi(\bar{x}, \bar{b}_{n_{k-1}})$$

is inconsistent. Setting $\varphi := \bigwedge \Phi$ it follows by indiscernibility that

$$T \vDash \neg\exists\bar{x}[\varphi(\bar{x}, \bar{b}_{i_0}) \wedge \dots \wedge \varphi(\bar{x}, \bar{b}_{i_{k-1}})],$$

for every increasing tuple $i_0 < \dots < i_{k-1}$. Hence, $\{\, \varphi(\bar{x}, \bar{b}_n) \mid n < \omega \,\}$ is k-inconsistent and φ divides over U. Consequently, $\bar{a} \not\mathop{\smile\!\!\!\!|}^{\mathrm{d}}_U \bar{b}$.

(3) ⇒ (2) Let $(\bar{b}_i)_{i \in I}$ be an infinite indiscernible sequence over U with $\bar{b}_{i_0} = \bar{b}$, for some $i_0 \in I$. Setting

$$\Psi := \{\, \psi(\bar{x}; \bar{b}[\bar{\imath}]) \leftrightarrow \psi(\bar{x}; \bar{b}[\bar{k}]) \mid \psi \text{ a formula over } U \text{ and } \operatorname{ord}(\bar{\imath}) = \operatorname{ord}(\bar{k}) \,\},$$

it is sufficient to prove that $\operatorname{tp}(\bar{a}/U\bar{b}) \cup \Psi$ is satisfiable.

Fix a dense linear order $J \supseteq I$ without end points. Using Lemma E5.3.9, we can extend $(\bar{b}_i)_{i \in I}$ to an indiscernible sequence $(\bar{b}_i)_{i \in J}$ over U. By (3) and compactness, there exists a tuple $\bar{a}' \equiv_{U\bar{b}} \bar{a}$ such that

$$\bar{b}_i \equiv_{U\bar{a}'} \bar{b}_j, \quad \text{for all } i, j \in J.$$

To show that $\operatorname{tp}(\bar{a}/U\bar{b}) \cup \Psi$ is satisfiable, let $\Psi_0 \subseteq \Psi$ be finite and let $I_0 \subseteq I$ be the finite set of all indices i such that Ψ_0 contains the constants $\bar{b}_i$. By the Theorem of Ramsey, there exist an order embedding $h_0 : I_0 \to J$ such that the sequence $(\bar{b}_{h(i)})_{i \in I_0}$ is indiscernible over $U \cup \bar{a}'$ with respect to the formulae in Ψ_0. We extend $h_0 : I_0 \to J$ to an order embedding $h : I_0 \cup \{i_0\} \to J$. There exists an automorphism $\pi \in \operatorname{Aut} \mathbb{M}_U$ mapping $\bar{b}_{h(i)}$ to $\bar{b}_i$, for $i \in I_0 \cup \{i_0\}$. Then the tuple $\pi(\bar{a}')$ satisfies $\bigcup_{i \in I_0 \cup \{i_0\}} \operatorname{tp}(\bar{a}/U\bar{b}_i) \cup \Psi_0$. In particular, it satisfies $\operatorname{tp}(\bar{a}/U\bar{b}) \cup \Psi_0$. □

Remark. Comparing the statement in (2) above with Lemma E5.3.11, we see that, when $\bar{a} \mathop{\smile\!\!\!\!|}^{\mathrm{d}}_U B$, we can choose $\bar{a}' \equiv_{UB} \bar{a}$ while, in general, we only find $\bar{a}' \equiv_U \bar{a}$.

Example. (a) Consider the structure $\langle \mathbb{Q}, < \rangle$ and let $b < a < c$ be elements. Then $bc \mathop{\smile\hspace{-0.9em}\mid}^{\mathrm{d}}_{\varnothing} a$ but $a \mathop{\not\smile\hspace{-0.9em}\mid}^{\mathrm{d}}_{\varnothing} bc$. In particular, $\mathop{\smile\hspace{-0.9em}\mid}^{\mathrm{d}}$ is not symmetric.

We have already seen above that $\varphi(x; b, c) := b < x \wedge x < c$ divides over $\downarrow b$ and, hence, also over the empty set. Consequently, $a \mathop{\not\smile\hspace{-0.9em}\mid}^{\mathrm{d}}_{\varnothing} bc$. To show that $bc \mathop{\smile\hspace{-0.9em}\mid}^{\mathrm{d}}_{\varnothing} a$, let $(a_i)_{i<\omega}$ be an indiscernible sequence over $\varnothing$. Choose elements b' and c' such that $b' < a < c'$ and $b' < a_i < c'$, for all $i < \omega$. Then $b'c' \equiv_a bc$ and $(a_i)_{i<\omega}$ is indiscernible over $\{b', c'\}$. By Lemma 1.3, it follows that $bc \mathop{\smile\hspace{-0.9em}\mid}^{\mathrm{d}}_{\varnothing} a$.

(b) Let $\langle A, \sim \rangle$ be a structure where $\sim$ is an equivalence relation with infinitely many classes all of which are infinite. Fix elements $a, b \in A$ and a set $U \subseteq A$. Then

$$a \mathop{\smile\hspace{-0.9em}\mid}^{\mathrm{d}}_{U} b \quad \text{iff} \quad \begin{array}{l} \{a\} \cap \{b\} \subseteq U \text{ and,} \\ a \nsim b \text{ or there is some } c \in U \text{ with } b \sim c\,. \end{array}$$

Let us show next that $\mathop{\smile\hspace{-0.9em}\mid}^{\mathrm{d}}$ is a preforking relation, that $\mathop{\smile\hspace{-0.9em}\mid}^{\mathrm{f}}$ is the corresponding forking relation, and that acl is the closure operator associated with them.

Proposition 1.4. $\mathop{\smile\hspace{-0.9em}\mid}^{\mathrm{d}}$ *is a preforking relation.*

Proof. Throughout the proof we will tacitly make use of the characterisation of $\mathop{\smile\hspace{-0.9em}\mid}^{\mathrm{d}}$ from Lemma 1.3.

(INV) follows immediately from the definition.

(MON) Suppose that $\bar{a}_0\bar{a}_1 \mathop{\smile\hspace{-0.9em}\mid}^{\mathrm{d}}_{U} B$ and let $B_0 \subseteq B$. For a contradiction, suppose that $\bar{a}_0 \mathop{\not\smile\hspace{-0.9em}\mid}^{\mathrm{d}}_{U} B_0$. Then we can find a formula $\varphi \in \mathrm{tp}(\bar{a}_0/UB_0)$ that divides over U. Hence, $\varphi \in \mathrm{tp}(\bar{a}_0\bar{a}_1/UB)$ implies that $\bar{a}_0\bar{a}_1 \mathop{\not\smile\hspace{-0.9em}\mid}^{\mathrm{d}}_{U} B$. A contradiction.

(NOR) Suppose that $\bar{a} \mathop{\smile\hspace{-0.9em}\mid}^{\mathrm{d}}_{\bar{c}} \bar{b}$. To show that $\bar{a}\bar{c} \mathop{\smile\hspace{-0.9em}\mid}^{\mathrm{d}}_{\bar{c}} \bar{b}\bar{c}$, let $(\bar{b}_n\bar{c}_n)_{n<\omega}$ be an indiscernible sequence over $\bar{c}$ with $\bar{b}_0\bar{c}_0 = \bar{b}\bar{c}$. Then $\bar{c}_n = \bar{c}$, for all n. Since $\bar{a} \mathop{\smile\hspace{-0.9em}\mid}^{\mathrm{d}}_{\bar{c}} \bar{b}$, there is a tuple $\bar{a}' \equiv_{\bar{b}\bar{c}} \bar{a}$ such that $(\bar{b}_n)_{n<\omega}$ is indiscernible over $\bar{a}'\bar{c}$. Hence, $(\bar{b}_n\bar{c})_{n<\omega}$ is also indiscernible over $\bar{a}'\bar{c}$. As $\bar{a}'\bar{c} \equiv_{\bar{b}\bar{c}} \bar{a}\bar{c}$, the claim follows.

(LRF) Let $\bar a, \bar b$ be tuples. To show that $\bar a \mathop{\smile\mkern-11mu|}^{\mathrm d}_{\bar a} \bar b$ it is sufficient to note that every indiscernible sequence $(\bar b_n)_{n<\omega}$ over $\bar a$ is also indiscernible over $\bar a \cup \bar a$.

(LTR) Suppose that $\bar a_0 \bar a_1 \bar a_2 \mathop{\smile\mkern-11mu|}^{\mathrm d}_{\bar a_0 \bar a_1} \bar b$ and $\bar a_0 \bar a_1 \mathop{\smile\mkern-11mu|}^{\mathrm d}_{\bar a_0} \bar b$. Let $(\bar b_n)_{n<\omega}$ be an infinite indiscernible sequence over $\bar a_0$ such that $\bar b_0 = \bar b$. We have to find tuples

$$\bar a_0'' \bar a_1'' \bar a_2'' \equiv_{\bar a_0 \bar b} \bar a_0 \bar a_1 \bar a_2$$

such that $(\bar b_n)_{n<\omega}$ is indiscernible over $\bar a_0'' \bar a_1'' \bar a_2''$. Since $\bar a_0 \bar a_1 \mathop{\smile\mkern-11mu|}^{\mathrm d}_{\bar a_0} \bar b$, there are tuples $\bar a_0' \bar a_1' \equiv_{\bar a_0 \bar b} \bar a_0 \bar a_1$ such that $(\bar b_n)_{n<\omega}$ is indiscernible over $\bar a_0' \bar a_1'$. Let $\bar a_2'$ be a tuple such that

$$\bar a_0' \bar a_1' \bar a_2' \equiv_{\bar a_0 \bar b} \bar a_0 \bar a_1 \bar a_2 \,.$$

Then $\bar a_0' \bar a_1' \bar a_2' \mathop{\smile\mkern-11mu|}^{\mathrm d}_{\bar a_0' \bar a_1'} \bar b$ and there are tuples

$$\bar a_0'' \bar a_1'' \bar a_2'' \equiv_{\bar a_0' \bar a_1' \bar b} \bar a_0' \bar a_1' \bar a_2'$$

such that $(\bar b_n)_{n<\omega}$ is indiscernible over $\bar a_0'' \bar a_1'' \bar a_2''$. Since

$$\bar a_0'' = \bar a_0 \quad \text{and} \quad \bar a_0'' \bar a_1'' \bar a_2'' \equiv_{\bar a_0 \bar b} \bar a_0 \bar a_1 \bar a_2$$

the claim follows.

(BMON) Suppose that $\bar a \mathop{\smile\mkern-11mu|}^{\mathrm d}_{\bar c} \bar b \bar d$. To show that $\bar a \mathop{\smile\mkern-11mu|}^{\mathrm d}_{\bar c \bar d} \bar b$, let $(\bar b_n)_{n<\omega}$ be a sequence of indiscernibles over $\bar c \bar d$ with $\bar b_0 = \bar b$. Then $(\bar b_n \bar d)_{n<\omega}$ is indiscernible over $\bar c$. Consequently, there is some tuple $\bar a' \equiv_{\bar c \bar b \bar d} \bar a$ such that $(\bar b_n \bar d)_{n<\omega}$ is indiscernible over $\bar a' \bar c$. It follows that $(\bar b_n)_{n<\omega}$ is indiscernible over $\bar a' \bar c \bar d$.

(DEF) Suppose that $\bar a \not\mathop{\smile\mkern-11mu|}^{\mathrm d}_U B$. Then there exists a formula $\varphi(\bar x; \bar b) \in \mathrm{tp}(\bar a/UB)$ that divides over U. For every $\bar a' \in \varphi(\bar x; \bar b)^{\mathbb M}$ it follows that $\mathrm{tp}(\bar a'/U\bar b)$ divides over U. □

Before proving that $\mathop{\smile\mkern-11mu|}^{\mathrm f}$ is the forking relation associated with $\mathop{\smile\mkern-11mu|}^{\mathrm d}$, let us show that forking satisfies the axiom (EXT) even for incomplete types.

Lemma 1.5. *A partial type Φ over $U \cup C$ forks over U if, and only if, every complete type $\mathfrak{p} \in \langle \Phi \rangle$ forks over U.*

Proof. Clearly, if Φ forks over U, then so does every type containing Φ. Conversely, suppose that every $\mathfrak{p} \in \langle \Phi \rangle$ forks over U. For each $\mathfrak{p} \in \langle \Phi \rangle$, we fix a formula $\varphi_{\mathfrak{p}} \in \mathfrak{p}$ that forks over U. By compactness,

$$\langle \Phi \rangle = \{ \mathfrak{p} \mid \mathfrak{p} \in \langle \Phi \rangle \} \subseteq \bigcup_{\mathfrak{p} \in \langle \Phi \rangle} \langle \varphi_{\mathfrak{p}} \rangle$$

implies that there are finitely many types $\mathfrak{p}_0, \dots, \mathfrak{p}_{n-1} \in \langle \Phi \rangle$ such that

$$\langle \Phi \rangle \subseteq \langle \varphi_{\mathfrak{p}_0} \rangle \cup \dots \cup \langle \varphi_{\mathfrak{p}_{n-1}} \rangle .$$

Consequently, $\Phi \vDash \varphi_{\mathfrak{p}_0} \vee \dots \vee \varphi_{\mathfrak{p}_{n-1}}$ and Φ forks over U. ☐

Proposition 1.6. $\overset{\mathrm{f}}{\surd} = {}^*(\overset{\mathrm{d}}{\surd})$

Proof. ($\subseteq$) To prove that $\overset{\mathrm{f}}{\surd} \subseteq {}^*(\overset{\mathrm{d}}{\surd})$, note that $\overset{\mathrm{f}}{\surd} \subseteq \overset{\mathrm{d}}{\surd}$ and that the operation $\surd \mapsto \overset{*}{\surd}$ is monotone. Therefore, it is sufficient to prove that $\overset{\mathrm{f}}{\surd} = {}^*(\overset{\mathrm{f}}{\surd})$, i.e., that $\overset{\mathrm{f}}{\surd}$ satisfies (EXT). Hence, suppose that $\bar{a} \overset{\mathrm{f}}{\surd}_U B$ and let C be an arbitrary set. By Lemma 1.5, there exists a complete type $\mathfrak{p}$ over $U \cup B \cup C$ that contains $\mathrm{tp}(\bar{a}/UB)$ and that does not fork over U. Fix a realisation $\bar{a}'$ of $\mathfrak{p}$. Then $\bar{a}' \equiv_{UB} \bar{a}$ and $\bar{a}' \overset{\mathrm{f}}{\surd}_U BC$.

($\supseteq$) Suppose that $\bar{a} \overset{\mathrm{f}}{\not\surd}_U B$. Then we can find finitely many formulae $\varphi_0(\bar{x}; \bar{c}_0), \dots, \varphi_{n-1}(\bar{x}; \bar{c}_{n-1})$ that each divide over U and such that

$$\mathrm{tp}(\bar{a}/UB) \vDash \varphi_0(\bar{x}; \bar{c}_0) \vee \dots \vee \varphi_{n-1}(\bar{x}; \bar{c}_{n-1}) .$$

For every tuple $\bar{a}' \equiv_{UB} \bar{a}$, there is some $i < n$ such that $\mathbb{M} \vDash \varphi_i(\bar{a}'; \bar{c}_i)$. Consequently,

$$\bar{a}' \overset{\mathrm{d}}{\not\surd}_U B\bar{c}_0 \dots \bar{c}_{n-1} , \quad \text{for all } \bar{a}' \equiv_{UB} \bar{a} .$$

Hence, $\bar{a} \; {}^*(\overset{\mathrm{d}}{\surd})_U \; B$ does not hold. ☐

Corollary 1.7. $\overset{\mathrm{f}}{\surd}$ *is a forking relation.*

Lemma 1.8. $\mathrm{cl}_{\overset{\mathrm{f}}{\surd}} = \mathrm{cl}_{\overset{\mathrm{d}}{\surd}} = \mathrm{acl}$

Proof. By Lemma F2.4.6, it is sufficient to prove that $\mathrm{cl}_{\overset{\mathrm{d}}{\surd}} = \mathrm{acl}$.

For one inclusion, let $a \notin \mathrm{acl}(U)$. Then there exists an indiscernible sequence $(a_n)_{n<\omega}$ over U with $a_0 = a$ and $a_i \neq a_k$, for $i \neq k$. Since a is the only element realising $\mathrm{tp}(a/Ua)$ and $(a_n)_n$ is not indiscernible over $U \cup \{a\}$ it follows by Lemma 1.3 that $a \overset{\mathrm{d}}{\not\surd}_U a$.

Conversely, suppose that there are sets B, C such that $a \overset{\mathrm{d}}{\not\surd}_{UC} B$. By Lemma 1.2, we can find a formula $\varphi(x;\bar{c}) \in \mathrm{tp}(a/UCB)$ and an indiscernible sequence $(\bar{c}_n)_{n<\omega}$ such that $\bar{c}_0 = \bar{c}$ and $\{\, \varphi(x;\bar{c}_n) \mid n < \omega \,\}$ is k-inconsistent, for some k. For every $n < \omega$, fix an element a_n such that $a_n\bar{c}_n \equiv_U a\bar{c}$. Since $\mathbb{M} \vDash \varphi(a_n;\bar{c}_n)$ and $\{\, \varphi(x;\bar{c}_n) \mid n < \omega \,\}$ is k-inconsistent, there exists an infinite subset $I \subseteq \omega$ such that $a_i \neq a_j$, for distinct $i, j \in I$. As each a_n satisfies $\mathrm{tp}(a/U)$ it follows that $a \notin \mathrm{acl}(U)$. □

At first sight, the definition of $\overset{\mathrm{d}}{\surd}$ might seem rather ad-hoc. The following result indicates that $\overset{\mathrm{d}}{\surd}$ plays a rather distinguished role: it is the largest preforking relation that is contained in every symmetric forking relation.

Theorem 1.9. $\overset{\mathrm{d}}{\surd} \subseteq \downarrow$, *for every symmetric forking relation* $\downarrow$.

Proof. Suppose that $\bar{a} \overset{\mathrm{d}}{\surd}_U \bar{b}$. Since $\downarrow$ is symmetric, (LRF) implies that $B \downarrow_U U$. Therefore, we can use Proposition F2.4.10 and Lemma F2.4.12 to construct a reverse $\downarrow$-Morley sequence $(\bar{b}_n)_{n<\omega}$ for $\mathrm{tp}(\bar{b}/U)$ over U. By (INV) we may assume that $\bar{b}_0 = \bar{b}$. Since $\bar{a} \overset{\mathrm{d}}{\surd}_U \bar{b}$ there is a tuple $\bar{a}' \equiv_{U\bar{b}} \bar{a}$ such that $(\bar{b}_n)_{n<\omega}$ is indiscernible over $U\bar{a}'$. Hence, $(\bar{b}_n)_{n<\omega}$ is a reverse $\downarrow$-Morley sequence for $\mathrm{tp}(\bar{b}/U)$ over $U\bar{a}'$. Since $\downarrow$ is right local, it follows by Lemma F2.4.16 that $\bar{a}' \downarrow_U \bar{b}$. By invariance we obtain $\bar{a} \downarrow_U \bar{b}$. □

Remark. In the next section we will show that there are theories where $\overset{\mathrm{d}}{\surd}$ is symmetric and a forking relation. For such theories, $\overset{\mathrm{d}}{\surd}$ is the largest preforking relation that is contained in every symmetric forking relation.

To conclude this section we compare $\overset{\mathrm{d}}{\surd}$ and $\overset{\mathrm{f}}{\surd}$ with the preforking relations introduced in Section F2.3. First, let us introduce the forking relation associated with the splitting relation $\overset{\mathrm{s}}{\surd}$.

Definition 1.10. $\overset{\mathrm{i}}{\surd} := {}^{*}(\overset{\mathrm{s}}{\surd})$.

Lemma 1.11. $\overset{\mathrm{i}}{\surd} \subseteq \overset{\mathrm{d}}{\surd}$

Proof. Suppose that $\bar{a} \overset{\mathrm{i}}{\surd}_U B$. To show that $\bar{a} \overset{\mathrm{d}}{\surd}_U B$, consider a formula $\varphi(\bar{x};\bar{c}) \in \mathrm{tp}(\bar{a}/UB)$ and let $(\bar{c}_n)_{n<\omega}$ be a sequence such that $\bar{c}_n \equiv_U \bar{c}$, for all n. We have to show that the set $\{\varphi(\bar{x};\bar{c}_n) \mid n < \omega\}$ is not k-inconsistent for any k.

There is a tuple $\bar{a}' \equiv_{UB} \bar{a}$ such that

$$\bar{a}' \overset{\mathrm{s}}{\surd}_U B\bar{c}[<\omega]\,.$$

Hence, $\varphi(\bar{x};\bar{c}) \in \mathrm{tp}(\bar{a}'/UB\bar{c}[<\omega])$ implies that

$$\varphi(\bar{x};\bar{c}_n) \in \mathrm{tp}(\bar{a}'/UB\bar{c}[<\omega])\,, \quad \text{for all } n\,.$$

Consequently, $\bar{a}'$ satisfies $\{\varphi(\bar{x};\bar{c}_n) \mid n < \omega\}$ and this set is not k-inconsistent. ☐

Proposition 1.12. $\overset{\mathrm{u}}{\surd} \subseteq \overset{\mathrm{i}}{\surd} \subseteq \overset{\mathrm{f}}{\surd} \subseteq \overset{\mathrm{d}}{\surd}$

Proof. The inclusions $\overset{\mathrm{u}}{\surd} \subseteq \overset{\mathrm{i}}{\surd} \subseteq \overset{\mathrm{d}}{\surd}$ follow from Theorem F2.3.13 and the preceding lemma, respectively. Since the operation $\surd \mapsto \overset{*}{\surd}$ is monotone and idempotent, we further have $\overset{\mathrm{i}}{\surd} = {}^{*}(\overset{\mathrm{i}}{\surd}) \subseteq {}^{*}(\overset{\mathrm{d}}{\surd}) = \overset{\mathrm{f}}{\surd}$. ☐

2. *Simple theories and the tree property*

The aim of this section is to characterise those theories where the relation $\overset{\mathrm{f}}{\surd}$ is symmetric. In the same way as stable theories are characterised by the absence of the order property, we will present a combinatorial property causing $\overset{\mathrm{f}}{\surd}$ to be non-symmetric.

Definition 2.1. A first-order theory T is *simple* if $\sqrt[\mathrm{f}]{}$ is symmetric. For simple theories we will write $\downfree^{\mathrm{f}}$ and $\downfree^{\mathrm{d}}$ instead of $\sqrt[\mathrm{f}]{}$ and $\sqrt[\mathrm{d}]{}$. In later chapters, where $\downfree^{\mathrm{f}}$ will be the only forking relation under consideration, we will frequently drop the superscript and just write $\downfree$.

Before giving a combinatorial characterisation of simple theories, let us note some special properties of the relation $\downfree^{\mathrm{f}}$ in such theories. It follows from Theorem 1.9 that, for complete types in simple theories, forking and dividing is the same. According to the next lemma this is also true for partial types.

Lemma 2.2. *Let T be a simple theory, $\Phi(\bar{x};\bar{y})$ a set of formulae over U, and $\bar{c} \subseteq \mathbb{M}$. The following statements are equivalent:*

(1) *$\Phi(\bar{x};\bar{c})$ forks over U.*

(2) *$\Phi(\bar{x};\bar{c})$ divides over U.*

(3) *For every $\downfree^{\mathrm{f}}$-Morley sequence $(\bar{c}_n)_{n<\omega}$ for $\mathrm{tp}(\bar{c}/U)$ over U, the set $\bigcup_{i<\omega}\Phi(\bar{x};\bar{c}_n)$ is inconsistent.*

Proof. (2) ⇒ (1) follows immediately from the definition of forking.

(3) ⇒ (2) Let $(\bar{c}_n)_{n<\omega}$ be a $\downfree^{\mathrm{f}}$-Morley sequence for $\mathrm{tp}(\bar{c}/U)$ over U. Applying a U-automorphism we can ensure that $\bar{c}_0 = \bar{c}$. By assumption, $\bigcup_{n<\omega}\Phi(\bar{x};\bar{c}_n)$ is inconsistent. Using compactness, we obtain a finite subset $\Phi_0 \subseteq \Phi$ such that $\bigcup_{n<\omega}\Phi_0(\bar{x};\bar{c}_n)$ is inconsistent. Set $\varphi := \bigwedge\Phi_0$. By Lemma 1.2, it follows that $\varphi(\bar{x};\bar{c})$ divides over U. Since $\Phi(\bar{x};\bar{c}) \vDash \varphi(\bar{x};\bar{c})$, so does $\Phi(\bar{x};\bar{c})$.

(1) ⇒ (3) Suppose that $(\bar{c}_n)_{n<\omega}$ is a $\downfree^{\mathrm{f}}$-Morley sequence for $\mathrm{tp}(\bar{c}/U)$ over U such that the set $\bigcup_{n<\omega}\Phi(\bar{x};\bar{c}_n)$ is consistent. Fix a regular cardinal $\kappa \geq \mathrm{loc}(\downfree^{\mathrm{f}}) \oplus |\bar{x}|^+$. By compactness, there exists a $\downfree^{\mathrm{f}}$-Morley sequence $(\bar{c}_i)_{i<\kappa}$ for $\mathrm{tp}(\bar{c}/U)$ over U such that $\bigcup_{i<\kappa}\Phi(\bar{x};\bar{c}_i)$ is consistent. Let $\bar{a}$ be a tuple satisfying this set. By Lemma F2.4.15, we can find an index $\alpha < \kappa$ such that

$$\bar{a}\bar{c}[<\alpha] \downfree^{\mathrm{f}}_U \bar{c}_\alpha .$$

Consequently, $\Phi(\bar{x};\bar{c}_\alpha)$ does not fork over U. By (INV), the same holds for $\Phi(\bar{x};\bar{c})$. $\square$

Next, we present an improved version of Lemma 1.3.

Proposition 2.3 (Kim). *Let T be a simple theory. The following statements are equivalent.*

(1) $\bar{a} \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{d}}_U \bar{b}$

(2) $\bar{a} \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}_U \bar{b}$

(3) *For every infinite* $\mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}$*-Morley sequence* $(\bar{b}_i)_{i\in I}$ *for* $\mathrm{tp}(\bar{b}/U)$ *over U there exists a tuple* $\bar{a}' \equiv_{U\bar{b}} \bar{a}$ *such that* $(\bar{b}_i)_{i\in I}$ *is a* $\mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}$*-Morley sequence over* $U \cup \bar{a}'$.

(4) *For some* $\mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}$*-Morley sequence* $(\bar{b}_i)_{i<\omega}$ *for* $\mathrm{tp}(\bar{b}/U)$ *over U there exists a tuple* $\bar{a}' \equiv_{U\bar{b}} \bar{a}$ *such that* $(\bar{b}_i)_{i<\omega}$ *is a* $\mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}$*-Morley sequence over* $U \cup \bar{a}'$.

Proof. (1) $\Leftrightarrow$ (2) has already been shown in Lemma 2.2 and (1) $\Rightarrow$ (3) is a special case of Lemma 1.3.

(3) $\Rightarrow$ (4) is trivial since we have seen in Corollary F2.4.11 that, for symmetric forking relations, Morley sequences always exist.

(4) $\Rightarrow$ (2) Let $(\bar{b}_i)_{i<\omega}$ be a $\mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}$-Morley sequence for $\mathrm{tp}(\bar{b}/U)$ over $U\cup\bar{a}'$, for some $\bar{a}' \equiv_{U\bar{b}} \bar{a}$. Set $\mathfrak{p}(\bar{x},\bar{x}') := \mathrm{tp}(\bar{a}\bar{b}/U)$. Then $\bar{a}'$ realises $\mathfrak{p}(\bar{x},\bar{b})$. Hence, $\bar{a}'$ is a realisation of $\bigcup_{i<\omega} \mathfrak{p}(\bar{x},\bar{b}_i)$ and it follows by Lemma 2.2 that $\mathfrak{p}(\bar{x},\bar{b})$ does not fork over U. $\square$

Right locality

Note that, if the relation $\sqrt[\mathrm{f}]{}$ is right local, then $\sqrt[\mathrm{f}]{} \subseteq \sqrt[\mathrm{d}]{}$ implies that $\sqrt[\mathrm{d}]{}$ is also right local. (This is also a consequence of Lemma 2.2.) In this section we will prove that the converse is also true: if $\sqrt[\mathrm{d}]{}$ is right local, then so is $\sqrt[\mathrm{f}]{}$. Recall the notion of a $\sqrt{}$-forking chain introduced in Section F2.3.

Definition 2.4. (a) We call $\sqrt[\mathrm{d}]{}$-forking chains and $\sqrt[\mathrm{f}]{}$-forking chains *dividing chains* and *forking chains,* respectively.

(b) A *specification* of a dividing chain $(\bar{b}_\alpha)_{\alpha<\gamma}$ for $\bar{a}$ over U is a sequence $\langle \varphi_\alpha, k_\alpha \rangle_{\alpha<\gamma}$ of pairs consisting of a formula $\varphi_\alpha(\bar{x};\bar{y}_\alpha)$ and a natural number k_α such that, for all $\alpha < \gamma$,

$$\mathbb{M} \vDash \varphi_\alpha(\bar{a};\bar{b}_\alpha) \quad \text{and} \quad \varphi_\alpha(\bar{x};\bar{b}_\alpha)\ k_\alpha\text{-divides over } U \cup \bar{b}[<\alpha]\,.$$

Similarly, a *specification* of a forking chain $(\bar{b}_\alpha)_{\alpha<\gamma}$ for $\bar{a}$ over U is a sequence $\langle \varphi_\alpha, \bar{\psi}_\alpha, \bar{k}_\alpha, m_\alpha \rangle_{\alpha<\gamma}$, where φ_α is a formula, m_α a natural number, $\bar{\psi}_\alpha$ an m_α-tuple of formulae, and $\bar{k}_\alpha$ is an m_α-tuple of natural numbers such that, for all $\alpha < \gamma$,

$$\mathbb{M} \vDash \varphi_\alpha(\bar{a};\bar{b}_\alpha)$$

and there are tuples $\bar{d}_0, \dots, \bar{d}_{m_\alpha-1}$ such that

$$\varphi_\alpha(\bar{x};\bar{b}_\alpha) \vDash \psi_{\alpha,0}(\bar{x},\bar{d}_0) \vee \dots \vee \psi_{\alpha,m_\alpha-1}(\bar{x},\bar{d}_{m_\alpha-1})$$

and each $\psi_{\alpha,i}(\bar{x},\bar{d}_i)$ $k_{\alpha,i}$-divides over $U \cup \bar{b}[<\alpha]$.

(c) A dividing chain is *uniform* if it has a specification $\langle \varphi_\alpha, k_\alpha \rangle_{\alpha<\gamma}$ where

$$\varphi_\alpha = \varphi_\beta \quad \text{and} \quad k_\alpha = k_\beta\,, \quad \text{for all } \alpha, \beta < \gamma\,.$$

Similarly, we say that a forking chain is *uniform* if it has a specification $\langle \varphi_\alpha, \bar{\psi}_\alpha, \bar{k}_\alpha, m_\alpha \rangle_{\alpha<\gamma}$ where

$$\varphi_\alpha = \varphi_\beta\,, \quad m_\alpha = m_\beta\,, \quad \psi_{\alpha,i} = \psi_{\beta,i}\,, \quad k_{\alpha,i} = k_{\beta,i}\,,$$

for all $\alpha, \beta < \gamma$ and $i < m_\alpha$.

Note that, according to Theorem F2.3.25, $\overset{\mathrm{d}}{\surd}$ is not right local if, and only if, there are arbitrarily long dividing chains. The same holds for $\overset{\mathrm{f}}{\surd}$ and forking chains. Our aim is therefore to show that, if a theory has arbitrarily long forking chains, then there are also arbitrarily long dividing chains. We start with the observation that any subsequence of a forking chain is again a forking chain. As a consequence we can use the Pigeon Hole Principle to construct uniform forking chains.

Lemma 2.5. *Let γ be an ordinal and $I \subseteq \gamma$.*

(a) *If $(\bar{b}_\alpha)_{\alpha<\gamma}$ is a dividing chain for $\bar{a}$ over U with the specification $\langle \varphi_\alpha, k_\alpha \rangle_{\alpha<\gamma}$, then $(\bar{b}_\alpha)_{\alpha \in I}$ is a dividing chain for $\bar{a}$ over U with specification $\langle \varphi_\alpha, k_\alpha \rangle_{\alpha \in I}$.*

(b) *If $(\bar{b}_\alpha)_{\alpha<\gamma}$ is a forking chain for $\bar{a}$ over U with the specification $\langle \varphi_\alpha, \bar{\psi}_\alpha, \bar{k}_\alpha, m_\alpha \rangle_{\alpha<\gamma}$, then $(\bar{b}_\alpha)_{\alpha \in I}$ is a forking chain for $\bar{a}$ over U with specification $\langle \varphi_\alpha, \bar{\psi}_\alpha, \bar{k}_\alpha, m_\alpha \rangle_{\alpha \in I}$.*

Proof. (a) Fix $\alpha \in I$ and set $B := \bigcup\{ \bar{b}_\beta \mid \beta \in I,\ \beta < \alpha \}$. It is sufficient to show that $\varphi_\alpha(\bar{x}; \bar{b}_\alpha)$ k_α-divides over $U \cup B$. This follows from the definition of dividing and the fact that $\varphi_\alpha(\bar{x}; \bar{b}_\alpha)$ k_α-divides over the superset $U \cup \bar{b}[<\alpha] \supseteq U \cup B$.

(b) follows analogously. □

Corollary 2.6. *Let $\kappa > |T|$ be a cardinal. If there exists a forking chain for $\bar{a}$ over U of length κ, then there also exists a uniform forking chain for $\bar{a}$ over U of length κ.*

Proof. Let $(\bar{b}_\alpha)_{\alpha<\kappa}$ be a forking chain for $\bar{a}$ over U with specification $\langle \varphi_\alpha, \bar{\psi}_\alpha, \bar{k}_\alpha, m_\alpha \rangle_{\alpha<\kappa}$. Since there are at most $|T| < \kappa$ formulae over $\varnothing$, there exist a subset $I \subseteq \kappa$ of size $|I| = \kappa$, formulae φ, $\bar{\psi}$, and numbers m, $\bar{k}$ such that

$$\varphi_\alpha = \varphi\,, \quad m_\alpha = m\,, \quad \psi_{\alpha,i} = \psi_i\,, \quad k_{\alpha,i} = k_i\,,$$

for all $\alpha < \kappa$ and $i < m$. By Lemma 2.5, the subsequence $(\bar{b}_\alpha)_{\alpha \in I}$ is a uniform forking chain for $\bar{a}$ over U. □

The key property of dividing which allows us to turn forking chains into dividing chains is contained in the following lemma.

Lemma 2.7. *Suppose that the formula $\varphi(\bar{x}; \bar{b})$ k-divides over a set U. For every set $C \subseteq \mathbb{M}$, there is some tuple $\bar{b}' \equiv_U \bar{b}$ such that $\varphi(\bar{x}; \bar{b}')$ k-divides over $U \cup C$.*

Proof. By Lemma 1.2, there exists an indiscernible sequence $(\bar{b}_n)_{n<\omega}$ over U such that $\bar{b}_0 = \bar{b}$ and the set $\{ \varphi(\bar{x};\bar{b}_n) \mid n < \omega \}$ is k-inconsistent. Using Lemma E5.3.11, we can find a set $C' \equiv_U C$ such that $(\bar{b}_n)_{n<\omega}$ is indiscernible over $U \cup C'$. Let $\pi \in \operatorname{Aut}\mathbb{M}_U$ be an automorphism with $\pi[C'] = C$, and set $\bar{b}'_n := \pi(\bar{b}_n)$. Then $(\bar{b}'_n)_{n<\omega}$ is indiscernible over $U \cup C$ and the set $\{ \varphi(\bar{x};\bar{b}'_n) \mid n < \omega \}$ is k-inconsistent. By Lemma 1.2, it follows that $\varphi(\bar{x};\bar{b}'_0)$ k-divides over $U \cup C$. Since $\bar{b}'_0 \equiv_U \bar{b}_0 = \bar{b}$, the claim follows. ☐

Corollary 2.8. *Let $(\bar{b}_i)_{i<n}$ be a dividing chain for $\bar{a}$ over U with finite length. For every set $C \subseteq \mathbb{M}$, there exist tuples*

$$\bar{a}'\bar{b}'_0 \dots \bar{b}'_{n-1} \equiv_U \bar{a}\bar{b}_0 \dots \bar{b}_{n-1}$$

such that $(\bar{b}'_i)_{i<n}$ is a dividing chain for $\bar{a}'$ over $U \cup C$ with the same specification as $(\bar{b}_i)_{i<n}$.

Proof. Let $\langle \varphi_i, k_i \rangle_{i<n}$ be a specification of $(\bar{b}_i)_{i<n}$. We prove the claim by induction on n. For $n = 0$, there is nothing to do. Hence, suppose that $n > 0$. We can use Lemma 2.7 to find a tuple $\bar{b}'_0 \equiv_U \bar{b}_0$ such that $\varphi_0(\bar{x};\bar{b}'_0)$ k_0-divides over $U \cup C$. Let $\pi \in \operatorname{Aut}\mathbb{M}_U$ be an automorphism with $\pi(\bar{b}_0) = \bar{b}'_0$. Then $(\pi(\bar{b}_i))_{0<i<n}$ is a dividing chain for $\pi(\bar{a})$ over $U \cup \bar{b}'_0$. Applying the inductive hypothesis to it, we obtain tuples

$$\bar{a}'\bar{b}'_1 \dots \bar{b}'_{n-1} \equiv_{U\bar{b}'_0} \pi(\bar{a})\pi(\bar{b}_1) \dots \bar{\pi}(b_{n-1})$$

such that $(\bar{b}'_i)_{0<i<n}$ is a dividing chain for $\bar{a}'$ over $U \cup C \cup \bar{b}'_0$. Since

$$\bar{a}'\bar{b}'_0\bar{b}'_1 \dots \bar{b}'_{n-1} \equiv_U \pi(\bar{a})\bar{b}'_0\pi(\bar{b}_1) \dots \bar{\pi}(b_{n-1}) \equiv_U \bar{a}\bar{b}_0\bar{b}_1 \dots \bar{b}_{n-1}\,,$$

it follows that $(\bar{b}'_i)_{i<n}$ is the desired dividing chain for $\bar{a}'$ over $U \cup C$. ☐

In order to turn a forking chain into a dividing chain, we iterate the following construction.

Lemma 2.9. *Let* $(\bar{b}_i)_{i<n}$ *be a dividing chain for* $\bar{a}$ *over* $U \cup C$ *with a finite length* n *and with the specification* $\langle \varphi_i, k_i \rangle_{i<n}$. *If*

$$\operatorname{tp}(\bar{a}/UC) \vDash \vartheta_0(\bar{x}; \bar{d}_0) \vee \dots \vee \vartheta_{m-1}(\bar{x}; \bar{d}_{m-1}),$$

where each $\vartheta_j(\bar{x}; \bar{d}_j)$ l_j*-divides over* U, *then there exist an index* $j < m$ *and a tuple* $\bar{d}' \equiv_U \bar{d}_j$ *such that* $\bar{d}', \bar{b}_0, \dots, \bar{b}_{n-1}$ *is a dividing chain for* $\bar{a}$ *over* U *with specification*

$$\langle \vartheta_j, l_j \rangle, \langle \varphi_0, k_0 \rangle, \dots, \langle \varphi_{n-1}, k_{n-1} \rangle.$$

Proof. We prove the claim by induction on n. For $n = 0$, pick an index j such that $\mathbb{M} \vDash \vartheta_j(\bar{a}; \bar{d}_j)$. Then $\bar{d}_j$ is a dividing chain for $\bar{a}$ over U with specification $\langle \vartheta_j, l_j \rangle$. Hence, suppose that $n > 0$. By Corollary 2.8, there exist tuples

$$\bar{a}' \bar{b}'_0 \dots \bar{b}'_{n-1} \equiv_{UC} \bar{a} \bar{b}_0 \dots \bar{b}_{n-1}$$

such that $(\bar{b}'_i)_{i<n}$ is a dividing chain for $\bar{a}'$ over $U \cup C \cup \bar{d}_0 \dots \bar{d}_{n-1}$. Since $\bar{a}' \equiv_{UC} \bar{a}$, there is some index $j < m$ such that

$$\mathbb{M} \vDash \vartheta_j(\bar{a}'; \bar{d}_j).$$

It follows that $\bar{d}_j, \bar{b}'_0, \dots, \bar{b}'_{n-1}$ is a dividing chain for $\bar{a}'$ over U with specification

$$\langle \vartheta_j, l_j \rangle, \langle \varphi_0, k_0 \rangle, \dots, \langle \varphi_{n-1}, k_{n-1} \rangle.$$

Fix a tuple $\bar{d}'$ such that

$$\bar{a} \bar{d}' \bar{b}_0 \dots \bar{b}_{n-1} \equiv_U \bar{a}' \bar{d}_j \bar{b}'_0 \dots \bar{b}'_{n-1}.$$

Then $\bar{d}', \bar{b}_0, \dots, \bar{b}_{n-1}$ is the desired dividing chain. □

Corollary 2.10. *Let* $(\bar{b}_i)_{i<n}$ *be a uniform forking chain for* $\bar{a}$ *over* U *with specification* $\langle \varphi, \bar{\psi}, \bar{k}, m \rangle_{i<n}$. *There exists a function* $g : [n] \to [m]$ *and a dividing chain* $(\bar{b}'_i)_{i<n}$ *for* $\bar{a}$ *over* U *with specification*

$$\langle \psi_{g(0)}, k_{g(0)} \rangle, \dots, \langle \psi_{g(n-1)}, k_{g(n-1)} \rangle.$$

Proof. We prove the claim by induction on n. For $n = 0$, there is nothing to do. Hence, suppose that $n > 0$. Applying the inductive hypothesis to the subchain $(\bar{b}_i)_{0<i<n}$ we obtain a dividing chain $(\bar{b}'_i)_{0<i<n}$ for $\bar{a}$ over $U \cup \bar{b}_0$ with specification

$$\langle \psi_{g(1)}, k_{g(1)} \rangle, \dots, \langle \psi_{g(n-1)}, k_{g(n-1)} \rangle.$$

Since $\mathbb{M} \vDash \varphi(\bar{a}; \bar{b}_0)$ and

$$\varphi(\bar{x}; \bar{b}_0) \vDash \psi_0(\bar{x}; \bar{d}_0) \vee \dots \vee \psi_{m-1}(\bar{x}; \bar{d}_{m-1}),$$

for suitable $\bar{d}_0, \dots, \bar{d}_{m-1}$, we can use Lemma 2.9 to find an index $j < m$ and a tuple $\bar{b}'_0 \equiv_U \bar{d}_j$ such that $(\bar{b}'_i)_{i<n}$ is a dividing chain for $\bar{a}$ over U with specification

$$\langle \psi_j, k_j \rangle, \langle \psi_{g(1)}, k_{g(1)} \rangle, \dots, \langle \psi_{g(n-1)}, k_{g(n-1)} \rangle. \qquad \square$$

Starting from a sufficiently long forking chain, we have constructed arbitrarily long finite dividing chains. According to the next lemma, this is sufficient to obtain dividing chains of every ordinal length.

Lemma 2.11. *Let φ be a formula and $k < \omega$ a number. If, for each $n < \omega$, there exists a uniform dividing chain for $\bar{a}$ over U of length n with specification $\langle \varphi, k \rangle_{i<n}$, then, for every ordinal γ, we can find a uniform dividing chain for $\bar{a}$ over U of length γ with specification $\langle \varphi, k \rangle_{\alpha<\gamma}$.*

Proof. Let γ be an ordinal. We define the following set of formulae with variables $\bar{x}$, $\bar{y}^\alpha$, $\bar{z}^\alpha_i$, for $\alpha < \gamma$ and $i < \omega$.

$$\begin{aligned} \Phi := {} & \{ \varphi(\bar{x}; \bar{y}^\alpha) \mid \alpha < \gamma \} \\ & \cup \{ \psi(\bar{z}^\alpha_i; \bar{y}^{\beta_0}, \dots, \bar{y}^{\beta_{n-1}}) \leftrightarrow \psi(\bar{y}^\alpha; \bar{y}^{\beta_0}, \dots, \bar{y}^{\beta_{n-1}}) \mid \\ & \qquad \psi \text{ a formula over } U,\ i, n < \omega, \text{ and} \\ & \qquad \beta_0 < \dots < \beta_{n-1} < \alpha < \gamma \} \\ & \cup \{ \neg\exists\bar{x}[\varphi(\bar{x}; \bar{z}^\alpha_{i_0}) \wedge \dots \wedge \varphi(\bar{x}; \bar{z}^\alpha_{i_{k-1}})] \mid \\ & \qquad \alpha < \gamma,\ i_0 < \dots < i_{k-1} < \omega \}. \end{aligned}$$

Note that, if $\bar{a}$, $\bar{b}^\alpha$, and $\bar{c}^\alpha_i$, for $\alpha < \gamma$ and $i < n$, satisfy Φ, then

$$\bar{c}^\alpha_i \equiv_{U\bar{b}[<\alpha]} \bar{b}^\alpha$$

and the set $\{\, \varphi(\bar{x};\bar{c}^\alpha_i) \mid i < \omega \,\}$ is k-inconsistent. Hence, the formula $\varphi(\bar{x};\bar{b}^\alpha)$ k-divides over $U\bar{b}[<\alpha]$. Consequently, $(\bar{b}^\alpha)_{\alpha<\gamma}$ is a dividing chain for $\bar{a}$ over U with specification $\langle \varphi, k\rangle_{\alpha<\gamma}$.

It therefore remains to show that Φ is satisfiable. Let $\Phi_0 \subseteq \Phi$ be finite and let $I \subseteq \gamma$ be the finite set of indices α such that Φ_0 contains some of the variables $\bar{y}^\alpha$ or $\bar{z}^\alpha_i$, for $i < \omega$. Choose a uniform dividing chain $(\bar{b}_i)_{i<n}$ for $\bar{a}$ over U of length $n := |I|$. We can satisfy Φ_0 by interpreting $\bar{x}$ by $\bar{a}$, $\bar{y}^\alpha$ by the corresponding $\bar{b}_i$, and $\bar{z}^\alpha_i$ by tuples witnessing the fact that $\varphi(\bar{x};\bar{b}_i)$ k-divides over $U \cup \bar{b}[<i]$. By the Compactness Theorem, it follows that Φ is satisfiable. □

Combining the results of this section, we have proved that, if $\overset{\mathrm{f}}{\surd}$ is not right local, then neither is $\overset{\mathrm{d}}{\surd}$.

Theorem 2.12. *Let T be a complete first-order theory. The following statements are equivalent:*

(1) $\overset{\mathrm{d}}{\surd}$ *is right local.*

(2) $\overset{\mathrm{f}}{\surd}$ *is right local.*

(3) *There is no dividing chain of length $|T|^+$.*

Proof. (2) ⇒ (1) If $\overset{\mathrm{f}}{\surd}$ is right local, then T is simple. Hence, it follows by Lemma 2.2 that $\overset{\mathrm{d}}{\surd} = \overset{\mathrm{f}}{\surd}$. In particular, $\overset{\mathrm{d}}{\surd}$ is right local.

(1) ⇒ (3) If there are arbitrarily long dividing chains, it follows by Theorem F2.3.25 that $\overset{\mathrm{d}}{\surd}$ is not right local.

(3) ⇒ (2) Suppose that $\overset{\mathrm{f}}{\surd}$ is not right local and set $\kappa := |T|^+$. By Theorem F2.3.25, there exists a forking chain of length κ for a suitable tuple $\bar{a}$ over the empty set $\varnothing$. Using Corollary 2.6 we obtain a uniform forking chain of the same length. Let $\langle \varphi, \bar{\psi}, \bar{k}, m\rangle_{\alpha<\kappa}$ be its specification. According to Corollary 2.10, there exists, for every $n < \omega$, a dividing

chain of length n with specification $\langle \vartheta_i, l_i \rangle_{i<n}$, where $\vartheta_i \in \bar{\psi}$ and $l_i \in \bar{k}$, for every $i < n$.

By the Pigeon Hole Principle and Lemma 2.5, we can find a formula $\vartheta \in \bar{\psi}$ and a number $l \in \bar{k}$ such that, for every $n < \omega$, there exists a uniform dividing chain of length n with specification $\langle \vartheta, l \rangle_{i<n}$. Consequently, it follows from Lemma 2.11 that there exist arbitrarily long dividing chains. □

The tree property

The following combinatorial property characterises simple theories in the same way as the order property characterises stable theories.

Definition 2.13. Let T be a first-order theory. A formula $\varphi(\bar{x}; \bar{y})$ has the *tree property* if there exists a family $(\bar{c}_\eta)_{\eta \in \omega^{<\omega}}$ of parameters and a number $k < \omega$ such that

- for every $\beta \in \omega^\omega$, the set $\{ \varphi(\bar{x}; \bar{c}_\eta) \mid \eta \prec \beta \}$ is consistent and
- for every $\eta \in \omega^{<\omega}$, the set $\{ \varphi(\bar{x}; \bar{c}_{\eta i}) \mid i < \omega \}$ is k-inconsistent.

Exercise 2.1. Prove that, in the theory of dense linear orders, the formula $\varphi(x; y_0, y_1) := y_0 < x \wedge x < y_1$ has the tree property.

Before proving that a theory is simple if, and only if, no formula has the tree property, let us note that the tree property implies the order property.

Lemma 2.14. *Every formula with the tree property has the order property.*

Proof. Let $(\bar{c}_\eta)_{\eta \in \omega^{<\omega}}$ be a family witnessing the tree property of the formula $\varphi(\bar{x}; \bar{y})$. For every $\beta \in \omega^\omega$, we choose a tuple $\bar{a}_\beta$ satisfying $\{ \varphi(\bar{x}; \bar{c}_\eta) \mid \eta \prec \beta \}$. To prove that φ has the order property it is sufficient to find indices $\eta_0 \prec \eta_1 \prec \dots$ in $\omega^{<\omega}$ and a sequence $(\beta_n)_{n<\omega}$ in ω^ω such that $\eta_n \prec \beta_n$ and

$$\mathbb{M} \vDash \varphi(\bar{a}_{\beta_i}; \bar{c}_{\eta_k}) \quad \text{iff} \quad i \leq k\,.$$

We proceed by induction on n, starting with $\eta_0 := \langle\rangle$ and an arbitrary $\beta_0 \in \omega^\omega$. For the inductive step, suppose that η_n and β_n are already defined. The k-inconsistency of $\{\, \varphi(\bar{x}; \bar{c}_{\eta_n i}) \mid i < \omega \,\}$ implies that, for each $m \leq n$, there are only finitely many $i < \omega$ such that

$$\mathbb{M} \vDash \varphi(\bar{a}_{\beta_m}; \bar{c}_{\eta_n i}) .$$

Hence, there is some $i < \omega$ such that

$$\mathbb{M} \vDash \neg\varphi(\bar{a}_{\beta_m}; \bar{c}_{\eta_n i}) , \quad \text{for all } m \leq n .$$

We set $\eta_{n+1} := \eta_n i$, for such an index i, and we choose some $\beta_{n+1} \in \omega^\omega$ such that $\eta_{n+1} \prec \beta_{n+1}$. Then $\eta_m \prec \beta_{n+1}$ implies that

$$\mathbb{M} \vDash \varphi(\bar{a}_{\beta_{n+1}}; \bar{c}_{\eta_m}) , \quad \text{for all } m \leq n + 1 . \qquad \square$$

To show that simple theories are exactly those where no formula has the tree property, we introduce a generalised form of the tree property.

Definition 2.15. Let κ be a cardinal, γ an ordinal, $(\varphi_\alpha)_{\alpha<\gamma}$ a sequence of formulae, and $(k_\alpha)_{\alpha<\gamma}$ a sequence of numbers.

(a) A family $(\bar{c}_\eta)_{\eta\in\kappa^{\leq\gamma}}$ of tuples $\bar{c}_\eta \subseteq \mathbb{M}$ is a *dividing κ-tree* with *specification* $\langle\varphi_\alpha, k_\alpha\rangle_{\alpha<\gamma}$ if

- for each $\beta \in \kappa^\gamma$, the set $\{\, \varphi_\alpha(\bar{x}; \bar{c}_{\beta\restriction(\alpha+1)}) \mid \alpha < \gamma \,\}$ is consistent,
- for each $\eta \in \kappa^{<\gamma}$, the set $\{\, \varphi_{|\eta|}(\bar{x}; \bar{c}_{\eta\alpha}) \mid \alpha < \kappa \,\}$ is $k_{|\eta|}$-inconsistent.

We call γ the *height* of the dividing κ-tree.

(b) A dividing κ-tree $(\bar{c}_\eta)_{\eta\in\kappa^{\leq\gamma}}$ with specification $\langle\varphi_\alpha, k_\alpha\rangle_{\alpha<\gamma}$ is *uniform* if

$$\varphi_\alpha = \varphi_\beta \quad \text{and} \quad k_\alpha = k_\beta , \quad \text{for all } \alpha, \beta < \gamma .$$

Remark. Note that a formula $\varphi(\bar{x}; \bar{y})$ has the tree property if, and only if, there exists a uniform dividing ω-tree of height ω with specification $\langle\varphi, k\rangle_{n<\omega}$, for some $k < \omega$.

Lemma 2.16. *Let $\kappa > |T|$ be a cardinal. If there exists a dividing ω-tree of height κ, then there also exists an uniform dividing ω-tree of height ω.*

Proof. Let $(\bar{b}_\eta)_{\eta \in \omega^{<\kappa}}$ be a dividing ω-tree of height κ and let $\langle \varphi_\alpha, k_\alpha \rangle_{\alpha < \kappa}$ be its specification. Since $\kappa > |T|$, there exist a subset $I \subseteq \kappa$ of size $|I| = \kappa$, a formula φ_*, and a number $k_* < \omega$ such that

$$\varphi_\alpha = \varphi_* \quad \text{and} \quad k_\alpha = k_*\,, \quad \text{for all } \alpha \in I\,.$$

Choose a strictly increasing map $h : \omega \to I$. We inductively define an embedding $g : \omega^{<\omega} \to \omega^{<\kappa}$ as follows. We start with $g(\langle\rangle) := \langle\rangle$. If $g(\eta)$ is already defined, we choose some $\zeta \in \omega^{<\kappa}$ with $g(\eta) \preceq \zeta$ and $|\zeta| = h(|\eta|)$, and we set $g(\eta i) := \zeta i$, for $i < \omega$.

We claim that the family $(\bar{b}_{g(\eta)})_{\eta \in \omega^{<\omega}}$ is a uniform dividing ω-tree of height ω. By construction, the set $\{\, \varphi_*(\bar{x}; \bar{b}_{g(\eta n)}) \mid n < \omega \,\}$ is k_*-inconsistent, for every $\eta \in \omega^{<\omega}$. Furthermore, for each $\beta \in \omega^\omega$, we can choose some $\beta' \in \omega^{<\kappa}$ with

$$\beta' \succeq g(\beta \upharpoonright \alpha)\,, \quad \text{for all } \alpha < \omega\,,$$

and we see that

$$\big\{\, \varphi_*(\bar{x}; \bar{b}_{g(\eta)}) \;\big|\; \eta \prec \beta \,\big\} \subseteq \big\{\, \varphi_\alpha(\bar{x}; \bar{b}_{\beta' \upharpoonright (\alpha+1)}) \;\big|\; \alpha < \gamma \,\big\}$$

is consistent. □

The following lemma contains the main technical argument we use to relate the tree property to dividing.

Lemma 2.17. *The following statements are equivalent:*

(1) *There exists a dividing ω-tree of height γ.*

(2) *There exists a dividing chain of length γ.*

Proof. (1) ⇒ (2) Set $\kappa := (2^{|T| \oplus |\gamma|})^+$. If there is a dividing ω-tree, we can use the Compactness Theorem to construct a dividing κ-tree $(\bar{b}_\eta)_{\eta \in \kappa^{\leq \gamma}}$.

Let $\langle \varphi_\alpha, k_\alpha \rangle_{\alpha<\gamma}$ be its specification. We define an embedding $h : \kappa^{\leq\gamma} \to \kappa^{\leq\gamma}$ as follows. We start with $h(\langle\rangle) := \langle\rangle$. If $|\eta|$ is a limit ordinal, we set

$$h(\eta) := \sup \{ h(\zeta) \mid \zeta \prec \eta \} .$$

For the successor step, we proceed as follows. Suppose that the value of $h(\eta)$ is already defined. Let $\bar{s}$ be the sorts of $\bar{b}_{\eta 0}$. As $|S^{\bar{s}}(\bigcup_{\zeta\leq\eta} \bar{b}_\zeta)| < \kappa$ there exists a subset $I \subseteq \kappa$ of size $|I| = \kappa$ such that

$$\bar{b}_{\eta i} \equiv_{\bigcup_{\zeta\leq\eta} \bar{b}_\zeta} \bar{b}_{\eta k} , \quad \text{for all } i, k \in I .$$

We fix a bijection $g : \kappa \to I$ and we set $h(\eta i) := h(\eta) g(i)$.

Having defined the embedding h, we fix some $\beta \in \kappa^{<\omega}$ and we set $\bar{c}_\alpha := \bar{b}_{h(\beta\restriction(\alpha+1))}$, for $\alpha < \gamma$. We claim that the sequence $(\bar{c}_\alpha)_{\alpha<\gamma}$ is a dividing chain for some $\bar{a}$ over $\varnothing$ with specification $\langle \varphi_\alpha, k_\alpha \rangle_{\alpha<\gamma}$.

Set $\beta' := \sup \{ h(\beta \restriction \alpha) \mid \alpha < \gamma \}$ and choose some tuple $\bar{a}$ satisfying

$$\{ \varphi_\alpha(\bar{x}; \bar{b}_{\beta'\restriction(\alpha+1)}) \mid \alpha < \gamma \} .$$

Then

$$\begin{aligned} \{ \varphi_\alpha(\bar{x}; \bar{c}_\alpha) \mid \alpha < \gamma \} &= \{ \varphi_\alpha(\bar{x}; \bar{b}_{h(\beta\restriction(\alpha+1))}) \mid \alpha < \gamma \} \\ &= \{ \varphi_\alpha(\bar{x}; \bar{b}_{\beta'\restriction(\alpha+1)}) \mid \alpha < \gamma \} , \end{aligned}$$

implies that

$$\mathbb{M} \vDash \varphi_\alpha(\bar{a}; \bar{c}_\alpha) , \quad \text{for all } \alpha < \gamma .$$

It therefore remains to show that $\varphi_\alpha(\bar{x}; \bar{c}_\alpha)$ k_α-divides over $\bar{c}[<\alpha]$. Let $\bar{a}_n := \bar{b}_{h((\beta\restriction\alpha)n)}$, for $n < \omega$. Then $\bar{a}_n \equiv_{\bar{c}[<\alpha]} \bar{b}_{h(\beta\restriction(\alpha+1))} = \bar{c}_\alpha$ and the set $\{ \varphi_\alpha(\bar{x}; \bar{a}_n) \mid n < \omega \}$ is k_α-inconsistent.

(2) ⇒ (1) Given a dividing chain $(\bar{c}_\alpha)_{\alpha<\gamma}$ for $\bar{a}$ over U with specification $\langle \varphi_\alpha, k_\alpha \rangle_{\alpha<\gamma}$, we construct a dividing ω-tree $(\bar{b}_\eta)_{\eta\in\omega^{\leq\gamma}}$ with the additional property that, for every $\eta \in \omega^{\leq\gamma}$,

$$(\bar{b}_{\eta\restriction(\alpha+1)})_{\alpha<|\eta|} \equiv_\varnothing (\bar{c}_\alpha)_{\alpha<|\eta|} .$$

If $\eta = \langle\rangle$ or if $|\eta|$ is a limit ordinal, we can choose an arbitrary tuple $\bar{b}_\eta$, since the definition of a dividing ω-tree places no constraint on such tuples. Hence, it remains to consider the successor step. Suppose that $\bar{b}_\eta$ has already been defined and set $\alpha := |\eta|$. Since

$$(\bar{b}_{\eta\restriction(i+1)})_{i<\alpha} \equiv_\varnothing (\bar{c}_i)_{i<\alpha} .$$

there exists some $\bar{b}'$ such that

$$(\bar{b}_{\eta\restriction(i+1)})_{i<\alpha}\bar{b}' \equiv_\varnothing (\bar{c}_i)_{i<\alpha}\bar{c}_\alpha .$$

Since $\varphi_\alpha(\bar{x};\bar{c}_\alpha)$ k_α-divides over $U \cup \bar{c}[<\alpha]$, we can find a sequence $(\bar{c}'_n)_{n<\omega}$ such that $\bar{c}'_n \equiv_{U\bar{c}[<\alpha]} \bar{c}_\alpha$ and $\{\, \varphi_\alpha(\bar{x};\bar{c}'_n) \mid n < \omega \,\}$ is k_α-inconsistent. By choice of $\bar{b}'$, we can therefore find a sequence $(\bar{b}'_n)_{n<\omega}$ such that

$$\bar{b}'_n \equiv_{\bigcup_{i<\alpha} \bar{b}_{\eta\restriction(i+1)}} \bar{b}'$$

and $\{\, \varphi_\alpha(\bar{x};\bar{b}'_n) \mid n < \omega \,\}$ is k_α-inconsistent. We set $\bar{b}_{\eta i} := \bar{b}'_i$, for $i < \omega$.

To see that the family $(\bar{b}_\eta)_{\eta\in\omega^{\leq\gamma}}$ constructed in this way is a dividing ω-tree, note that, for each $\beta \in \omega^\gamma$, $(\bar{b}_{\eta\restriction(\alpha+1)})_{\alpha<\gamma} \equiv_\varnothing (\bar{c}_\alpha)_{\alpha<\gamma}$ implies that the set $\{\, \varphi_\alpha(\bar{x};\bar{b}_{\beta\restriction(\alpha+1)}) \mid \alpha < \gamma \,\}$ is consistent. □

Using these two lemmas, we obtain the following characterisation of simple theories.

Theorem 2.18. *Let T be a complete first-order theory. The following statements are equivalent:*

(1) *T is simple.*

(2) *$\sqrt[\mathrm{d}]{}$ is right local.*

(3) *No formula has the tree property.*

(4) *There is no dividing chain of length $|T|^+$.*

(5) *For some cardinal κ, there is no dividing chain of length κ.*

Proof. (4) ⇒ (5) is trivial and (1) ⇔ (2) ⇔ (4) was already shown in Theorem 2.12.

(5) ⇒ (3) Suppose that there exists a formula $\varphi(\bar{x};\bar{y})$ with the tree property. Fix a family $(\bar{c}_\eta)_{\eta\in\omega^{<\omega}}$ and a number $k < \omega$ witnessing this fact.

For every cardinal κ, we will construct a dividing chain of length κ. Given κ, we use compactness to find a family $(\bar{b}_\eta)_{\eta\in\omega^{<\kappa}}$ such that

- for every $\beta \in \omega^\kappa$, the set $\{\, \varphi(\bar{x};\bar{b}_\eta) \mid \eta \prec \beta \,\}$ is consistent and
- for every $\eta \in \omega^{<\kappa}$, the set $\{\, \varphi(\bar{x};\bar{b}_{\eta i}) \mid i < \omega \,\}$ is k-inconsistent.

In particular, $(\bar{b}_\eta)_{\eta\in\omega^{<\kappa}}$ is a uniform dividing ω-tree of height κ. Hence, we can use Lemma 2.17 to obtain a dividing chain of length κ. A contradiction.

(3) ⇒ (4) Suppose that there exists a dividing chain of length $\kappa := |T|^+$. We will show that some formula has the tree property. By Lemma 2.17, there exists a dividing ω-tree $(\bar{b}_\eta)_{\eta\in\omega^{<\kappa}}$ of height κ. Hence, we can use Lemma 2.16 to obtain a uniform dividing ω-tree $(\bar{b}'_\eta)_{\eta\in\omega^{<\omega}}$ of height ω. Let $\langle\varphi, k\rangle_{n<\omega}$ be its specification. Then the formula φ has the tree property. A contradiction. □

Corollary 2.19. *Every stable theory is simple.*

Proof. This follows by Theorem 2.18 and Lemma 2.14. □

Corollary 2.20. *A theory T is simple if, and only if, T^{eq} is simple.*

Proof. Clearly, if φ has the tree property with respect to T, it also has the tree property with respect to T^{eq}. Conversely, if φ has the tree property with respect to T^{eq} we can use Proposition E2.2.10 to construct a formula φ' that has the tree property with respect to T. □

Finally, we show that no simple theory has the strict order property. Consequently, all simple theories that are not stable have the independence property.

Proposition 2.21. *No simple theory has the strict order property.*

Proof. Suppose that the formula $\varphi(\bar{x};\bar{y})$ has the strict order property. We will show that the formula

$$\psi(\bar{x};\bar{y}_0\bar{y}_1) := \neg\varphi(\bar{x};\bar{y}_0) \wedge \varphi(\bar{x};\bar{y}_1)$$

has the tree property. By compactness, there exists a sequence $(\bar{c}_i)_{i\in\mathbb{Q}}$ such that

$$\varphi(\bar{x};\bar{c}_i)^{\mathbb{M}} \subset \varphi(\bar{x};\bar{c}_k)^{\mathbb{M}}, \quad \text{for all } i < k\,.$$

We define two functions $\lambda, \rho : \omega^{<\omega} \to \mathbb{Q}$ such that $\lambda(\eta) < \rho(\eta)$, for all η. We proceed by induction on $\eta \in \omega^{<\omega}$ starting with $\lambda(\langle\rangle) := 0$ and $\rho(\langle\rangle) := 1$. If $\lambda(\eta) < \rho(\eta)$ are already defined, we choose a strictly increasing sequence $\lambda(\eta) < z_0 < z_1 < \cdots < \rho(\eta)$ and we set $\lambda(\eta i) := z_i$ and $\rho(\eta i) := z_{i+1}$, for $i < \omega$.

Having defined λ and ρ, we set $\bar{b}_\eta := \bar{c}_{\lambda(\eta)}\bar{c}_{\rho(\eta)}$, for $\eta \in \omega^{<\omega}$. To show that this family witnesses the tree property of ψ, note that

$$\psi(\bar{x};\bar{b}_\eta)^{\mathbb{M}} = \varphi(\bar{x};\bar{c}_{\rho(\eta)})^{\mathbb{M}} \setminus \varphi(\bar{x};\bar{c}_{\lambda(\eta)})^{\mathbb{M}}.$$

Hence,

$$\begin{aligned} &\psi(\bar{x};\bar{b}_\eta)^{\mathbb{M}} \subseteq \psi(\bar{x};\bar{b}_\zeta)^{\mathbb{M}}, && \text{for } \eta \leq \zeta\,, \\ \text{and}\quad &\psi(\bar{x};\bar{b}_\eta)^{\mathbb{M}} \cap \psi(\bar{x};\bar{b}_\zeta)^{\mathbb{M}} = \emptyset, && \text{for incomparable } \eta \text{ and } \zeta\,. \end{aligned}$$

Consequently, the set $\{\, \psi(\bar{x};\bar{b}_{\eta i}) \mid i < \omega \,\}$ is 2-inconsistent, for every η. Furthermore, for every $\beta \in \omega^\omega$, we can use compactness and the fact that $\psi(\bar{x};\bar{b}_\eta)^{\mathbb{M}} \neq \emptyset$, for all η, to show that $\{\, \psi(\bar{x};\bar{b}_\eta) \mid \eta \prec \beta \,\}$ is satisfiable. □

Strongly minimal theories

We conclude this section by considering the example of strongly minimal theories. Note that such theories are stable and, hence, simple. We will show that, for strongly minimal theories, the relations $\overset{\mathrm{f}}{\smile}\!\!\!\!|$ and $\overset{\mathrm{acl}}{\smile}\!\!\!\!|$ coincide. One of the inclusions holds in general.

Lemma 2.22. *If $\surd$ is a forking relation, then $\surd \subseteq \overset{\mathrm{cl}_\surd}{\surd}$.*

Proof. Suppose that $A \surd_U B$. To show that $A \overset{\mathrm{cl}_\surd}{\surd}_U B$, consider a set $I \subseteq B$ that is not $\mathrm{cl}_\surd$-independent over $U \cup A$. We have to show that I is not $\mathrm{cl}_\surd$-independent over U. There exists an element $b \in I$ such that $b \in \mathrm{cl}_\surd(UAI_0)$ where $I_0 := I \setminus \{b\}$. Consequently, $b \surd_{UAI_0} B$. By (BMON), $A \surd_U B$ implies $A \surd_{UI_0} B$. Hence, it follows by Lemma F2.2.3 that $Ab \surd_{UI_0} B$. In particular, we have $b \surd_{UI_0} b$ which, by Lemma F2.3.5, implies that $b \in \mathrm{cl}_\surd(UI_0)$. Therefore, I is not $\mathrm{cl}_\surd$-independent over U. □

The converse is given by the following lemma.

Lemma 2.23. *Let T be a simple theory and $\mathbb{S}$ a U-definable strongly minimal set. Then*

$$A \overset{\mathrm{acl}}{\surd}_U B \quad \textit{implies} \quad A \mathop{\downarrow}^{\mathrm{f}}_U B\,, \quad \textit{for all } A, B, U \subseteq \mathbb{S}\,.$$

Proof. Recall that we have shown in Lemma F1.4.3 that $\langle \mathbb{S}, \mathrm{acl} \rangle$ forms a matroid. By (DEF), it is sufficient to prove the claim for finite sets A and B. Hence, suppose that A and B are finite sets with $A \overset{\mathrm{acl}}{\surd}_U B$. We choose bases $I \subseteq A$ and $J \subseteq B$ of, respectively, A over U and B over U, and enumerations $\bar a$ of I and $\bar b$ of J. Then $\bar a \overset{\mathrm{acl}}{\surd}_U \bar b$. Since $\bar b$ is independent over U, it follows that it is also independent over $U \cup \bar a$. Hence, $\bar a \bar b$ is independent over U.

To show that $\bar a \mathop{\downarrow}^{\mathrm{f}}_U \bar b$, let $(\bar b_n)_{n<\omega}$ be an indiscernible sequence over U with $\bar b_0 = \bar b$. Note that the union $\bar b[<\omega]$ is independent over U. We choose a tuple $\bar a' \subseteq \mathbb{S}$ such that $|\bar a'| = |\bar a|$ and $\bar a'$ is independent over $U \cup \bar b[<\omega]$. According to Proposition F1.4.6, we have $\bar a' \equiv_{U\bar b} \bar a$. Since $\bar b[<\omega]$ is independent over $U \cup \bar a'$, it follows by the same proposition that the sequence $(\bar b_n)_{n<\omega}$ is indiscernible over $U \cup \bar a'$. By Lemma 1.3, it follows that $\bar a \overset{\mathrm{d}}{\surd}_U \bar b$. Since T is simple, this implies that $\bar a \mathop{\downarrow}^{\mathrm{f}}_U \bar b$. Hence, we can use Lemma F2.2.14 to show that $\mathrm{acl}(\bar a U) \mathop{\downarrow}^{\mathrm{f}}_U \mathrm{acl}(\bar b U)$. By monotonicity, it follows that $A \mathop{\downarrow}^{\mathrm{f}}_U B$. □

Corollary 2.24. *For a strongly minimal theory T, we have* $\overset{\text{acl}}{\not\bigvee} = \mathop{\smile\!\!\!\!\mid}^{\text{f}} = \mathop{\smile\!\!\!\!\mid}^{\text{d}}$. *In particular, T is simple and* $\mathop{\smile\!\!\!\!\mid}^{\text{f}}$ *is a geometric independence relation.*

Proof. First, note that, according to Lemma F1.4.3, $\langle \mathbb{M}, \text{acl} \rangle$ is a matroid. Hence, it follows from Proposition F2.2.8 that $\overset{\text{acl}}{\not\bigvee}$ is a geometric independence relation. We have seen in Corollary F1.4.14 that a strongly minimal theory T is κ-categorical, for every $\kappa > |T|$. Consequently, it follows by Theorem E6.3.16 that T is stable. Using Corollary 2.19, we see that T is simple. Therefore, the equality $\overset{\text{acl}}{\not\bigvee} = \mathop{\smile\!\!\!\!\mid}^{\text{f}} = \mathop{\smile\!\!\!\!\mid}^{\text{d}}$ follows the two preceding lemmas. $\square$

Exercise 2.2. Prove that, in an arbitrary theory, $\overset{\text{acl}}{\not\bigvee}$ satisfies (INV) and (DEF).

F4. Theories without the independence property

1. Honest definitions

Alternation numbers

We have seen in Proposition E5.4.2 that the independence property can be characterised by counting the number of segments of sets of the form $[\![\varphi(\bar{a}_i)]\!]_{i \in I}$ for an indiscernible sequence $(\bar{a}_i)_{i \in I}$. In this section we will use this characterisation to derive various properties of theories without the independence property. We start by setting up the required combinatorial machinery.

Definition 1.1. Let $\varphi(\bar{x})$ be a formula over $\mathbb{M}$.

(a) The *φ-alternation number* $\mathrm{alt}_\varphi(\alpha)$ of a sequence $\alpha = (\bar{a}_i)_{i \in I}$ is the maximal number $n < \omega$ such that there are indices $\bar{k} \in [I]^{n+1}$ with

$$\mathbb{M} \vDash \varphi(\bar{a}_{k_i}) \leftrightarrow \neg\varphi(\bar{a}_{k_{i+1}}) , \quad \text{for all } i < n .$$

If this maximum does not exist, we set $\mathrm{alt}_\varphi(\alpha) := \infty$.

(b) The *alternation rank* of φ is

$$\mathrm{rk}_{\mathrm{alt}}(\varphi) := \max \{ \mathrm{alt}_\varphi(\alpha) \mid \alpha \text{ an indiscernible sequence in } \mathbb{M} \} .$$

If this maximum does not exist, we set $\mathrm{rk}_{\mathrm{alt}}(\varphi) := \infty$.

(c) A sequence $\alpha = (\bar{a}_i)_{i \in I}$ is *maximally φ-alternating over U* if it is indiscernible over U and

$$\mathrm{alt}_\varphi(\alpha) = \mathrm{alt}_\varphi(\alpha\beta) < \infty ,$$

for every extension $\alpha\beta$ of α that is still indiscernible over U.

Using these notions, we can characterise the independence property as follows.

Proposition 1.2. *Let $\varphi(\bar{x};\bar{y})$ be a formula without parameters and let $U \subseteq \mathbb{M}$. The following statements are equivalent.*

(1) *$\varphi(\bar{x};\bar{y})$ does not have the independence property.*

(2) $\mathrm{rk}_{\mathrm{alt}}(\varphi(\bar{x};\bar{c})) < \infty$, *for all $\bar{c} \subseteq \mathbb{M}$.*

(3) *There exists some number $n < \omega$ such that*

$$\mathrm{rk}_{\mathrm{alt}}(\varphi(\bar{x};\bar{c})) \leq n\,, \quad \text{for all } \bar{c} \subseteq \mathbb{M}\,.$$

(4) $\mathrm{alt}_{\varphi(\bar{x};\bar{c})}(\alpha) < \infty$, *for every indiscernible sequence α over U and every tuple $\bar{c} \subseteq \mathbb{M}$.*

(5) *Let $\bar{c} \subseteq \mathbb{M}$. Every indiscernible sequence α over U has an extension $\alpha\beta$ that is maximally $\varphi(\bar{x};\bar{c})$-alternating over U.*

Proof. (3) ⇒ (2) is trivial.

(2) ⇒ (5) Suppose that $\mathrm{rk}_{\mathrm{alt}}(\varphi(\bar{x};\bar{c})) < \infty$ and let α be an indiscernible sequence over U. We construct a maximally $\varphi(\bar{x};\bar{c})$-alternating extension of α by induction on the difference

$$\mathrm{rk}_{\mathrm{alt}}(\varphi(\bar{x};\bar{c})) - \mathrm{alt}_{\varphi(\bar{x};\bar{c})}(\alpha)\,.$$

If α is already maximally $\varphi(\bar{x};\bar{c})$-alternating, there is nothing to do. Hence, suppose otherwise. Then we can find some extension $\alpha\beta$ that is indiscernible over U such that $\mathrm{alt}_{\varphi(\bar{x};\bar{c})}(\alpha\beta) > \mathrm{alt}_{\varphi(\bar{x};\bar{c})}(\alpha)$. By inductive hypothesis, this sequence has an extension $\alpha\beta\gamma$ that is maximally $\varphi(\bar{x};\bar{c})$-alternating over U.

(5) ⇒ (4) Let $\alpha\beta$ be a maximally $\varphi(\bar{x};\bar{c})$-alternating extension of α over U. Then $\mathrm{alt}_{\varphi(\bar{x};\bar{c})}(\alpha) \leq \mathrm{alt}_{\varphi(\bar{x};\bar{c})}(\alpha\beta) < \infty$.

(4) ⇒ (1) Suppose that $\varphi(\bar{x};\bar{y})$ has the independence property. By Proposition E5.4.2, there exists an indiscernible sequence $\alpha = (\bar{a}_n)_{n<\omega}$

and a tuple $\bar{c}$ such that

$$\mathbb{M} \vDash \varphi(\bar{a}_n; \bar{c}) \quad \text{iff} \quad n \text{ is even.}$$

Hence, $\mathrm{alt}_{\varphi(\bar{x};\bar{c})}(\alpha) = \infty$.

(1) ⇒ (3) Suppose that, for every number $n < \omega$, there exists some tuple $\bar{c} \subseteq \mathbb{M}$ such that $\mathrm{rk}_{\mathrm{alt}}(\varphi(\bar{x}; \bar{c})) > n$. We claim that φ has the independence property. Let Ψ be a set of formulae stating that the sequence $(\bar{x}_i)_{i<\omega}$ is indiscernible and set

$$\Phi := \Psi \cup \{ \varphi(\bar{x}_{2i}; \bar{y}) \mid i < \omega \} \cup \{ \neg\varphi(\bar{x}_{2i+1}; \bar{y}) \mid i < \omega \} .$$

We will show that Φ is satisfiable. Then there exists an indiscernible sequence $(\bar{a}_i)_{i<\omega}$ and a tuple $\bar{b}$ such that

$$\mathbb{M} \vDash \varphi(\bar{a}_i; \bar{b}) \quad \text{iff} \quad i \text{ is even,}$$

and it follows by Proposition E5.4.2 that φ has the independence property.

Thus, let $\Phi_0 \subseteq \Phi$ be finite. Then there exists a number $n < \omega$ such that all variables occurring in Φ_0 are among $\bar{x}_0, \dots, \bar{x}_{2n-1}$. By assumption, we can find a tuple $\bar{c}$ and an indiscernible sequence $\alpha = (\bar{a}_i)_{i\in I}$ such that

$$\mathrm{alt}_{\varphi(\bar{x};\bar{c})}(\alpha) \geq 2n .$$

We choose indices $\bar{m} \in [I]^{2n+1}$ such that

$$\mathbb{M} \vDash \varphi(\bar{a}_{m_i}; \bar{c}) \leftrightarrow \neg\varphi(\bar{a}_{m_{i+1}}; \bar{c}) , \quad \text{for all } i < 2n .$$

Depending on whether or not $\mathbb{M} \vDash \varphi(\bar{a}_{m_0}; \bar{c})$, it follows that either the sequence $(\bar{a}_{m_i})_{0\leq i<2n}$ or the sequence $(\bar{a}_{m_i})_{1\leq i<2n+1}$ satisfies Φ_0 together with the tuple $\bar{c}$. □

Below we will frequently make use of the following consequence of this characterisation.

Corollary 1.3. *Let T be a theory without the independence property and let Δ be a finite set of formulae over $\mathbb{M}$. Every indiscernible sequence α over U has an extension $\alpha\beta$ that is maximally φ-alternating over U, for all $\varphi \in \Delta$.*

Proof. Let α be indiscernible over U. We construct the desired extension by induction on $|\Delta|$. If $\Delta = \emptyset$, we can take the sequence α itself. Hence, we may assume that there is some formula $\varphi \in \Delta$. Suppose that $\varphi(\bar{x}) = \varphi_0(\bar{x};\bar{c})$ where $\bar{c} \subseteq \mathbb{M}$ and $\varphi_0(\bar{x};\bar{y})$ is a formula without parameters. As $\varphi_0(\bar{x};\bar{y})$ does not have the independence property, it follows by Proposition 1.2 that α has a maximally φ-alternating extension $\alpha\beta$. By inductive hypothesis, this sequence has an extension $\alpha\beta\gamma$ that is maximally ψ-alternating, for every $\psi \in \Delta \setminus \{\varphi\}$. Since $\mathrm{alt}_\varphi(\alpha\beta) \leq \mathrm{alt}_\varphi(\alpha\beta\gamma)$, this extension is also maximally φ-alternating. Hence, $\alpha\beta\gamma$ is the desired extension of α. □

Honest definitions

Stable theories have the property that every set $\mathbb{A} \subseteq \mathbb{M}$ is self-contained as far as definable relations are concerned, that is, all parameter-definable relations $\mathbb{R} \subseteq \mathbb{A}^{\bar{s}}$ are definable with parameters from $\mathbb{A}$ itself. In this section, we will prove that theories without the independence property have a similar, but weaker property: the parameters are not necessarily in the set $\mathbb{A}$, but in some elementary extension. We start by taking a look at the stable case.

Definition 1.4. A set $\mathbb{A} \subseteq \mathbb{M}$ is *stably embedded* if, for every parameter-definable relation $\mathbb{R} \subseteq \mathbb{M}^{\bar{s}}$, there is a formula $\varphi(\bar{x})$ over $\mathbb{A}$ such that

$$\mathbb{R} \cap \mathbb{A}^{\bar{s}} = \varphi^{\mathbb{M}} \cap \mathbb{A}^{\bar{s}} .$$

Proposition 1.5. *In a stable theory, every set $\mathbb{A} \subseteq \mathbb{M}$ is stably embedded.*

Proof. Let $\psi(\bar{x};\bar{c})$ be a formula with parameters $\bar{c} \subseteq \mathbb{M}$. As T is stable, it follows by Theorem C3.5.17 that the type $\mathrm{tp}(\bar{c}/\mathbb{A})$ is definable over $\mathbb{A}$. Consequently, there exists a formula $\delta_\psi(\bar{y})$ over $\mathbb{A}$ such that

$$\mathbb{M} \vDash \delta_\psi(\bar{a}) \quad \text{iff} \quad \mathbb{M} \vDash \psi(\bar{a};\bar{c}) .$$

This implies that $\psi(\bar{x};\bar{c})^{\mathbb{M}} \cap \mathbb{A}^{\bar{s}} = \delta_\psi(\bar{x})^{\mathbb{M}} \cap \mathbb{A}^{\bar{s}}$. □

For theories with the independence property, we need to consider elementary extensions of the given structure to find the desired parameters. Alternatively, we can also use the following finitary version of stable embeddedness.

Definition 1.6. An *honest definition* of a relation $\mathbb{R} \subseteq \mathbb{M}^{\bar{s}}$ over a set U is a formula $\varphi(\bar{x};\bar{y})$ without parameters such that, for every finite $U_0 \subseteq U$, there is some tuple $\bar{c} \subseteq U$ with

$$\mathbb{R} \cap U_0^{\bar{s}} \subseteq \varphi(\bar{x};\bar{c})^{\mathbb{M}} \cap U^{\bar{s}} \subseteq \mathbb{R} \cap U^{\bar{s}}.$$

Example. The set $\mathbb{Q}$ of rationals is not stably embedded in $\langle \mathbb{R}, \leq \rangle$. For instance, for the parameter-definable relation $(0, \sqrt{2})$, there is no formula $\varphi(x)$ over $\mathbb{Q}$ with

$$\varphi^{\mathbb{R}} \cap \mathbb{Q} = (0, \sqrt{2}) \cap \mathbb{Q}\,.$$

But $(0, \sqrt{2})$ *does* have an honest definition over $\mathbb{Q}$. For every finite subset $A \subseteq (0, \sqrt{2})$, we have

$$(0, \sqrt{2}) \cap A \subseteq \varphi(x; a, b)^{\mathbb{R}} \cap \mathbb{Q} \subseteq (0, \sqrt{2}) \cap \mathbb{Q}\,,$$

where $\varphi(x; y, z) :\!= y \leq x \wedge x \leq z$ and a and b are, respectively, the minimal and the maximal element of A.

Below we will prove that these two weaker version of stable embeddedness are equivalent and that they hold in theories without the independence property. The key argument is contained in the following lemma.

Lemma 1.7. *Let $\kappa > |T|$ be a cardinal and let $\langle \mathfrak{M}, C \rangle \preceq \langle \mathfrak{M}_+, C_+ \rangle$ be structures where the former one has size $|M| < \kappa$ and the latter one is κ-saturated. For all sets $A, B \subseteq M_+$ of size $|A|, |B| < \kappa$ with $A \overset{\mathrm{u}}{\surd}_C B$, there exists some $A' \subseteq C_+$ such that $A' \equiv_B A$.*

Proof. Let $\bar{a} = (a_i)_{i<\lambda}$ be an enumeration of A and let $\mathbb{C} \subseteq \mathbb{M}$ be a set such that $\langle \mathbb{M}, \mathbb{C} \rangle \succeq \langle \mathfrak{M}_+, C_+ \rangle$. Set

$$\Phi(\bar{x}) :\!= \mathrm{Th}(\langle \mathbb{M}, \mathbb{C} \rangle) \cup \mathrm{tp}(\bar{a}/B) \cup \{\, Px_i \mid i < \lambda \,\}\,,$$

where the type $\mathrm{tp}(\bar a/B)$ is taken with respect to the structure $\mathbb{M}$ and P is the predicate symbol of $\langle \mathbb{M}, \mathbb{C} \rangle$ corresponding to the set $\mathbb{C}$. If $\Phi(\bar x)$ is satisfiable, it follows by κ-saturation of $\langle \mathfrak{M}_+, C_+ \rangle$ that there is some tuple $\bar a' \subseteq M_+$ with $\langle \mathfrak{M}_+, C_+ \rangle \vDash \Phi(\bar a')$. By definition of Φ, this implies that $\bar a' \subseteq C_+$ and $\bar a' \equiv_B \bar a$. Hence, it remains to prove that Φ is satisfiable.

Let $\Phi_0 \subseteq \Phi$ be finite. Then

$$\Phi_0(\bar x) \equiv \psi \wedge \varphi(\bar x) \wedge \bigwedge_{i \in I} P x_i \,,$$

for suitable formulae $\psi \in \mathrm{Th}(\langle \mathbb{M}, \mathbb{C} \rangle)$, $\varphi(\bar x) \in \mathrm{tp}(\bar a/B)$, and some finite set $I \subseteq \lambda$. Since $\bar a \overset{\mathrm{u}}{\surd}_C B$, we can find some tuple $\bar a' \subseteq C \subseteq C_+$ with $\mathbb{M} \vDash \varphi(\bar a')$. Consequently,

$$\langle \mathfrak{M}_+, C_+ \rangle \vDash \psi \wedge \varphi(\bar a') \wedge \bigwedge_{i \in I} P a'_i \,,$$

and $\bar a'$ satisfies $\Phi_0(\bar x)$. □

A second technical ingredient we need in the proof below is the notion of a type *generating* a sequence.

Definition 1.8. Let $\mathfrak{p}$ be a type. A sequence $(\bar a_i)_{i \in I}$ is *generated by* $\mathfrak{p}$ *over* U if $\bar a_i$ realises $\mathfrak{p} \upharpoonright U \bar a[<i]$, for all $i \in I$.

Exercise 1.1. Prove that, for every type $\mathfrak{p} \in S^{\bar s}(\mathbb{M})$ and every small index set I, there is some sequence $(\bar a_i)_{i \in I}$ generated by $\mathfrak{p}$.

When using a suitable type, the generated sequence is automatically a Morley sequence.

Lemma 1.9. *Let $\surd$ be a preforking relation and $\mathfrak{p}$ a global type that is $\surd$-free over U. Every sequence generated by $\mathfrak{p}$ over a set $U \cup C$ is a $\surd$-Morley sequence for $\mathfrak{p} \upharpoonright UC$ over U.*

The existence of honest definitions turns out to being equivalent to not having the independence property.

Theorem 1.10. *Let $\varphi(\bar x)$ be a formula over $\mathbb{M}$ and let $\bar s$ be the sorts of $\bar x$. The following statements are equivalent:*

(1) $\mathrm{rk}_{\mathrm{alt}}(\varphi) < \infty$.

(2) *For every set $C \subseteq \mathbb{M}$, there is a honest definition of $\varphi^{\mathbb{M}}$ over C.*

(3) *For every model $\mathfrak{M}$ containing the parameters of φ, every set $C \subseteq M$ of parameters, and every $(|T| \oplus |M|)^+$-saturated elementary extension $\langle \mathfrak{M}_+, C_+ \rangle \succeq \langle \mathfrak{M}, C \rangle$, there exists a formula $\varphi_+(\bar{x})$ over C_+ such that*

$$\varphi(\bar{x})^{\mathbb{M}} \cap C^{\bar{s}} \subseteq \varphi_+(\bar{x})^{\mathbb{M}} \cap C_+^{\bar{s}} \subseteq \varphi(\bar{x})^{\mathbb{M}} \cap C_+^{\bar{s}}.$$

Proof. (3) ⇒ (2) Fix a model $\mathfrak{M}$ containing the parameters of φ, a set $C \subseteq M$, and a $(|T| \oplus |M|)^+$-saturated elementary extension $\langle \mathfrak{M}_+, C_+ \rangle \succeq \langle \mathfrak{M}, C \rangle$. By (3), there is some formula $\varphi_+(\bar{x}; \bar{c})$ with parameters $\bar{c} \subseteq C_+$ such that

$$\varphi(\bar{x})^{\mathbb{M}} \cap C^{\bar{s}} \subseteq \varphi_+(\bar{x}; \bar{c})^{\mathbb{M}} \cap C_+^{\bar{s}} \subseteq \varphi(\bar{x})^{\mathbb{M}} \cap C_+^{\bar{s}}.$$

We claim that $\varphi_+(\bar{x}; \bar{y})$ is a honest definition of $\varphi^{\mathbb{M}}$ over C. Let $C_0 \subseteq C$ be finite. Then

$$\langle \mathfrak{M}_+, C_+ \rangle \vDash \bigwedge_{\bar{a} \in C_0^{\bar{s}}} [\varphi_+(\bar{a}; \bar{c}) \leftrightarrow \varphi(\bar{a})] \wedge (\forall \bar{x}. \textstyle\bigwedge_i P x_i)[\varphi_+(\bar{x}; \bar{c}) \rightarrow \varphi(\bar{x})].$$

Consequently,

$$\langle \mathfrak{M}, C \rangle \vDash (\exists \bar{y}. \textstyle\bigwedge_i P y_i)\Big[\displaystyle\bigwedge_{\bar{a} \in C_0^{\bar{s}}} [\varphi_+(\bar{a}; \bar{y}) \leftrightarrow \varphi(\bar{a})] \wedge (\forall \bar{x}. \textstyle\bigwedge_i P x_i)[\varphi_+(\bar{x}; \bar{y}) \rightarrow \varphi(\bar{x})]\Big],$$

and there is some tuple $\bar{c}' \subseteq C$ such that

$$\varphi^{\mathbb{M}} \cap C_0^{\bar{s}} \subseteq \varphi_+(\bar{x}; \bar{c}')^{\mathbb{M}} \cap C^{\bar{s}} \subseteq \varphi^{\mathbb{M}} \cap C^{\bar{s}}.$$

(2) ⇒ (1) For a contradiction, suppose that $\mathrm{rk}_{\mathrm{alt}}(\varphi(\bar{x})) = \infty$ but $\varphi^{\mathbb{M}}$ has honest definitions over all sets $C \subseteq \mathbb{M}$. By compactness there

exists an indiscernible sequence $\alpha = (\bar{a}_n)_{n<\omega}$ such that $\mathrm{alt}_\varphi(\alpha) = \infty$. Omitting some elements of α we may assume that

$$\mathbb{M} \vDash \varphi(\bar{a}_n) \quad \text{iff} \quad n \text{ is even.}$$

Let $\psi(\bar{x}; \bar{y})$ be an honest definition of $\varphi^{\mathbb{M}}$ over the set $C := \bar{a}[<\omega]$ and let $C_0 := \bar{a}[<2k+2]$ where $k := |\bar{y}|$. By assumption, there is some tuple $\bar{c} \subseteq C$ such that

$$\varphi^{\mathbb{M}} \cap C_0^{\bar{s}} \subseteq \psi(\bar{x}; \bar{c})^{\mathbb{M}} \cap C^{\bar{s}} \subseteq \varphi^{\mathbb{M}} \cap C^{\bar{s}}.$$

Fix some tuple $\bar{\jmath} \in [\omega]^k$ such that $\bar{c} \subseteq \bar{a}[\bar{\jmath}]$. Then there is some index $i < 2k+1$ such that

$$\mathrm{ord}(i\bar{\jmath}) = \mathrm{ord}((i+1)\bar{\jmath}).$$

Consequently,

$$\mathbb{M} \vDash \psi(\bar{a}_i; \bar{c}) \leftrightarrow \psi(\bar{a}_{i+1}; \bar{c}).$$

If i is even, then

$$\psi(\bar{x}; \bar{c})^{\mathbb{M}} \cap C^{\bar{s}} \subseteq \varphi^{\mathbb{M}} \cap C^{\bar{s}} \qquad \text{implies} \quad \bar{a}_{i+1} \notin \psi(\bar{x}; \bar{c})^{\mathbb{M}},$$

while $\quad \varphi^{\mathbb{M}} \cap C_0^{\bar{s}} \subseteq \psi(\bar{x}; \bar{c})^{\mathbb{M}} \cap C^{\bar{s}} \quad \text{implies} \quad \bar{a}_i \in \psi(\bar{x}; \bar{c})^{\mathbb{M}}.$

A contradiction. In the case where i is odd, we can show in the same way that $\bar{a}_i \notin \psi(\bar{x}; \bar{c})^{\mathbb{M}}$ and $\bar{a}_{i+1} \in \psi(\bar{x}; \bar{c})^{\mathbb{M}}$.

(1) ⇒ (3) Let $F \subseteq S^{\bar{s}}(M_+)$ be the set of all types over M_+ that are finitely satisfiable in C and let $F_\varphi := F \cap \langle\varphi\rangle$ be the subset of those types containing φ. As $\mathrm{rk}_{\mathrm{alt}}(\varphi) < \infty$, we can choose, for every type $\mathfrak{p} \in F$, a sequence $\alpha_\mathfrak{p} \subseteq C_+$ that is generated by $\mathfrak{p}$ over C and such that $\mathrm{alt}_{\varphi(\bar{x})}(\alpha_\mathfrak{p})$ is maximal (among all such sequences in C_+).

Let $\bar{a}' \subseteq C_+$ be a tuple realising $\mathfrak{p} \upharpoonright C\alpha_\mathfrak{p}$, for some $\mathfrak{p} \in F$. We claim that

$$\mathbb{M} \vDash \varphi(\bar{a}') \quad \text{iff} \quad \varphi(\bar{x}) \in \mathfrak{p}.$$

By Lemma 1.7, there is some $\bar{a}'' \in C_+^{\bar{s}}$ realising $\mathfrak{p} \upharpoonright M\alpha_\mathfrak{p}\bar{a}'$. Then the sequence $\alpha_\mathfrak{p}\bar{a}'\bar{a}''$ is generated by $\mathfrak{p}$ over C and our choice of $\alpha_\mathfrak{p}$ implies that

$$\mathrm{alt}_\varphi(\alpha_\mathfrak{p}\bar{a}'\bar{a}'') = \mathrm{alt}_\varphi(\alpha_\mathfrak{p})\,.$$

As φ is over M, it follows by choice of $\bar{a}''$ that

$$\mathbb{M} \vDash \varphi(\bar{a}') \quad \text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}'') \quad \text{iff} \quad \varphi(\bar{x}) \in \mathfrak{p}\,,$$

as desired.

For types $\mathfrak{p} \in F_\varphi$, the claim we have just proved implies that

$$\mathrm{Th}(\langle \mathbb{M}_M, \mathbb{C}\rangle) \cup \mathfrak{p} \upharpoonright C\alpha_\mathfrak{p} \cup \{Px_0, \dots, Px_{n-1}\} \vDash \varphi(\bar{x})\,,$$

where $\bar{x} = x_0 \dots x_{n-1}$, $\mathbb{C}$ is a set such that $\langle \mathbb{M}, \mathbb{C}\rangle \succeq \langle \mathfrak{M}_+, C_+\rangle$, and P is the predicate symbol corresponding to $\mathbb{C}$. Therefore, we can use compactness to find a formula $\vartheta_\mathfrak{p}(\bar{x}) \in \mathfrak{p} \upharpoonright C\alpha_\mathfrak{p}$ such that

$$\mathrm{Th}(\langle \mathbb{M}_M, \mathbb{C}\rangle) \cup \{\vartheta_\mathfrak{p}(\bar{x}), Px_0, \dots, Px_{n-1}\} \vDash \varphi(\bar{x})\,.$$

Note that $\vartheta_\mathfrak{p} \in \mathfrak{p}$ implies $\mathfrak{p} \in \langle \vartheta_\mathfrak{p}\rangle$. Hence,

$$F_\varphi \subseteq \bigcup_{\mathfrak{p} \in F_\varphi} \langle \vartheta_\mathfrak{p}\rangle\,.$$

By Lemma F2.3.7, F is a closed set. Hence, so is $F_\varphi = F \cap \langle \varphi(\bar{x})\rangle$. As closed sets in Hausdorff spaces are compact, it follows that there exists a finite subset $F_0 \subseteq F_\varphi$ such that

$$F_\varphi \subseteq \bigcup_{\mathfrak{p} \in F_0} \langle \vartheta_\mathfrak{p}\rangle\,.$$

We claim that

$$\varphi_+(\bar{x}) := \bigvee_{\mathfrak{p} \in F_0} \vartheta_\mathfrak{p}$$

is the desired formula.

Consider a tuple $\bar{a} \in C^{\bar{s}}$ with $\mathbb{M} \vDash \varphi(\bar{a})$. Then $\mathfrak{p} := \mathrm{tp}(\bar{a}/M_+)$ is trivially finitely satisfiable in C. Hence, $\mathfrak{p} \in F_\varphi$ and we have $\vartheta_\mathfrak{q} \in \mathfrak{p}$, for some $\mathfrak{q} \in F_0$. This implies that $\varphi_+(\bar{x}) \in \mathfrak{p}$. Consequently,

$$\varphi(\bar{x})^{\mathbb{M}} \cap C^{\bar{s}} \subseteq \varphi_+(\bar{x})^{\mathbb{M}} \cap C^{\bar{s}} \subseteq \varphi_+(\bar{x})^{\mathbb{M}} \cap C_+^{\bar{s}}.$$

For the second inclusion, let $\bar{a} \in C_+^{\bar{s}}$ be a tuple with $\mathbb{M} \vDash \vartheta_\mathfrak{p}(\bar{a})$, for some $\mathfrak{p} \in F_0$. Then we have $\mathbb{M} \vDash \varphi(\bar{a})$, by choice of $\vartheta_\mathfrak{p}$. Hence,

$$\varphi_+(\bar{x})^{\mathbb{M}} \cap C_+^{\bar{s}} \subseteq \varphi(\bar{x})^{\mathbb{M}} \cap C_+^{\bar{s}}. \qquad \square$$

As a corollary, we obtain the following weak variant of stable embeddedness for theories without the independence property.

Corollary 1.11. *For every model $\mathfrak{M}$, every set $C \subseteq M$, and every formula $\varphi(\bar{x})$ over M with $\mathrm{rk}_{\mathrm{alt}}(\varphi) < \infty$, there exists an elementary extension $\langle \mathfrak{M}_+, C_+ \rangle \succeq \langle \mathfrak{M}, C \rangle$ and a formula $\varphi_+(\bar{x})$ over C_+ such that*

$$\varphi(\bar{x})^{\mathbb{M}} \cap C^{\bar{s}} = \varphi_+(\bar{x})^{\mathbb{M}} \cap C^{\bar{s}}\,.$$

Another convenient consequence of Theorem 1.10 is contained in the proposition below. Again we isolate the main argument in a lemma.

Lemma 1.12. *Let T be a theory without the independence property and κ an infinite cardinal. Let $\mathfrak{M}$ be a model of T of size $|M| < \kappa$, $B \subseteq M$ a set, and $\langle \mathfrak{M}_+, B_+ \rangle \succeq \langle \mathfrak{M}, B \rangle$ a κ-saturated elementary extension. For every set $C \subseteq M$, there exists a set $U \subseteq B_+$ of size $|U| \leq |T| \oplus |C|$ such that*

$$\bar{b} \equiv_U \bar{b}' \quad \text{implies} \quad \bar{b} \equiv_C \bar{b}'\,, \quad \text{for all } \bar{b}, \bar{b}' \subseteq B\,.$$

Proof. For every formula $\varphi(\bar{x})$ over C, we use Theorem 1.10 to find a formula φ_+ over B_+ such that

$$\varphi(\bar{x})^{\mathbb{M}} \cap B^{\bar{s}} \subseteq \varphi_+(\bar{x})^{\mathbb{M}} \cap (B_+)^{\bar{s}} \subseteq \varphi(\bar{x})^{\mathbb{M}} \cap (B_+)^{\bar{s}}.$$

Let $U \subseteq B_+$ be a set of size $|U| \leq |T| \oplus |C|$ containing the parameters of each of these formulae φ_+.

To show that U has the desired properties, consider tuples $\bar{b}, \bar{b}' \subseteq B$ with $\bar{b} \equiv_U \bar{b}'$. For every formula $\varphi(\bar{x})$ over C and every finite set I of indices, it follows that

$$\begin{aligned} \mathbb{M} \vDash \varphi(\bar{b}|_I) \quad &\text{iff} \quad \mathbb{M} \vDash \varphi_+(\bar{b}|_I) \\ &\text{iff} \quad \mathbb{M} \vDash \varphi_+(\bar{b}|'_I) \quad \text{iff} \quad \mathbb{M} \vDash \varphi(\bar{b}|'_I)\,. \end{aligned}$$

Consequently, $\bar{b} \equiv_C \bar{b}'$. ☐

Proposition 1.13. *Let T be a theory without the independence property, $\mathfrak{M}$ a model of T, and $B \subseteq M$. Then there exists an elementary extension $\langle \mathfrak{M}_+, B_+ \rangle \succeq \langle \mathfrak{M}, B \rangle$ such that, for every set $A \subseteq M$, there exists a set $U \subseteq B_+$ of size $|U| \leq |T| \oplus |A|$ with $A \sqrt[s]{}_U B_+$.*

Proof. We iterate the preceding lemma. Let $\langle \mathfrak{M}_+, B_+ \rangle$ be the union of an elementary chain $\langle \mathfrak{M}_n, B_n \rangle_{n<\omega}$ starting with $\langle \mathfrak{M}_0, B_0 \rangle := \langle \mathfrak{M}, B \rangle$ where each $\langle \mathfrak{M}_{n+1}, B_{n+1} \rangle \succeq \langle \mathfrak{M}_n, B_n \rangle$ is $(|T| \oplus |M_n|)^+$-saturated. We inductively construct a sequence $(U_n)_{n<\omega}$ of sets $U_n \subseteq B_{n+1}$ of size $|U_n| \leq |T| \oplus |A|$ as follows. Suppose that we have already defined $U_0, \dots, U_{n-1} \subseteq B_n \subseteq M_n$. By Lemma 1.12, there exists some set $U_n \subseteq B_{n+1}$ of size

$$|U_n| \leq |T| \oplus |A| \oplus |U_0| \oplus \dots \oplus |U_{n-1}| = |T| \oplus |A|$$

such that

$$\bar{b} \equiv_{U_n} \bar{b}' \quad \text{implies} \quad \bar{b} \equiv_{A \cup U_0 \cup \dots \cup U_{n-1}} \bar{b}'\,, \quad \text{for all } \bar{b}, \bar{b}' \subseteq B_n\,.$$

Set $U := \bigcup_{n<\omega} U_n$ and let $\bar{b}, \bar{b}' \subseteq B_+$ be finite tuples with $\bar{b} \equiv_U \bar{b}'$. Then $|U| \leq |T| \oplus |A|$ and there is some $k < \omega$ such that $\bar{b}, \bar{b}' \subseteq B_k$. It follows that

$$\bar{b} \equiv_{A \cup U_0 \cup \dots \cup U_{n-1}} \bar{b}'\,, \quad \text{for all } n \geq k\,.$$

Consequently, $\bar{b} \equiv_{AU} \bar{b}'$, as desired.

For infinite tuples $\bar{b}, \bar{b}' \subseteq B_+$ with $\bar{b} \equiv_U \bar{b}'$, it therefore follows that

$$\bar{b}|_I \equiv_U \bar{b}'|_I \quad \text{implies} \quad \bar{b}|_I \equiv_{AU} \bar{b}'|_I\,, \quad \text{for all finite sets } I\,.$$

Consequently, $\bar{b} \equiv_{AU} \bar{b}'$. ☐

Convex equivalence relations

As an application we study the structure of indiscernible sequences in theories without the independence property.

Definition 1.14. Let $\mathfrak{I} = \langle I, \leq \rangle$ be a linear order and $\sim$ an equivalence relation on I.

(a) $\sim$ is *convex* if

$$i \sim j \quad \text{implies} \quad i \sim k \quad \text{for all } i \leq k \leq j\,.$$

(b) $\sim$ is *finite* if it has only finitely many classes.

(c) The *intersection number* $\mathrm{in}(\sim)$ of a convex equivalence relation $\sim$ is the least cardinal κ such that $\sim$ can be written as an intersection of κ finite convex equivalence relations.

(d) For tuples $\bar{\imath}, \bar{\jmath} \in I^{<\omega}$, we set

$$\bar{\imath} \sim \bar{\jmath} \quad :\text{iff} \quad \mathrm{ord}(\bar{\imath}) = \mathrm{ord}(\bar{\jmath}) \quad \text{and} \quad i_s \sim j_s \quad \text{for all } s\,.$$

(e) For a subset $C \subseteq I$ and tuples $\bar{\imath}, \bar{\jmath} \subseteq I$, we define

$$\bar{\imath} \equiv^{\circ}_{C} \bar{\jmath} \quad :\text{iff} \quad \mathfrak{I}, \bar{\imath}\bar{c} \equiv^{\circ} \mathfrak{I}, \bar{\jmath}\bar{c} \quad \text{where } \bar{c} \text{ is an enumeration of } C\,.$$

Let us note that the relation $\equiv^{\circ}_{C}$ is convex and that its definition for tuples is consistent with the notation introduced in (d) above.

Lemma 1.15. *$\equiv^{\circ}_{C}$ is a convex equivalence relation with* $\mathrm{in}(\equiv^{\circ}_{C}) \leq |C|$ *that satisfies*

$$\bar{\imath} \equiv^{\circ}_{C} \bar{\jmath} \quad :\text{iff} \quad \mathrm{ord}(\bar{\imath}) = \mathrm{ord}(\bar{\jmath}) \quad \textit{and} \quad i_s \equiv^{\circ}_{C} j_s \quad \textit{for all } s\,.$$

Proof. For the bound on the intersection number, note that

$$\equiv^{\circ}_{C} = \bigcap_{c \in C} \equiv^{\circ}_{\{c\}}\,.$$

The other claims are straightforward. □

The statement of the preceding lemma has a weak converse: every convex equivalence relation can be obtained as a coarsening of a relation of the form $\equiv^{\circ}_{C}$.

Lemma 1.16. *Let $\sim$ be a convex equivalence relation on a linear order I and J a complete linear order containing I. Then there exists a set $C \subseteq J$ of size $|C| \leq \mathrm{in}(\sim) \oplus \aleph_0$ such that the restriction of $\equiv^{\circ}_{C}$ to I refines $\sim$.*

Proof. Set $\kappa := \mathrm{in}(\sim) \oplus \aleph_0$ and let F be a set of finite convex equivalence relations of size $|F| \leq \kappa$ such that $\sim = \bigcap F$. We set

$$\begin{aligned} C := {} & \{ \inf E \mid E \text{ an } \approx\text{-class for some } \approx \in F \} \\ & \cup \{ \sup E \mid E \text{ an } \approx\text{-class for some } \approx \in F \} , \end{aligned}$$

where we take the infima and suprema in the ordering J. Then $|C| \leq |F| \otimes \aleph_0 \leq \kappa$ and the restriction of $\equiv^{\circ}_{C}$ to I refines $\sim$. □

Theorem 1.17. *Let T be a theory without the independence property and $\alpha = (\bar{a}^i)_{i \in I}$ an indiscernible sequence over U. For every set $C \subseteq \mathbb{M}$, there exist a linear order $J \supseteq I$, an indiscernible sequence $\alpha_+ = (\bar{a}^j)_{j \in J}$ over U with $\alpha_+ \restriction I = \alpha$, and a subset $K \subseteq J$ of size $|K| \leq |T| \oplus |C|$ such that*

$$\bar{\imath} \equiv^{\circ}_{K} \bar{\jmath} \quad \textit{implies} \quad \bar{a}[\bar{\imath}] \equiv_{UC} \bar{a}[\bar{\jmath}] , \quad \textit{for all } \bar{\imath}, \bar{\jmath} \in [J]^{<\omega} .$$

Proof. Let $\mathfrak{M}$ be a model containing $U \cup C \cup \alpha$. Suppose that the sequence α consists of γ-tuples $\bar{a}^i = (a^i_k)_{k<\gamma}$ and set

$$\begin{aligned} P &:= U \cup \{ a^i_k \mid i \in I ,\ k < \gamma \} , \\ E &:= \{ \langle a^i_k, a^i_l \rangle \mid i \in I ,\ k, l < \gamma \} , \\ F &:= \{ \langle a^i_k, a^j_k \rangle \mid i, j \in I ,\ k < \gamma \} , \\ R &:= \{ \langle a^i_k, a^j_l \rangle \mid i < j \text{ in } I ,\ k, l < \gamma \} . \end{aligned}$$

Fix an $|M|^+$-saturated elementary extension

$$\langle \mathfrak{M}_+, P_+, U_+, E_+, F_+, R_+ \rangle \succeq \langle \mathfrak{M}, P, U, E, F, R \rangle .$$

Using the relations E_+, F_+, and R_+, we see that there are a linear order $I_+ \supseteq I$, an ordinal $\gamma_+ \geq \gamma$, and a family $(b^i_k)_{i \in I_+, k < \gamma_+}$ of elements such that

- $P_+ = U_+ \cup \{ b^i_k \mid i \in I_+,\ k < \gamma_+ \}$,
- $b^i_k = a^i_k\,, \quad$ for $i \in I$ and $k < \gamma\,,$
- the sequence $(\bar b^i)_{i \in I_+}$ consisting of $\bar b^i := (b^i_k)_{k < \gamma_+}$, $i \in I_+$, is indiscernible over U_+.

By Lemma 1.12, we can find a set $W \subseteq P_+$ of size $|W| \leq |T| \oplus |C|$ such that

$$\bar a \equiv_W \bar a' \quad \text{implies} \quad \bar a \equiv_C \bar a'\,, \quad \text{for all } \bar a, \bar a' \subseteq P\,.$$

We claim that the sequence $\alpha' := (\bar b^i|_\gamma)_{i \in I_+}$ and the set

$$K := \{ i \in I_+ \mid \bar b^i \cap W \neq \varnothing \}$$

have the desired properties. Consider tuples $\bar\imath, \bar\jmath \in [I_+]^{<\omega}$ with $\bar\imath \equiv^{\circ}_K \bar\jmath$ and let $\bar k$ be an enumeration of K. Since $(\bar b^i)_{i \in I_+}$ is indiscernible over U, it follows that

$$\bar\imath \equiv^{\circ}_K \bar\jmath \quad \Rightarrow \quad \mathfrak{I}_+, \bar\imath\bar k \equiv^{\circ} \mathfrak{I}_+, \bar\jmath\bar k \quad \Rightarrow \quad \bar b[\bar\imath\bar k] \equiv_U \bar b[\bar\jmath\bar k]\,.$$

Fix an enumeration $\bar c$ of U. Since $\bar a[\bar\imath], \bar a[\bar\jmath], \bar c \subseteq P$, it follows by choice of W that

$$\bar a[\bar\imath]\bar c \equiv_W \bar a[\bar\jmath]\bar c \quad \text{implies} \quad a[\bar\imath]\bar c \equiv_C \bar a[\bar\jmath]\bar c\,.$$

Hence, $\bar a[\bar\imath] \equiv_{UC} \bar a[\bar\jmath]$ and the claim follows. □

Corollary 1.18. *Let T be a theory without the independence property and $\alpha = (\bar a^i)_{i \in I}$ an indiscernible sequence over U. For every set $C \subseteq \mathbb{M}$, there exists a convex equivalence relation $\approx$ on I with* $\mathrm{in}(\approx) \leq |T| \oplus |C|$ *such that*

$$\bar\imath \approx \bar\jmath \quad \textit{implies} \quad \bar a[\bar\imath] \equiv_{UC} \bar a[\bar\jmath]\,.$$

Proof. Let $\alpha' = (\bar{a}^j)_{j \in J}$ and $K \subseteq J$ be the sequence and the set obtained from Theorem 1.17. We claim that the restriction $\approx$ of $\equiv^o_K$ to I has the desired properties. By Lemma 1.15, $\approx$ is convex and

$$\operatorname{in}(\approx) \leq |K| \leq |T| \oplus |C| .$$

Consider tuples $\bar{\imath}, \bar{\jmath} \subseteq I$ with $\bar{\imath} \approx \bar{\jmath}$. Then

$$\operatorname{ord}(\bar{\imath}) = \operatorname{ord}(\bar{\jmath}) \quad \text{and} \quad i_s \approx j_s \quad \text{for all } s ,$$

and it follows by Lemma 1.15 that $\bar{\imath} \equiv^o_K \bar{\jmath}$. By choice of α' and K, this implies that $\bar{a}[\bar{\imath}] \equiv_{UC} \bar{a}[\bar{\jmath}]$. □

Corollary 1.19. *Let T be a theory without the independence property, $\alpha = (\bar{a}^i)_{i \in I}$ an indiscernible sequence over U, and $C \subseteq \mathbb{M}$ a set of parameters. If* $\operatorname{cf} I > |T| \oplus |C|$, *then there exists an index $k \in I$ such that the subsequence $(\bar{a}_i)_{i \geq k}$ is indiscernible over $U \cup C \cup \bar{a}[<k]$.*

Proof. Let $\alpha' = (\bar{a}^j)_{j \in J}$ and $K \subseteq J$ be the sequence and the set obtained from Theorem 1.17. Since $\operatorname{cf} I > |K|$, there exists some index $k \in J \setminus K$ that is greater than all elements of K. This index has the desired properties.

□

2. Lascar invariant types

As forking is less well-behaved in non-simple theories, we need additional tools to investigate theories without the independence properties.

Lascar strong types

We start by studying the question of when two tuples $\bar{a}$, $\bar{b}$ can appear as elements of the same indiscernible sequence.

Definition 2.1. For two tuples $\bar{a}$ and $\bar{b}$, we write

$$\bar{a} \approx^{\mathrm{ls}}_U \bar{b} \quad :\text{iff} \quad \text{there is some indiscernible sequence } (\bar{c}_n)_{n<\omega} \text{ over } U \text{ such that } \bar{c}_0 = \bar{a} \text{ and } \bar{c}_1 = \bar{b} .$$

We denote the transitive closure of $\approx^{\mathrm{ls}}_U$ by $\equiv^{\mathrm{ls}}_U$. If $\bar a \equiv^{\mathrm{ls}}_U \bar b$, we say that $\bar a$ and $\bar b$ have the same *Lascar strong type* over U.

Remark. Clearly, $\bar a \equiv^{\mathrm{ls}}_U \bar b$ implies $\bar a \equiv_U \bar b$.

Example. If $b \in \mathrm{acl}(Ua)$, then $a \approx^{\mathrm{ls}}_U b$ iff $a = b$.

Exercise 2.1. Prove that $\approx^{\mathrm{ls}}_U$ is reflexive and symmetric, but in general not transitive.

Let us start by giving an alternative characterisation of the relation $\approx^{\mathrm{ls}}_U$ in terms of formulae that are *chain-bounded.*

Definition 2.2. A formula $\varphi(\bar x, \bar y)$ where $\bar x$ and $\bar y$ have the same sorts is *chain-bounded* if there exists a number $n < \omega$ such that

$$\mathbb{M} \vDash \neg\exists \bar x_0 \cdots \exists \bar x_n \bigwedge_{0 \le i < k \le n} \varphi(\bar x_i, \bar x_k)\,.$$

Remark. Let $\varphi(\bar x, \bar y)$ be a formula where $\bar x$ and $\bar y$ both have sorts $\bar s$. By compactness, it follows that the formula φ is not chain-bounded if, and only if, for every strict linear order $\langle I, < \rangle$, there exist a homomorphism $\langle I, < \rangle \to \langle \mathbb{M}^{\bar s}, \varphi^{\mathbb{M}} \rangle$.

Example. If $\chi(\bar x, \bar y) \in \mathrm{FE}^{\bar s}(U)$, then $\neg\chi(\bar x, \bar y)$ is chain-bounded.

Lemma 2.3. *The following statements are equivalent:*

(1) $\bar a \approx^{\mathrm{ls}}_U \bar b$

(2) $\bar a \approx^{\mathrm{ls}}_C \bar b\,,$ *for all finite $C \subseteq U$.*

(3) $\bar a \approx^{\mathrm{ls}}_M \bar b\,,$ *for some model $M \supseteq U$.*

(4) *For every set C, there exists some set $C' \equiv_U C$ such that $\bar a \approx^{\mathrm{ls}}_{UC'} \bar b$.*

(5) $\mathbb{M} \vDash \neg\varphi(\bar a, \bar b)\,,$ *for every chain-bounded formula φ over U.*

(6) $\bigcup_{0 \le i < k < \omega} \mathfrak{p}(\bar x_i, \bar x_k)$ *is satisfiable, where* $\mathfrak{p}(\bar x, \bar x') := \mathrm{tp}(\bar a \bar b / U)$.

Proof. (4) ⇒ (3) Fix an arbitrary model $\mathfrak{M}$ containing U. By (4), there is some $M' \equiv_U M$ such that $\bar{a} \approx^{\mathrm{ls}}_{M'} \bar{b}$.

(3) ⇒ (1) ⇒ (2) If $(\bar{c}_i)_{i<\omega}$ is an indiscernible sequence over a model $M \supseteq U$ with $\bar{c}_0 = \bar{a}$ and $\bar{c}_1 = \bar{b}$, then $(\bar{c}_i)_{i<\omega}$ is also indiscernible over U.

Similarly, if $(\bar{c}_i)_{i<\omega}$ is indiscernible over U, it is also indiscernible over every subset $C \subseteq U$.

(2) ⇒ (5) Consider a chain-bounded formula $\varphi(\bar{x}, \bar{y})$ over U. Fix a finite set $C \subseteq U$ such that φ is over C. Since $\bar{a} \approx^{\mathrm{ls}}_C \bar{b}$, there exists an indiscernible sequence $(\bar{c}_n)_{n<\omega}$ over C such that $\bar{c}_0 = \bar{a}$ and $\bar{c}_1 = \bar{b}$. If $\mathbb{M} \vDash \varphi(\bar{a}, \bar{b})$, then φ would not be chain-bounded since indiscernibility would imply that

$$\mathbb{M} \vDash \varphi(\bar{c}_i, \bar{c}_k)\,, \quad \text{for all } i < k < \omega\,.$$

Therefore, $\mathbb{M} \vDash \neg\varphi(\bar{a}, \bar{b})$.

(5) ⇒ (6) Suppose that $\bigcup_{0\leq i<k<\omega} \mathfrak{p}(\bar{x}_i, \bar{x}_k)$ is inconsistent. By compactness, there exists a number $n < \omega$ and a finite subset $\Phi \subseteq \mathfrak{p}$ such that $\bigcup_{0\leq i<k<n} \Phi(\bar{x}_i, \bar{x}_k)$ is inconsistent. Setting $\varphi(\bar{x}, \bar{x}') := \bigwedge \Phi$ we have

$$\mathbb{M} \vDash \neg\exists\bar{x}_0 \cdots \exists\bar{x}_{n-1} \bigwedge_{0\leq i<k<n} \varphi(\bar{x}_i, \bar{x}_k)\,.$$

Hence, φ is chain-bounded formula, and $\varphi \in \mathfrak{p}$ implies $\mathbb{M} \nvDash \neg\varphi(\bar{a}, \bar{b})$.

(6) ⇒ (4) Let $(\bar{c}_n)_{n<\omega}$ be a sequence satisfying $\bigcup_{0\leq i<k<\omega} \mathfrak{p}(\bar{x}_i, \bar{x}_k)$. By Proposition E5.3.6, there exists an indiscernible sequence $(\bar{d}_n)_{n<\omega}$ over U with

$$\mathrm{Av}((\bar{c}_n)_{n<\omega}/U) \subseteq \mathrm{Av}((\bar{d}_n)_{n<\omega}/U)\,.$$

Since $\mathfrak{p}(\bar{x}_0, \bar{x}_1) \subseteq \mathrm{Av}((\bar{c}_n)_n/U)$, the sequence $(\bar{d}_n)_{n<\omega}$ also satisfies $\bigcup_{0\leq i<k<\omega} \mathfrak{p}(\bar{x}_i, \bar{x}_k)$. In particular, $\bar{d}_0\bar{d}_1 \equiv_U \bar{a}\bar{b}$ and there exists an automorphism $\pi \in \mathrm{Aut}\,\mathbb{M}_U$ such that $\pi(\bar{d}_0) = \bar{a}$ and $\pi(\bar{d}_1) = \bar{b}$. We can use Lemma E5.3.11 to find a set $C' \equiv_U C$ such that $(\pi(\bar{d}_n))_{n<\omega}$ is indiscernible over $U \cup C'$. It follows that $\bar{a} \approx^{\mathrm{ls}}_{UC'} \bar{b}$. □

Our next goal is to show that, for a model $\mathfrak{M}$, the relation $\equiv^{\mathrm{ls}}_M$ coincides with $\equiv_M$. We start with a technical lemma.

Lemma 2.4. *If $\varphi(\bar{x}, \bar{y})$ and $\psi(\bar{x}, \bar{y})$ are chain-bounded, then so is $\varphi \vee \psi$.*

Proof. Suppose that $\varphi \vee \psi$ is not chain-bounded. Then there exists a sequence $(\bar{c}_n)_{n<\omega}$ such that

$$\mathbb{M} \vDash (\varphi \vee \psi)(\bar{c}_i, \bar{c}_k), \quad \text{for all } i < k < \omega .$$

By the Theorem of Ramsey, we can find an infinite subset $I \subseteq \omega$ such that

$$\begin{aligned} & \mathbb{M} \vDash \varphi(\bar{c}_i, \bar{c}_k), \quad \text{for all } i < k \text{ in } I, \\ \text{or} \quad & \mathbb{M} \vDash \psi(\bar{c}_i, \bar{c}_k), \quad \text{for all } i < k \text{ in } I. \end{aligned}$$

In the first case, φ is not chain-bounded; in the second case, ψ is not chain-bounded. □

Proposition 2.5. *For a model $\mathfrak{M}$, the following statements are equivalent:*

(1) $\bar{a} \equiv^{\mathrm{ls}}_M \bar{b}$

(2) $\bar{a} \equiv_M \bar{b}$

(3) $\bar{a} \approx^{\mathrm{ls}}_M \bar{c} \approx^{\mathrm{ls}}_M \bar{b}$, *for some* $\bar{c}$.

(4) *There exist tuples $\bar{c}_0, \bar{c}_1, \dots$ such that the sequences $\bar{a}, \bar{c}_0, \bar{c}_1, \bar{c}_2, \dots$ and $\bar{b}, \bar{c}_0, \bar{c}_1, \bar{c}_2, \dots$ are both indiscernible over M.*

(5) $\mathbb{M} \vDash \exists \bar{y}[\neg\varphi(\bar{a}, \bar{y}) \wedge \neg\varphi(\bar{b}, \bar{y})]$, *for every chain-bounded formula $\varphi(\bar{x}, \bar{y})$ over M.*

Proof. (3) ⇒ (1) is trivial.

(1) ⇒ (2) By definition of $\equiv^{\mathrm{ls}}_M$, there are tuples $\bar{c}_0, \dots, \bar{c}_n$ such that

$$\bar{a} = \bar{c}_0 \approx^{\mathrm{ls}}_M \cdots \approx^{\mathrm{ls}}_M \bar{c}_n = \bar{b} .$$

For each $k < n$, there is an indiscernible sequence $(\bar{d}^k_i)_{i<\omega}$ over M with $\bar{d}^k_0 = \bar{c}_k$ and $\bar{d}^k_1 = \bar{c}_{k+1}$. Consequently, $\bar{c}_k \equiv_M \bar{c}_{k+1}$ and the claim follows.

(2) ⇒ (4) Suppose that $\bar{a} \equiv_M \bar{b}$. By Lemma F2.3.15, we have $\bar{a} \overset{\mathrm{u}}{\surd}_M M$. As $\overset{\mathrm{u}}{\surd}$ is a forking relation, the type $\mathrm{tp}(\bar{a}/M)$ has some $\overset{\mathrm{u}}{\surd}$-free extension $\mathfrak{p} \in S^{\bar{s}}(\mathbb{M})$. We construct a sequence $\beta = (\bar{c}_n)_{n<\omega}$ by inductively

choosing a tuple $\bar{c}_n$ realising $\mathfrak{p} \upharpoonright M\bar{a}\bar{a}'\bar{c}[<n]$. Since $\sqrt[\mathrm{u}]{} \subseteq \sqrt[\mathrm{s}]{}$, the type $\mathfrak{p}$ is invariant over M and the sequences $\alpha := \bar{a}\beta$ and $\alpha' := \bar{a}'\beta$ both satisfy the conditions of Lemma F2.4.14 (b). Hence, they are indiscernible over M.

(4) ⇒ (3) Suppose that $\bar{a}, \bar{c}_0, \bar{c}_1, \bar{c}_2, \dots$ and $\bar{b}, \bar{c}_0, \bar{c}_1, \bar{c}_2, \dots$ are indiscernible sequences over M. Then

$$\bar{a} \approx^{\mathrm{ls}}_M \bar{c}_0 \quad \text{and} \quad \bar{b} \approx^{\mathrm{ls}}_M \bar{c}_0\,,$$

and the claim follows by symmetry of $\approx^{\mathrm{ls}}_M$.

(2) ⇒ (5) Suppose that $\bar{a} \equiv_M \bar{b}$. Let $\varphi(\bar{x}, \bar{y})$ be a chain-bounded formula over M and let n be the minimal number such that

$$\mathbb{M} \models \neg\exists\bar{x}_0 \cdots \exists\bar{x}_n \bigwedge_{0 \leq i < k \leq n} \varphi(\bar{x}_i, \bar{x}_k)\,.$$

Then

$$\mathbb{M} \models \exists\bar{x}_0 \cdots \exists\bar{x}_{n-1} \bigwedge_{0 \leq i < k < n} \varphi(\bar{x}_i, \bar{x}_k)\,.$$

As the same formula holds in $\mathfrak{M}$, there are tuples $\bar{c}_0, \dots, \bar{c}_{n-1}$ in M such that

$$\mathfrak{M} \models \bigwedge_{0 \leq i < k < n} \varphi(\bar{c}_i, \bar{c}_k)\,.$$

By choice of n, there is an index $k < n$ such that $\mathbb{M} \not\models \varphi(\bar{a}, \bar{c}_k)$. Since $\bar{a} \equiv_M \bar{b}$ we also have $\mathbb{M} \not\models \varphi(\bar{b}, \bar{c}_k)$. Consequently,

$$\mathbb{M} \models \neg\varphi(\bar{a}, \bar{c}_k) \wedge \neg\varphi(\bar{b}, \bar{c}_k)\,.$$

(5) ⇒ (3) Set

$$\Phi(\bar{y}) := \big\{\, \neg\varphi(\bar{a}, \bar{y}) \wedge \neg\varphi(\bar{b}, \bar{y}) \;\big|\; \varphi(\bar{x}, \bar{y}) \text{ a chain-bounded formula over } M \,\big\}\,.$$

If there is a tuple $\bar{c}$ satisfying Φ, then it follows from Lemma 2.3 that

$$\bar{a} \approx^{\mathrm{ls}}_M \bar{c} \quad \text{and} \quad \bar{b} \approx^{\mathrm{ls}}_M \bar{c}\,.$$

Hence, it remains to show that $T(\mathbb{M}) \cup \Phi$ is satisfiable. Let $\Phi_0 \subseteq \Phi$ be finite. Then there are chain-bounded formulae $\varphi_0, \dots, \varphi_{n-1}$ over M such that

$$\Phi_0 = \{\, \neg\varphi_i(\bar{a}, \bar{y}) \wedge \neg\varphi_i(\bar{b}, \bar{y}) \mid i < n \,\}\,.$$

By Lemma 2.4 the disjunction $\psi := \varphi_0 \vee \dots \vee \varphi_{n-1}$ is also chain-bounded. Therefore, (5) implies that there is some tuple $\bar{c}$ with

$$\mathbb{M} \vDash \neg\psi(\bar{a}, \bar{c}) \wedge \neg\psi(\bar{b}, \bar{c})\,.$$

Consequently, $\bar{c}$ satisfies $T(\mathbb{M}) \cup \Phi_0$. By compactness, it follows that $T(\mathbb{M}) \cup \Phi$ is satisfiable. □

Finally we provide several characterisations of the relation $\equiv^{\mathrm{ls}}_U$ for arbitrary sets U. One of them is in terms of bounded equivalence relations, where boundedness is an analog to the notion of chain-boundedness, but for the complement of the relation.

Definition 2.6. Let $\mathbb{R} \subseteq \mathbb{M}^{\bar{s}} \times \mathbb{M}^{\bar{s}}$ be a relation.

(a) $\mathbb{R}$ is *U-invariant* if

$$\bar{a}\bar{b} \equiv_U \bar{a}'\bar{b}' \quad \text{implies} \quad \langle \bar{a}, \bar{b} \rangle \in \mathbb{R} \Leftrightarrow \langle \bar{a}', \bar{b}' \rangle \in \mathbb{R}\,.$$

(b) $\mathbb{R}$ is *co-chain-bounded* if there exists a small cardinal κ such that, for every sequence $\alpha = (\bar{a}_i)_{i<\kappa}$ in $\mathbb{M}^{\bar{s}}$, there are indices $i < j$ with $\langle \bar{a}_i, \bar{a}_j \rangle \in \mathbb{R}$. A co-chain-bounded equivalence relation is simply called *bounded.*

Before concentrating on equivalence relations, let us first give several characterisations of co-chain-boundedness for arbitrary relations.

Proposition 2.7. *Let $\mathbb{R} \subseteq \mathbb{M}^{\bar{s}} \times \mathbb{M}^{\bar{s}}$ be a U-invariant relation. The following statements are equivalent.*

(1) $\mathbb{R}$ *is co-chain-bounded.*

(2) $\approx^{\mathrm{ls}}_U \subseteq \mathbb{R}$

(3) *For every indiscernible sequence* $(\bar{a}_n)_{n<\omega}$ *over* U *with* $\bar{a}_n \in \mathbb{M}^{\bar{s}}$, *we have* $\langle \bar{a}_i, \bar{a}_j \rangle \in \mathbb{R}$, *for all* $i < j < \omega$.

Proof. (2) ⇒ (3) Let $(\bar{a}_n)_{n<\omega}$ be an indiscernible sequence over U. For every pair of indices $i < j < \omega$, we obtain an indiscernible sequence $\bar{a}_i, \bar{a}_j, \bar{a}_{j+1}, \dots$ over U, which witnesses that $\bar{a}_i \approx^{\mathrm{ls}}_U \bar{a}_j$. By (2), this implies that $\langle \bar{a}_i, \bar{a}_j \rangle \in \mathbb{R}$.

(3) ⇒ (2) Let $\bar{a} \approx^{\mathrm{ls}}_U \bar{b}$. By definition, there exists an indiscernible sequence $(\bar{c}_n)_{n<\omega}$ over U with $\bar{c}_0 = \bar{a}$ and $\bar{c}_1 = \bar{b}$. Hence, it follows by (3) that $\langle \bar{c}_0, \bar{c}_1 \rangle \in \mathbb{R}$.

(1) ⇒ (3) Let $\mathbb{R}$ be co-chain-bounded and let κ be the corresponding cardinal. For a contradiction, suppose that there exists an indiscernible sequence $\alpha = (\bar{a}_n)_{n<\omega}$ such that $\langle \bar{a}_i, \bar{a}_j \rangle \notin \mathbb{R}$, for some $i < j$. We extend α to an indiscernible sequence $(\bar{a}_i)_{i<\kappa}$ of length κ. By U-invariance, it follows that $\langle \bar{a}_i, \bar{a}_j \rangle \notin \mathbb{R}$, for all $i < j < \kappa$. This contradicts our choice of κ.

(3) ⇒ (1) Suppose that $\mathbb{R}$ is not co-chain-bounded. Then there exists a sequence $(\bar{a}_i)_{i<\kappa}$ of length $\kappa := \beth_{\lambda^+}$ where $\lambda := 2^{|T| \oplus |U| \oplus |\bar{s}|}$ such that

$$\langle \bar{a}_i, \bar{a}_j \rangle \notin \mathbb{R}, \quad \text{for all } i < j < \kappa .$$

We can use Theorem E5.3.7 to find an indiscernible sequence $(\bar{b}_n)_{n<\omega}$ over U such that, for every $\bar{\imath} \in [\omega]^{<\omega}$, there is some $\bar{\jmath} \in [\kappa]^{<\omega}$ with

$$\bar{b}[\bar{\imath}] \equiv_U \bar{a}[\bar{\jmath}] .$$

By U-invariance, it follows that $\langle \bar{b}_i, \bar{b}_j \rangle \notin \mathbb{R}$, for all $i < j < \omega$. This contradicts (3). □

For equivalence relations, we obtain the following characterisation.

Proposition 2.8. *Let* $\approx$ *be a* U*-invariant equivalence relation on* $\mathbb{M}^{\bar{s}}$. *The following statements are equivalent:*

(1) $\approx$ *is bounded.*

(2) $\approx$ *has at most* $2^{|T| \oplus |U| \oplus |\bar{s}|}$ *classes.*

(3) $\equiv^{\mathrm{ls}}_U \subseteq \approx$

(4) *For every indiscernible sequence* $(\bar{a}_n)_{n<\omega}$ *over* U *with* $\bar{a}_n \in \mathbb{M}^{\bar{s}}$, *we have* $\bar{a}_i \approx \bar{a}_j$, *for all* $i, j < \omega$.

(5) *For every model* $\mathfrak{M}$ *containing* U,

$$\bar{a} \equiv_M \bar{b} \quad \textit{implies} \quad \bar{a} \approx \bar{b}\,, \quad \textit{for all } \bar{a}, \bar{b} \in \mathbb{M}^{\bar{s}}.$$

Proof. (2) ⇒ (1) is trivial, and the equivalence (1) ⇔ (4) has already been proved in Proposition 2.7. The equivalence (1) ⇔ (3) also follows by Proposition 2.7 since $\approx$ is an equivalence relation and $\equiv^{\mathrm{ls}}_U$ is the transitive closure of $\approx^{\mathrm{ls}}_U$. Consequently, we have

$$\equiv^{\mathrm{ls}}_U \subseteq \approx \quad \text{iff} \quad \approx^{\mathrm{ls}}_U \subseteq \approx\,.$$

(4) ⇒ (5) Suppose that $\bar{a} \equiv_M \bar{b}$. By Proposition 2.5 (4), we can find a sequence $\gamma = (\bar{c}_n)_{n<\omega}$ such that $\bar{a}\gamma$ and $\bar{b}\gamma$ are both indiscernible over M. By (4), this implies that $\bar{a} \approx \bar{c}_0 \approx \bar{b}$.

(5) ⇒ (2) Fix a model $\mathfrak{M}$ containing U of size $|M| \leq |T| \oplus |U|$. Then $\equiv_M \subseteq \approx$ implies that $\approx$ has at most as many classes as $\equiv_M$. The latter number is $|S^{\bar{s}}(M)| \leq 2^{|T| \oplus |M| \oplus |\bar{s}|} = 2^{|T| \oplus |U| \oplus |\bar{s}|}$. □

Corollary 2.9. *Let* $U \subseteq \mathbb{M}$.

(a) $\approx^{\mathrm{ls}}_U$ *is the finest relation that is co-chain-bounded and* U*-invariant.*

(b) $\equiv^{\mathrm{ls}}_U$ *is the finest equivalence relation that is bounded and* U*-invariant.*

Over arbitrary sets U, we can characterise the relation $\equiv^{\mathrm{ls}}_U$ as follows.

Proposition 2.10. *Let* $\bar{a}, \bar{b} \in \mathbb{M}^{\bar{s}}$ *and* $U \subseteq \mathbb{M}$. *The following statements are equivalent:*

(1) $\bar{a} \equiv^{\mathrm{ls}}_U \bar{b}$

(2) *$\bar a \approx \bar b$, for every equivalence relation $\approx$ on $\mathbb{M}$ that is bounded and U-invariant.*

(3) *There are tuples $\bar c_0, \dots, \bar c_n$ and models $M_0, \dots, M_{n-1} \supseteq U$, for some $n < \omega$, such that*

$$\bar a = \bar c_0 \equiv_{M_0} \bar c_1 \equiv_{M_1} \cdots \equiv_{M_{n-2}} \bar c_{n-1} \equiv_{M_{n-1}} \bar c_n = \bar b\,.$$

(4) *There are models $M_0, \dots, M_{n-1} \supseteq U$, for some $n < \omega$, and automorphisms $\pi_i \in \operatorname{Aut} \mathbb{M}_{M_i}$ such that*

$$\bar b = (\pi_{n-1} \circ \cdots \circ \pi_0)(\bar a)\,.$$

Proof. (3) $\Leftrightarrow$ (4) follows from the fact that $\bar c_i \equiv_{M_i} \bar c_{i+1}$ if, and only if, there exists some automorphism $\pi_i \in \operatorname{Aut} \mathbb{M}_{M_i}$ with $\bar c_{i+1} = \pi(\bar c_i)$.

(1) $\Rightarrow$ (2) follows by Proposition 2.8 (3).

(2) $\Rightarrow$ (3) Let $\sim^*$ be the transitive closure of the relation

$$\bar c \sim \bar d \quad :\text{iff} \quad \bar c \equiv_M \bar d\,, \quad \text{for some model } \mathfrak{M} \text{ containing } U\,.$$

This relation is clearly U-invariant. Furthermore, it is bounded since it satisfies property (4) of Proposition 2.8. By (2), it follows that $\bar a \sim^* \bar b$.

(3) $\Rightarrow$ (1) By Proposition 2.5, there are tuples $\bar d_i$, for $i < n$, such that

$$\bar c_i \approx^{\text{ls}}_{M_i} \bar d_i \approx^{\text{ls}}_{M_i} \bar c_{i+1}\,.$$

According to Lemma 2.3 this implies that

$$\bar c_i \approx^{\text{ls}}_{U} \bar d_i \approx^{\text{ls}}_{U} \bar c_{i+1}\,, \quad \text{for all } i < n\,.$$

Hence, $\bar a = \bar c_0 \equiv^{\text{ls}}_U \bar c_n = \bar b$. □

Two tuples are said to have the same *strong type* over a set U if they are elementarily equivalent over $\operatorname{acl}^{\text{eq}}(U)$. The next result shows that having the same Lascar strong type implies having the same strong type.

Corollary 2.11. $\bar a \equiv^{\text{ls}}_U \bar b$ *implies* $\bar a \equiv_{\operatorname{acl}^{\text{eq}}(U)} \bar b\,.$

Proof. Suppose that $\bar{a} \equiv^{\mathrm{ls}}_U \bar{b}$. We can use Proposition 2.10 to find tuples $\bar{c}_0, \dots, \bar{c}_n$ and models $M_0, \dots, M_{m-1} \supseteq U$ such that

$$\bar{a} = \bar{c}_0 \equiv_{M_0} \cdots \equiv_{M_{n-1}} \bar{c}_n = \bar{b}\,.$$

This implies that

$$\bar{a} = \bar{c}_0 \equiv_{M_0^{\mathrm{eq}}} \cdots \equiv_{M_{n-1}^{\mathrm{eq}}} \bar{c}_n = \bar{b}\,.$$

Since $\mathrm{acl}^{\mathrm{eq}}(U) \subseteq M_i^{\mathrm{eq}}$, for all i, it follows that

$$\bar{a} = \bar{c}_0 \equiv_{\mathrm{acl}^{\mathrm{eq}}(U)} \cdots \equiv_{\mathrm{acl}^{\mathrm{eq}}(U)} \bar{c}_n = \bar{b}\,. \qquad \square$$

We conclude our investigation of Lascar strong types by two technical results. The first one shows that the relation $\approx^{\mathrm{ls}}_U$ satisfies a restricted form of the back-and-forth property.

Lemma 2.12. *If $\bar{a} \approx^{\mathrm{ls}}_U \bar{b}$ and $\bar{c} \sqrt[\mathrm{d}]{}_{U\bar{a}}\ \bar{b}$, there exists a tuple $\bar{d}$ such that $\bar{a}\bar{c} \approx^{\mathrm{ls}}_U \bar{b}\bar{d}$.*

Proof. Let $(\bar{a}_i)_{i<\omega}$ be an indiscernible sequence over U with $\bar{a}_0 = \bar{a}$ and $\bar{a}_1 = \bar{b}$. Since the subsequence $(\bar{a}_i)_{0<i<\omega}$ is indiscernible over $U \cup \bar{a}$ and $\bar{c} \sqrt[\mathrm{d}]{}_{U\bar{a}}\ \bar{b}$, we can use Lemma F3.1.3 to find an element $\bar{c}' \equiv_{U\bar{a}\bar{b}} \bar{c}$ such that $(\bar{a}_i)_{0<i<\omega}$ is indiscernible over $U\bar{a}\bar{c}'$. Applying an $U\bar{a}\bar{b}$-automorphism mapping $\bar{c}'$ to $\bar{c}$, we obtain an indiscernible sequence $(\bar{a}'_i)_{0<i<\omega}$ over $U\bar{a}\bar{c}$ such that

$$(\bar{a}'_i)_{0<i<\omega} \equiv_{U\bar{a}\bar{b}} (\bar{a}_i)_{0<i<\omega}\,.$$

Replacing $\bar{a}_i$ by $\bar{a}'_i$, for $0 < i < \omega$, we may therefore assume that the sequence $(\bar{a}_i)_{0<i<\omega}$ is indiscernible over $U\bar{a}\bar{c}$.

For every $i < \omega$, we choose an automorphism $\pi_i \in \mathrm{Aut}\,\mathbb{M}_U$ such that $\pi_i(\bar{a}_n) = \bar{a}_{n+i}$, for all n, and we set $\bar{c}_i := \pi_i(\bar{c})$. Since $(\bar{a}_i)_{0<i<\omega}$ is indiscernible over $U\bar{a}\bar{c}$, it follows that

$$\bar{c}\bar{a}\bar{b} \equiv_U \bar{c}\bar{a}\bar{a}_n \equiv_U \bar{c}_i\bar{a}_i\bar{a}_{n+i}\,, \quad \text{for all } i < \omega \text{ and } 0 < n < \omega\,.$$

By Proposition E5.3.6, there exists an indiscernible sequence $(\bar c'_i \bar a'_i)_{i<\omega}$ over U such that

$$\mathrm{Av}((\bar c_i \bar a_i)_{i<\omega}/U) \subseteq \mathrm{Av}((\bar c'_i \bar a'_i)_{i<\omega}/U)\,.$$

In particular, we have

$$\bar c'_i \bar a'_i \bar a'_{n+i} \equiv_U \bar c_i \bar a_i \bar a_{n+i} \equiv_U \bar c \bar a \bar b\,.$$

Let σ be an U-automorphism such that $\sigma(\bar c'_0) = \bar c$, $\sigma(\bar a'_0) = \bar a$, and $\sigma(\bar a'_1) = \bar b$. The tuple $\bar d := \sigma(\bar c'_1)$ has the desired properties. □

The second observation contains a strengthening of the extension axiom.

Lemma 2.13. *Let $\surd$ be a forking relation and suppose that $\bar a \surd_U U$. For every set B, there exists a tuple $\bar a' \approx^{\mathrm{ls}}_U \bar a$ such that $\bar a' \surd_U B$.*

Proof. Since $\bar a \surd_U U$, we can use Proposition F2.4.10 to construct a $\surd$-Morley sequence $(\bar a_n)_{n<\omega}$ for $\mathrm{tp}(\bar a/U)$ over U. Applying a suitable automorphism we may assume that $\bar a_0 = \bar a$. Since $\bar a[>0] \surd_U \bar a_0$, there exists a sequence $\alpha' \equiv_{U\bar a_0} \bar a[>0]$ such that $\alpha' \surd_U B\bar a_0$. Let $\alpha' = (\bar a'_i)_{0<i<\omega}$. As $\bar a_0\alpha'$ is indiscernible over U, we have $\bar a_0 \approx^{\mathrm{ls}}_U \bar a'_1$. Since $\bar a'_1 \surd_U B$, the claim follows. □

Lascar invariance

To study theories without the independence property, we introduce variants of the relations $\overset{\mathrm{s}}{\surd}$ and $\overset{\mathrm{i}}{\surd}$ that are based on Lascar strong types instead of elementary equivalence.

Definition 2.14. For $A, B, U \subseteq \mathbb{M}$ we define

$$\begin{aligned}
A \overset{\mathrm{q}}{\surd}_U B \quad &:\text{iff}\quad \bar b \approx^{\mathrm{ls}}_U \bar b' \Rightarrow \bar b \approx^{\mathrm{ls}}_{UA} \bar b' \quad \text{for all } \bar b, \bar b' \subseteq B\,,\\
A \overset{\mathrm{ls}}{\surd}_U B \quad &:\text{iff}\quad \bar b \equiv^{\mathrm{ls}}_U \bar b' \Rightarrow \bar b \equiv_{UA} \bar b' \quad \text{for all } \bar b, \bar b' \subseteq B\,,\\
A \overset{\mathrm{li}}{\surd}_U B \quad &:\text{iff}\quad A \,{}^*(\overset{\mathrm{ls}}{\surd})_U\, B\,.
\end{aligned}$$

If $\bar{a} \sqrt[\mathrm{ls}]{}_U B$, we say that $\mathrm{tp}(\bar{a}/UB)$ is *Lascar invariant* over U.

Note that $\sqrt[\mathrm{s}]{} \subseteq \sqrt[\mathrm{ls}]{} \subseteq \sqrt[\mathrm{q}]{}$. Unfortunately, the relation $\sqrt[\mathrm{ls}]{}$ is not a preforking relation since it fails transitivity. But $\sqrt[\mathrm{q}]{}$ is. Hence, in order to show that $\sqrt[\mathrm{li}]{}$ is a forking relation, we will prove below that $\sqrt[\mathrm{li}]{} = {}^*(\sqrt[\mathrm{q}]{})$.

Exercise 2.2. Prove that $\sqrt[\mathrm{ls}]{}$ satisfies all axioms of a preforking relation except for (LTR).

Before turning to $\sqrt[\mathrm{li}]{}$, we take a look at the relation $\sqrt[\mathrm{q}]{}$.

Lemma 2.15. $\sqrt[\mathrm{q}]{}$ *is a preforking relation.*

Proof. (INV) follows immediately from the definition.

(MON) Suppose that $A \sqrt[\mathrm{q}]{}_U B$ and let $A_0 \subseteq A$ and $B_0 \subseteq B$. For tuples $\bar{b}, \bar{b}' \subseteq B_0 \subseteq B$, we have

$$\bar{b} \approx^{\mathrm{ls}}_U \bar{b}' \;\Rightarrow\; \bar{b} \approx^{\mathrm{ls}}_{UA} \bar{b}' \;\Rightarrow\; \bar{b} \approx^{\mathrm{ls}}_{UA_0} \bar{b}' .$$

(BMON) Suppose that $A \sqrt[\mathrm{q}]{}_U BC$ and let $\bar{b}, \bar{b}' \subseteq B$. Fixing an enumeration $\bar{c}$ of C, we have

$$\bar{b} \approx^{\mathrm{ls}}_{UC} \bar{b}' \;\Rightarrow\; \bar{b}\bar{c} \approx^{\mathrm{ls}}_U \bar{b}'\bar{c} \;\Rightarrow\; \bar{b}\bar{c} \approx^{\mathrm{ls}}_{UA} \bar{b}'\bar{c} \;\Rightarrow\; \bar{b} \approx^{\mathrm{ls}}_{UCA} \bar{b}' .$$

(NOR) Suppose that $A \sqrt[\mathrm{q}]{}_U B$. To show that $AU \sqrt[\mathrm{q}]{}_U BU$, consider tuples $\bar{b}, \bar{b}' \subseteq U \cup B$ with $\bar{b} \approx^{\mathrm{ls}}_U \bar{b}'$. Reordering $\bar{b}$ and $\bar{b}'$, we may assume that $\bar{b} = \bar{b}_0\bar{c}$ and $\bar{b}' = \bar{b}'_0\bar{c}$ for $\bar{b}_0, \bar{b}'_0 \subseteq B$ and $\bar{c} \subseteq U$. Consequently,

$$\bar{b}_0\bar{c} \approx^{\mathrm{ls}}_U \bar{b}'_0\bar{c} \;\Rightarrow\; \bar{b}_0 \approx^{\mathrm{ls}}_U \bar{b}'_0 \;\Rightarrow\; \bar{b}_0 \approx^{\mathrm{ls}}_{UA} \bar{b}'_0 \;\Rightarrow\; \bar{b}_0\bar{c} \approx^{\mathrm{ls}}_{UA} \bar{b}'_0\bar{c} .$$

(LRF) To show that $A \sqrt[\mathrm{q}]{}_A B$, let $\bar{b}, \bar{b}' \subseteq B$. Since, trivially,

$$\bar{b} \approx^{\mathrm{ls}}_A \bar{b}' \quad \text{implies} \quad \bar{b} \approx^{\mathrm{ls}}_A \bar{b}' ,$$

the claim follows.

(LTR) Suppose that $A_2 \mathop{\overset{\mathrm{q}}{\smile\!\!\!\!\!|}}_{A_1} B$ and $A_1 \mathop{\overset{\mathrm{q}}{\smile\!\!\!\!\!|}}_{A_0} B$ for $A_0 \subseteq A_1 \subseteq A_2$. To show that $A_2 \mathop{\overset{\mathrm{q}}{\smile\!\!\!\!\!|}}_{A_0} B$, consider two tuples $\bar b, \bar b' \subseteq B$. Then

$$\bar b \approx^{\mathrm{ls}}_{A_0} \bar b' \quad\Rightarrow\quad \bar b \approx^{\mathrm{ls}}_{A_1} \bar b' \quad\Rightarrow\quad \bar b \approx^{\mathrm{ls}}_{A_2} \bar b'\,.$$

(FIN) Suppose that $A_0 \mathop{\overset{\mathrm{q}}{\smile\!\!\!\!\!|}}_{U} B$, for all finite $A_0 \subseteq A$. To show that $A \mathop{\overset{\mathrm{q}}{\smile\!\!\!\!\!|}}_{U} B$, consider two tuples $\bar b, \bar b' \subseteq B$. Then

$$\bar b \approx^{\mathrm{ls}}_{U} \bar b' \quad \text{implies} \quad \bar b \approx^{\mathrm{ls}}_{UA_0} \bar b'\,, \quad \text{for all finite } A_0 \subseteq A\,.$$

By Lemma 2.3, it follows that $\bar b \approx^{\mathrm{ls}}_{UA} \bar b'$.

(DEF) Suppose that $\bar a \not\mathop{\overset{\mathrm{q}}{\smile\!\!\!\!\!|}}_{U} B$. Then there are tuples $\bar b, \bar b' \subseteq B$ such that

$$\bar b \approx^{\mathrm{ls}}_{U} \bar b' \quad \text{and} \quad \bar b \not\approx^{\mathrm{ls}}_{U\bar a} \bar b'\,.$$

By Lemma 2.3, there exists some formula $\varphi(\bar x, \bar y; \bar z)$ over U such that $\varphi(\bar x, \bar y; \bar a)$ is chain-bounded and $\mathbb{M} \vDash \varphi(\bar b, \bar b'; \bar a)$. Let n be the minimal number such that

$$\mathbb{M} \vDash \neg\exists \bar x_0 \cdots \exists \bar x_{n-1} \bigwedge_{0 \le i < k < n} \varphi(\bar x_i, \bar x_k; \bar a)\,,$$

and set

$$\psi(\bar z) := \varphi(\bar b, \bar b'; \bar z) \wedge \neg\exists \bar x_0 \cdots \exists \bar x_{n-1} \bigwedge_{0 \le i < k < n} \varphi(\bar x_i, \bar x_k; \bar z)\,.$$

If $\bar a'$ is a tuple satisfying $\psi(\bar x)$, then $\varphi(\bar x, \bar y; \bar a')$ is chain-bounded and it follows by Lemma 2.3 that $\bar b \not\approx^{\mathrm{ls}}_{U\bar a'} \bar b'$. Hence, $\bar a' \not\mathop{\overset{\mathrm{q}}{\smile\!\!\!\!\!|}}_{U} B$. □

There is also a characterisation of $\mathop{\overset{\mathrm{q}}{\smile\!\!\!\!\!|}}$ in terms of indiscernible sequences, which is obtained by simply replacing the relation $\approx^{\mathrm{ls}}_{U}$ by its definition.

Lemma 2.16. *$A \mathop{\overset{\mathrm{q}}{\smile\!\!\!\!\!|}}_{U} B$ if, and only if, for every indiscernible sequence $(\bar b_i)_{i<\omega}$ over U with $\bar b_0, \bar b_1 \subseteq B$, we can find some indiscernible sequence $(\bar b'_i)_{i<\omega}$ over $U \cup A$ with $\bar b'_0 = \bar b_0$ and $\bar b'_1 = \bar b_1$.*

Proof. (⇐) To show that $A \overset{\mathrm{q}}{\smile}_U B$, consider two tuples $\bar b, \bar b' \subseteq B$ with $\bar b \approx^{\mathrm{ls}}_U \bar b'$. Then there is some indiscernible sequence $(\bar c_i)_{i<\omega}$ over U with $\bar c_0 = \bar b$ and $\bar c_1 = \bar b'$. By assumption, we can find an indiscernible sequence $(\bar c'_i)_{i<\omega}$ over $U \cup A$ with $\bar c'_0 = \bar c_0$ and $\bar c'_1 = \bar c_1$. This implies that $\bar b = \bar c'_0 \approx^{\mathrm{ls}}_{U\cup A} \bar c'_1 = \bar b'$.

(⇒) Suppose that $A \overset{\mathrm{q}}{\smile}_U B$ and let $(\bar b_i)_{i<\omega}$ be an indiscernible sequence over U with $\bar b_0, \bar b_1 \subseteq B$. Then $\bar b_0 \approx^{\mathrm{ls}}_U \bar b_1$, which implies that $\bar b_0 \approx^{\mathrm{ls}}_{UA} \bar b_1$. Consequently, there is some indiscernible sequence $(\bar b'_i)_{i<\omega}$ over $U \cup A$ with $\bar b'_0 = \bar b_0$ and $\bar b'_1 = \bar b_1$. □

Before proving that $\overset{\mathrm{li}}{\smile}$ is a forking relation, we collect several different characterisations of this relation. We start with the following one.

Lemma 2.17. *$A \overset{\mathrm{li}}{\smile}_U B$ if, and only if, for every finite set of indiscernible sequences $\alpha_0, \dots, \alpha_{n-1}$ over U, there exists a set $A' \equiv_{UB} A$ such that each α_i is indiscernible over $U \cup A'$.*

Proof. (⇒) Suppose that $A \overset{\mathrm{li}}{\smile}_U B$ and let $\alpha_0, \dots, \alpha_{n-1}$ be indiscernible over U. W.l.o.g. we may assume that each α_i is indexed by a dense order I_i. By definition of $\overset{\mathrm{li}}{\smile}$, there exists a set $A' \equiv_{UB} A$ such that

$$A' \overset{\mathrm{ls}}{\smile}_U B\alpha_0 \dots \alpha_{n-1}\,.$$

We claim that each sequence α_i is indiscernible over $U \cup A'$. Suppose that $\alpha_i = (\bar a^i_j)_{j\in I_i}$. By Lemma E5.3.12, it is sufficient to prove that

$$\bar a^i[\bar k] \equiv_{UA'} \bar a^i[\bar l]\,, \quad \text{for all } \bar k, \bar l \in [I_i]^n \text{ such that } \bar k = \bar u s \bar v \text{ and } \bar l = \bar u t \bar v \text{ with } s < t\,.$$

Given $\bar u, \bar v, s, t$, we fix a strictly increasing function $g : \omega \to I_i$ such that

$$g(0) = s\,, \quad g(1) = t\,, \quad \text{and} \quad g(j) < \bar v\,, \quad \text{for all } j < \omega\,.$$

The sequence $(\bar a^i[\bar u g(j) \bar v])_{j<\omega}$ witnesses that

$$\bar a^i[\bar u s \bar v] \approx^{\mathrm{ls}}_U \bar a^i[\bar u t \bar v]\,.$$

Therefore, $A' \mathop{\sqrt{}}^{\mathrm{ls}}_U B\alpha_0 \dots \alpha_{n-1}$ implies that $\bar a^i[\bar u s \bar v] \equiv_{UA'} \bar a^i[\bar u t \bar v]$.

($\Leftarrow$) Let $\bar a$, B, and U be sets such that, for all indiscernible sequences $\alpha_0, \dots, \alpha_{n-1}$ over U, there is some tuple $\bar a' \equiv_{UB} \bar a$ such that each α_i is indiscernible over $U \cup \bar a'$. To show that $\bar a \mathop{\sqrt{}}^{\mathrm{li}}_U B$, consider some set $C \subseteq \mathbb{M}$. We have to find some tuple $\bar a' \equiv_{UB} \bar a$ such that $\bar a' \mathop{\sqrt{}}^{\mathrm{ls}}_U BC$. To do so, it is sufficient to prove that the set

$$\begin{aligned}\Phi(\bar x) &:= \mathrm{tp}(\bar a/UB)\\ &\qquad \cup \{\, \varphi(\bar x; \bar b) \leftrightarrow \varphi(\bar x; \bar b') \mid \bar b, \bar b' \subseteq UBC\,,\ \bar b \equiv^{\mathrm{ls}}_U \bar b' \,\}\end{aligned}$$

is satisfiable. Hence, consider a finite subset $\Phi_0 \subseteq \Phi$. Then there are formulae $\varphi_0(\bar x; \bar y_0), \dots, \varphi_n(\bar x; \bar y_n)$ and parameters $\bar b_0, \bar b'_0, \dots, \bar b_n, \bar b'_n \subseteq U \cup B \cup C$ such that $\bar b_i \equiv^{\mathrm{ls}}_U \bar b'_i$, for all $i \leq n$, and

$$\Phi_0 \subseteq \mathrm{tp}(\bar a/UB) \cup \{\, \varphi_i(\bar x; \bar b_i) \leftrightarrow \varphi_i(\bar x; \bar b'_i) \mid i \leq n \,\}\,.$$

For each $i \leq n$, we fix a finite sequence $\bar c^i_0 \approx^{\mathrm{ls}}_U \dots \approx^{\mathrm{ls}}_U \bar c^i_{m(i)}$ with $\bar c^i_0 = \bar b_i$ and $\bar c^i_{m(i)} = \bar b'_i$ and, for every $j < m(i)$, we choose an indiscernible sequence β^i_j over U starting with the tuples $\bar c^i_j$ and $\bar c^i_{j+1}$. By assumption, there exists a tuple $\bar a' \equiv_{UB} \bar a$ such that every β^i_j is indiscernible over $U \cup \bar a'$. This implies that

$$\bar c^i_j \approx^{\mathrm{ls}}_{U\bar a'} \bar c^i_{j+1}\,.$$

Hence, $\bar b_i \equiv^{\mathrm{ls}}_{U\bar a'} \bar b'_i$, which implies that $\bar b_i \equiv_{U\bar a'} \bar b'_i$. Consequently, $\bar a'$ realises Φ_0. □

It follows that $\mathop{\sqrt{}}^{\mathrm{li}}$ is the coarsest forking relation that preserves indiscernibility.

Proposition 2.18. *Let $\sqrt{}$ be a forking relation. Then $\sqrt{} \subseteq \mathop{\sqrt{}}^{\mathrm{li}}$ if, and only if, whenever β is an indiscernible sequence over some set U and $A \sqrt{}_U \beta$, then β is indiscernible over $U \cup A$.*

Proof. ($\Rightarrow$) Suppose that $\sqrt{} \subseteq \sqrt[\mathrm{li}]{}$ and that $A \sqrt{}_U \beta$, for some indiscernible sequence β over U. Then $A \sqrt[\mathrm{li}]{}_U \beta$ and we can use Lemma 2.17 to find a set $A' \equiv_{U\beta} A$ such that β is indiscernible over $U \cup A'$. Since $A'\beta \equiv_U A\beta$, it follows that β is also indiscernible over $U \cup A$.

($\Leftarrow$) To show that $\sqrt{} \subseteq \sqrt[\mathrm{li}]{}$, suppose that $A \sqrt{}_U B$. We use the characterisation of Lemma 2.17 to prove that $A \sqrt[\mathrm{li}]{}_U B$. Hence, consider indiscernible sequences $\alpha_0, \dots, \alpha_{n-1}$ over U. By (EXT), there exists a set $A' \equiv_{UB} A$ such that

$$A' \sqrt{}_U B\alpha_0 \dots \alpha_{n-1}\,.$$

By assumption, $A' \sqrt{}_U \alpha_i$ implies that α_i is indiscernible over $U \cup A'$. $\square$

We also need the following technical lemma about the splitting relation $\sqrt[\mathrm{s}]{}$.

Lemma 2.19. *Let $\bar a \sqrt[\mathrm{s}]{}_U M$ where $\mathfrak{M}$ is a κ-saturated model and $U \subseteq M$ a set of size $|U| < \kappa$. For every set C, there exists a unique extension of* $\mathrm{tp}(\bar a/M)$ *over $M \cup C$ that is $\sqrt[\mathrm{s}]{}$-free over U.*

Proof. For uniqueness, suppose that there are two extension $\mathfrak{p}$ and $\mathfrak{p}'$ of $\mathrm{tp}(\bar a/M)$ over $C \supseteq M$ that are both $\sqrt[\mathrm{s}]{}$-free over U. Fix realisations $\bar b$ and $\bar b'$ of these two types and consider a finite tuple $\bar c \subseteq C$. Since $\mathfrak{M}$ is κ-saturated, we can find some tuple $\bar d \subseteq M$ with $\bar d \equiv_U \bar c$. Then

$$\bar b \sqrt[\mathrm{s}]{}_U C\,, \quad \bar b' \sqrt[\mathrm{s}]{}_U C\,, \quad \text{and} \quad \bar c \equiv_U \bar d$$

implies $\bar c \equiv_{U\bar b} \bar d$ and $\bar c \equiv_{U\bar b'} \bar d$. Furthermore,

$$\bar b \equiv_M \bar a \equiv_M \bar b' \quad \text{implies} \quad \bar b \equiv_{U\bar d} \bar b'\,.$$

Consequently,

$$\bar b\bar c \equiv_U \bar b\bar d \equiv_U \bar b'\bar d \equiv_U \bar b'\bar c\,.$$

Hence, $\bar{b} \equiv_{U\bar{c}} \bar{b}'$, for all finite $\bar{c} \subseteq C$, which implies that $\bar{b} \equiv_{UC} \bar{b}'$. Consequently, $\mathfrak{p} = \mathrm{tp}(\bar{b}/C) = \mathrm{tp}(\bar{b}'/C) = \mathfrak{p}'$.

It remains to prove the existence of a $\sqrt[s]{}$-free extension. As $\mathfrak{M}$ is κ-saturated, it realises every type over U. Hence, there exists a function $g : C^{<\omega} \to M^{<\omega}$ such that

$$g(\bar{c}) \equiv_U \bar{c}\,, \quad \text{for all } \bar{c} \in C^{<\omega}.$$

We claim that

$$\mathfrak{p} := \left\{ \varphi(\bar{x};\bar{c}) \;\middle|\; \varphi(\bar{x};\bar{y}) \text{ a formula, } \bar{c} \in C^{<\omega},\ \mathbb{M} \vDash \varphi(\bar{a}; g(\bar{c})) \right\}$$

is the desired type.

Let us start by showing that the set $\mathfrak{p}$ is satisfiable. Consider finitely many formulae $\varphi_0(\bar{x};\bar{c}_0), \dots, \varphi_n(\bar{x};\bar{c}_n) \in \mathfrak{p}$ and set $\bar{c} := \bar{c}_0 \dots \bar{c}_n$ and $\bar{d} := g(\bar{c}_0) \dots g(\bar{c}_n)$. By definition of $\mathfrak{p}$, we have

$$\mathbb{M} \vDash \varphi_0(\bar{a}; g(\bar{c}_0)) \wedge \dots \wedge \varphi_n(\bar{a}; g(\bar{c}_n))\,.$$

By κ^+-saturation of $\mathfrak{M}$, there exists a tuple $\bar{b} \subseteq M$ with $\bar{b} \equiv_U \bar{c}$. Then

$$g(\bar{c}) \equiv_U \bar{c} \equiv_U \bar{b} \quad \text{and} \quad \bar{a} \sqrt[s]{}_U M \quad \text{implies} \quad g(\bar{c}) \equiv_{U\bar{a}} \bar{b}\,.$$

Choosing some tuple $\bar{a}'$ such that $\bar{a}\bar{b} \equiv_U \bar{a}'\bar{c}$, it follows that

$$\bar{a} g(\bar{c}) \equiv_U \bar{a}\bar{b} \equiv_U \bar{a}'\bar{c}\,.$$

Suppose that $g(\bar{c}) = \bar{d}_0 \dots \bar{d}_n$. Then

$$\mathbb{M} \vDash \varphi_i(\bar{a}; g(\bar{c}_i)) \quad \text{and} \quad \bar{a} \sqrt[s]{}_U M \quad \text{implies} \quad \mathbb{M} \vDash \varphi_i(\bar{a}; \bar{d}_i)\,.$$

By choice of $\bar{a}'$, it follows that

$$\mathbb{M} \vDash \varphi_0(\bar{a}';\bar{c}_0) \wedge \dots \wedge \varphi_n(\bar{a}';\bar{c}_n)\,.$$

Thus, $\bar{a}'$ is the desired tuple satisfying every $\varphi_i(\bar{x};\bar{c}_i)$.

Furthermore, note that $\mathfrak{p}$ is a complete type over C since, for every formula $\varphi(\bar{x};\bar{c})$ with parameters $\bar{c} \subseteq C$, we have

$$\begin{aligned}\varphi(\bar{x};\bar{c}) \in \mathfrak{p} \quad &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a};g(\bar{c})) \\ &\text{iff} \quad \mathbb{M} \nvDash \neg\varphi(\bar{a};g(\bar{c})) \quad \text{iff} \quad \neg\varphi(\bar{x};\bar{c}) \notin \mathfrak{p}\,.\end{aligned}$$

To see that $\mathfrak{p}$ is $\sqrt[\mathrm{s}]{}$-free over U, consider two tuples $\bar{c}, \bar{c}' \subseteq C$ such that $\bar{c} \equiv_U \bar{c}'$. Then

$$g(\bar{c}) \equiv_U \bar{c} \equiv_U \bar{c}' \equiv_U g(\bar{c}') \quad \text{and} \quad \bar{a} \sqrt[\mathrm{s}]{}_U\, M$$

implies that $g(\bar{c}) \equiv_{U\bar{a}} g(\bar{c}')$. For a formula $\varphi(\bar{x};\bar{y})$ over U, it follows that

$$\begin{aligned}\varphi(\bar{x};\bar{c}) \in \mathfrak{p} \quad &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a};g(\bar{c})) \\ &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a};g(\bar{c}')) \quad \text{iff} \quad \varphi(\bar{x};\bar{c}') \in \mathfrak{p}\,. \qquad \square\end{aligned}$$

Proposition 2.20. *Let $\bar{a}, U \subseteq \mathbb{M}$ and let $\mathfrak{M}$ be a model containing U that is $(|T| \oplus |U|)^+$-saturated and strongly $(|T| \oplus |U|)^+$-homogeneous. The following statements are equivalent:*

(1) $\bar{a} \sqrt[\mathrm{li}]{}_U\, M$.

(2) $\bar{a} \sqrt[\mathrm{ls}]{}_U\, M$.

(3) $\bar{a} \sqrt[\mathrm{q}]{}_U\, M$.

(4) $\bar{b} \equiv^{\mathrm{ls}}_U \bar{b}' \Rightarrow \bar{b} \equiv^{\mathrm{ls}}_{U\bar{a}} \bar{b}'$ *for all finite* $\bar{b}, \bar{b}' \subseteq M$.

(5) $\bar{a} \sqrt[\mathrm{s}]{}_N\, M$, *for all models* $\mathfrak{N} \preceq \mathfrak{M}$ *containing* U.

(6) *For all models* $\mathfrak{N} \preceq \mathfrak{M}$ *containing* U*, we have*

$$\bar{b} \equiv_N \bar{b}' \Rightarrow \bar{b} \equiv_{U\bar{a}} \bar{b}'\,, \quad \text{for all } \bar{b}, \bar{b}' \subseteq M\,.$$

(7) $\mathrm{tp}(\bar{a}/M)$ *is invariant under all automorphisms of* $\mathfrak{M}$ *that fix some model* $\mathfrak{N} \preceq \mathfrak{M}$ *containing* U.

(8) *Every indiscernible sequence* $(\bar{b}_i)_{i<\omega}$ *over* U *that is contained in* M *is also indiscernible over* $U \cup \bar{a}$.

(9) *For every indiscernible sequence* $(\bar{b}_i)_{i<\omega}$ *over* U *with* $\bar{b}_0, \bar{b}_1 \subseteq M$, *we can find some indiscernible sequence* $(\bar{b}'_i)_{i<\omega}$ *over* $U \cup \bar{a}$ *with* $\bar{b}'_0 = \bar{b}_0$ *and* $\bar{b}'_1 = \bar{b}_1$.

(10) $\bar{b}_0 \equiv_{U\bar{a}} \bar{b}_1$, *for every indiscernible sequence* $(\bar{b}_i)_{i<\omega}$ *over* U *with* $\bar{b}_0, \bar{b}_1 \subseteq M$.

Proof. Set $\kappa := |T| \oplus |U|$.

(3) $\Leftrightarrow$ (9) was already proved in Lemma 2.16.

(3) $\Rightarrow$ (4) Consider two finite tuples $\bar{b}, \bar{b}' \subseteq M$ with $\bar{b} \equiv^{\mathrm{ls}}_U \bar{b}'$. By definition of $\equiv^{\mathrm{ls}}$, there are tuples $\bar{c}_0, \dots, \bar{c}_n$ such that $\bar{c}_0 = \bar{b}$, $\bar{c}_n = \bar{b}'$ and $\bar{c}_i \approx^{\mathrm{ls}}_U \bar{c}_{i+1}$, for all $i < n$. As $\mathfrak{M}$ is κ^+-saturated, we may assume that $\bar{c}_0, \dots, \bar{c}_n$ are contained in M. By (3), it follows that $\bar{c}_i \approx^{\mathrm{ls}}_{U\bar{a}} \bar{c}_{i+1}$, for all $i < n$. This implies that $\bar{b} \equiv^{\mathrm{ls}}_{U\bar{a}} \bar{b}'$.

(4) $\Rightarrow$ (7) Let $\pi \in \operatorname{Aut} \mathfrak{M}_N$, for some model $\mathfrak{N} \preceq \mathfrak{M}$ containing U. For every finite $\bar{b} \subseteq M$, it follows by Proposition 2.5 that

$$\begin{aligned} \bar{b} \equiv_N \pi(\bar{b}) \quad &\Rightarrow \quad \bar{b} \equiv^{\mathrm{ls}}_N \pi(\bar{b}) \\ &\Rightarrow \quad \bar{b} \equiv^{\mathrm{ls}}_{N\bar{a}} \pi(\bar{b}) \quad \Rightarrow \quad \bar{b} \equiv_{\bar{a}} \pi(\bar{b})\,. \end{aligned}$$

Consequently, for every formula $\varphi(\bar{x}; \bar{y})$,

$$\varphi(\bar{x}; \bar{b}) \in \operatorname{tp}(\bar{a}/M) \quad \text{iff} \quad \varphi(\bar{x}; \pi(\bar{b})) \in \operatorname{tp}(\bar{a}/M)\,.$$

(7) $\Rightarrow$ (2) Let $\bar{b}, \bar{b}' \subseteq M$ be tuples with $\bar{b} \equiv^{\mathrm{ls}}_U \bar{b}'$. First, we consider the case where $\bar{b}$ and $\bar{b}'$ are finite. By Proposition 2.10, there are tuples $\bar{c}_0, \dots, \bar{c}_n$ and models $N_0, \dots, N_{m-1} \supseteq U$ such that

$$\bar{a} = \bar{c}_0 \equiv_{N_0} \bar{c}_1 \equiv_{N_1} \cdots \equiv_{N_{n-2}} \bar{c}_{n-1} \equiv_{N_{n-1}} \bar{c}_n = \bar{b}\,.$$

Replacing each model $\mathfrak{N}_i$ by a suitable elementary substructure, we can ensure that $|N_i| = \kappa$. By κ^+-saturation of $\mathfrak{M}$, we may therefore assume that $N_i \subseteq M$. Hence, κ^+-homogeneity of $\mathfrak{M}$ implies that there are automorphisms $\pi_i \in \operatorname{Aut} \mathfrak{M}_{N_i}$ with $\pi_i(\bar{c}_i) = \bar{c}_{i+1}$. By (7) it follows that $\bar{c}_i \equiv_{N_i \bar{a}} \bar{c}_{i+1}$. Consequently, $\bar{b} \equiv_{U\bar{a}} \bar{b}'$.

For infinite tuples $\bar{b}, \bar{b}' \subseteq M$, it follows that

$$\begin{aligned} \bar{b} \equiv^{\mathrm{ls}}_U \bar{b}' \quad &\Rightarrow \quad \bar{b}|_I \equiv_{U\bar{a}} \bar{b}'|_I \,, \quad \text{for all finite sets of indices } I \\ &\Rightarrow \quad \bar{b} \equiv_{U\bar{a}} \bar{b}' \,. \end{aligned}$$

Consequently, $\bar{a} \mathop{\sqrt{}}^{\mathrm{ls}}_U M$.

(2) $\Rightarrow$ (5) Let $\mathfrak{N} \preceq \mathfrak{M}$ be a model containing U and consider two tuples $\bar{b}, \bar{b}' \subseteq M$ with $\bar{b} \equiv_N \bar{b}'$. Let $\bar{c}$ be an enumeration of N. By (2) and Proposition 2.5, it follows that

$$\begin{aligned} \bar{b} \equiv_N \bar{b}' \quad &\Rightarrow \quad \bar{b}\bar{c} \equiv_N \bar{b}'\bar{c} \\ &\Rightarrow \quad \bar{b}\bar{c} \equiv^{\mathrm{ls}}_N \bar{b}'\bar{c} \\ &\Rightarrow \quad \bar{b}\bar{c} \equiv^{\mathrm{ls}}_U \bar{b}'\bar{c} \\ &\Rightarrow \quad \bar{b}\bar{c} \equiv_{U\bar{a}} \bar{b}'\bar{c} \\ &\Rightarrow \quad \bar{b} \equiv_{U\bar{a}\bar{c}} \bar{b}' \quad \Rightarrow \quad \bar{b} \equiv_{N\bar{a}} \bar{b}' \,. \end{aligned}$$

(5) $\Rightarrow$ (6) is trivial.

(6) $\Rightarrow$ (10) Let $(\bar{b}_i)_{i<\omega}$ be an indiscernible sequence over U such that $\bar{b}_0, \bar{b}_1 \subseteq M$. We fix an arbitrary model $\mathfrak{N} \preceq \mathfrak{M}$ of size $|N| = \kappa$ containing U. By Lemma E5.3.11, there is some model $N' \equiv_U N$ such that $(\bar{b}_i)_{i<\omega}$ is indiscernible over N'. In particular, we have $\bar{b}_0 \equiv_{N'} \bar{b}_1$. By κ^+-saturation of $\mathfrak{M}$, we can find some set $N'' \subseteq M$ with $N'' \equiv_{U\bar{b}_0\bar{b}_1} N'$. Hence, $\bar{b}_0 \equiv_{N''} \bar{b}_1$ and (6) implies that $\bar{b}_0 \equiv_{U\bar{a}} \bar{b}_1$.

(10) $\Rightarrow$ (8) Let $(\bar{b}_i)_{i<\omega}$ be an indiscernible sequence over U that is contained in M. To show that $(\bar{b}_i)_{i<\omega}$ is indiscernible over $U \cup \bar{a}$, we will prove that

$$\bar{b}[\bar{\imath}] \equiv_{U\bar{a}} \bar{b}[\bar{k}] \,, \quad \text{for all } \bar{\imath}, \bar{k} \in [\omega]^n, \ n < \omega \,.$$

It is sufficient to consider the case where $\bar{\imath} < \bar{k}$. Hence, let $\bar{\imath} < \bar{k}$ be elements of $[\omega]^n$. Fix some increasing sequence $\bar{l}_0 < \bar{l}_1 < \dots$ in $[\omega]^n$

with $\bar{l}_0 = \bar{\imath}$ and $\bar{l}_1 = \bar{k}$. We set $\bar{c}_j := \bar{b}[\bar{l}_j]$. Then $(\bar{c}_j)_{j<\omega}$ is indiscernible over U and it follows by (10) that $\bar{b}[\bar{\imath}] = \bar{c}_0 \equiv_{U\bar{a}} \bar{c}_1 = \bar{b}[\bar{k}]$.

(8) ⇒ (9) Let $(\bar{b}^n)_{n<\omega}$ be an indiscernible sequence over U such that $\bar{b}^0, \bar{b}^1 \subseteq M$. We first consider the special case where the tuples $\bar{b}^n$ are finite. Since $\mathfrak{M}$ is κ^+-saturated, it contains some sequence $(\bar{b}'_i)_{i<\omega}$ with $\bar{b}'[\omega] \equiv_{U\bar{b}_0\bar{b}_1} \bar{b}[\omega]$. Then $\bar{b}'_0 = \bar{b}_0$, $\bar{b}'_1 = \bar{b}_1$ and it follows by (8) that $(\bar{b}'_i)_{i<\omega}$ is indiscernible over $U \cup \bar{a}$.

For the general case, let $\Phi((\bar{x}^n)_{n<\omega})$ be a set of formulae stating that the sequence $(\bar{x}^n)_{n<\omega}$ is indiscernible over $U \cup \bar{a}$ and that $\bar{x}^0 = \bar{b}^0$ and $\bar{x}^1 = \bar{b}^1$. We have to show that Φ is satisfiable. Thus, consider a finite subset $\Phi_0 \subseteq \Phi$. Then there is a finite set I of indices such that the formulae in Φ_0 only contain variables x^n_i with $i \in I$. Applying the special case we have proved above to the sequence $(\bar{b}^n|_I)_{n<\omega}$, we obtain an indiscernible sequence $(\bar{b}'_n)_{n<\omega}$ over $U \cup \bar{a}$ with $\bar{b}'_0 = \bar{b}^0$ and $\bar{b}'_1 = \bar{b}^1$. This sequence satisfies Φ_0.

(1) ⇒ (2) follows since $\mathop{\smile\hskip-0.9em\mid}^{\mathrm{li}} = {}^*(\mathop{\smile\hskip-0.9em\mid}^{\mathrm{ls}}) \subseteq \mathop{\smile\hskip-0.9em\mid}^{\mathrm{ls}}$.

(5) ⇒ (1) Fix some set $C \subseteq \mathbb{M}$. We have to show that there is some tuple $\bar{a}' \equiv_M \bar{a}$ with $\bar{a}' \mathop{\smile\hskip-0.9em\mid}^{\mathrm{ls}}_U MC$. Let $\mathfrak{N} \preceq \mathfrak{M}$ be a model containing U of size $|N| = \kappa$. Then $\bar{a} \mathop{\smile\hskip-0.9em\mid}^{\mathrm{s}}_N M$ and we can use Lemma 2.19 to find some tuple $\bar{a}_N \equiv_M \bar{a}$ such that $\bar{a}_N \mathop{\smile\hskip-0.9em\mid}^{\mathrm{s}}_N MC$ and $\mathrm{tp}(\bar{a}_N/MC)$ is the unique $\mathop{\smile\hskip-0.9em\mid}^{\mathrm{s}}$-free extension of $\mathrm{tp}(\bar{a}/M)$. Furthermore, if we are given two such models $\mathfrak{N}, \mathfrak{N}' \preceq \mathfrak{M}$, we can find some model $\mathfrak{N}^+ \preceq \mathfrak{M}$ containing $N \cup N'$ of size $|N^+| = \kappa$. Then

$$\bar{a}_N \mathop{\smile\hskip-0.9em\mid}^{\mathrm{s}}_{N^+} MC\,, \quad \bar{a}_{N'} \mathop{\smile\hskip-0.9em\mid}^{\mathrm{s}}_{N^+} MC\,, \quad \text{and} \quad \bar{a}_N \equiv_M \bar{a}_{N'}\,,$$

and it follows by uniqueness that $\bar{a}_N \equiv_{MC} \bar{a}_{N'}$. Consequently, choosing $\bar{a}' := \bar{a}_{N_0}$, for an arbitrary model $\mathfrak{N}_0$, we have

$$\bar{a}' \equiv_M \bar{a} \quad \text{and} \quad \bar{a}' \mathop{\smile\hskip-0.9em\mid}^{\mathrm{s}}_N MC\,, \quad \text{for all models } U \subseteq N \subseteq M \text{ of size } |N| = \kappa\,.$$

We claim that $\bar{a}' \mathop{\smile\hskip-0.9em\mid}^{\mathrm{ls}}_N MC$. Consider two tuples $\bar{b}, \bar{b}' \subseteq MC$ with $\bar{b} \approx^{\mathrm{ls}}_U \bar{b}'$. By Lemma 2.3, there is some model $N \supseteq U$ with $\bar{b} \equiv_N \bar{b}'$. We can choose N of size $|N| = \kappa$ and, by κ^+-saturation of $\mathfrak{M}$, we may assume that $N \subseteq M$. Consequently,

$$\bar{a}' \overset{\mathrm{s}}{\surd}_N MC \quad \text{implies} \quad \bar{b} \equiv_{N\bar{a}'} \bar{b}' ,$$

as desired. □

Corollary 2.21. $\overset{\mathrm{li}}{\surd} = {}^*(\overset{\mathrm{q}}{\surd})$ *is a forking relation.*

Proof. We have seen in Lemma 2.15 that $\overset{\mathrm{q}}{\surd}$ is a preforking relation. Consequently, ${}^*(\overset{\mathrm{q}}{\surd})$ is a forking relation and it remains to prove that it coincides with $\overset{\mathrm{li}}{\surd}$. The inclusion $\overset{\mathrm{ls}}{\surd} \subseteq \overset{\mathrm{q}}{\surd}$ follows immediately from the respective definitions. Consequently, $\overset{\mathrm{li}}{\surd} = {}^*(\overset{\mathrm{ls}}{\surd}) \subseteq {}^*(\overset{\mathrm{q}}{\surd})$. Conversely, by the implication (3) ⇒ (1) of Proposition 2.20, we have

$$A \ {}^*(\overset{\mathrm{q}}{\surd})_U \ M \quad \text{implies} \quad A \overset{\mathrm{li}}{\surd}_U M ,$$

for sufficiently saturated models $\mathfrak{M}$. According to Lemma F2.4.7, this implies that ${}^*(\overset{\mathrm{q}}{\surd}) \subseteq \overset{\mathrm{li}}{\surd}$. □

Corollary 2.22. $\overset{\mathrm{s}}{\surd} \subseteq \overset{\mathrm{ls}}{\surd} \subseteq \overset{\mathrm{q}}{\surd}$ *and* $\overset{\mathrm{i}}{\surd} \subseteq \overset{\mathrm{li}}{\surd} \subseteq \overset{\mathrm{f}}{\surd}$

Proof. The first two inclusions follow immediately from the respective definitions. For the thrid one, it follows that

$$\overset{\mathrm{i}}{\surd} = {}^*(\overset{\mathrm{s}}{\surd}) \subseteq {}^*(\overset{\mathrm{ls}}{\surd}) = \overset{\mathrm{li}}{\surd} .$$

For the last inclusion, it is sufficient to prove that

$$A \overset{\mathrm{li}}{\surd}_U M \quad \text{implies} \quad A \overset{\mathrm{d}}{\surd}_U M ,$$

for every sufficiently saturated model $\mathfrak{M}$, since Lemma F2.4.7 then implies that $\overset{\mathrm{li}}{\surd} = {}^*(\overset{\mathrm{li}}{\surd}) \subseteq {}^*(\overset{\mathrm{d}}{\surd}) = \overset{\mathrm{f}}{\surd}$.

Hence, suppose that $A \overset{\mathrm{li}}{\surd}_U M$ where $\mathfrak{M}$ is a $(|T| \oplus |U|)^+$-saturated and strongly $(|T| \oplus |U|)^+$-homogeneous model containing U. By finite character it is sufficient to show that $A \overset{\mathrm{d}}{\surd}_U B$, for every finite subset $B \subseteq M$.

Hence, let $B \subseteq M$ be finite, and consider an indiscernible sequence $(\bar{b}'_i)_{i<\omega}$ over U where $\bar{b}'_0$ is an enumeration of B. By $(|T| \oplus |U|)^+$-saturation of $\mathfrak{M}$, we can find an indiscernible sequence $(\bar{b}_i)_{i<\omega}$ over U such that $\bar{b}[\omega] \subseteq M$ and $\bar{b}[\omega] \equiv_{U\bar{b}'_0} \bar{b}'[\omega]$. By Proposition 2.20 (8), this sequence is indiscernible over $U \cup A$. Let A' be some set such that

$$A\bar{b}[\omega] \equiv_{U\bar{b}'_0} A'\bar{b}'[\omega] .$$

Then $(\bar{b}'_i)_{i<\omega}$ is indiscernible over $U \cup A'$ and it follows by Lemma F3.1.3 that $A \overset{\mathrm{d}}{\surd}_U \bar{b}'_0$. □

It the remainder of this section we compare the relations $\overset{\mathrm{li}}{\surd}$ and $\overset{\mathrm{f}}{\surd}$.

Definition 2.23. We call an independence relation $\surd$ *weakly bounded* if, there exists a function $f : \mathrm{Cn} \to \mathrm{Cn}$ such that

$$\mathrm{mult}_{\surd}(\mathfrak{p}) \leq f\big(|T| \oplus |U|\big) , \quad \text{for all } \mathfrak{p} \in S^{<\omega}(U) .$$

In this case we also say that $\surd$ is weakly bounded *by* f.

We can characterise $\overset{\mathrm{li}}{\surd}$ as the coarsest weakly bounded forking relation.

Proposition 2.24.

(a) $\overset{\mathrm{li}}{\surd}$ *is weakly bounded by* $f(\kappa) = 2^{2^\kappa}$.

(b) $\surd \subseteq \overset{\mathrm{li}}{\surd}$, *for every weakly bounded forking relation* $\surd$.

Proof. (a) Fix a type $\mathfrak{p} \in S^{<\omega}(U)$ and some set $C \supseteq U$. We have to show that $\mathfrak{p}$ has at most $\kappa := 2^{2^{|T| \oplus |U|}}$ $\overset{\mathrm{li}}{\surd}$-free extensions over C. For $\mathfrak{q} \in S^{<\omega}(C)$, let $g_{\mathfrak{q}}$ be the function mapping a formula $\varphi(\bar{x};\bar{y})$ over U to the set

$$g_{\mathfrak{q}}(\varphi) := \big\{ [\bar{b}]_{\equiv^{\mathrm{ls}}_U} \bigm| \varphi(\bar{x};\bar{b}) \in \mathfrak{q} \big\} .$$

We claim that $g_{\mathfrak{q}} = g_{\mathfrak{q}'}$ implies $\mathfrak{q} = \mathfrak{q}'$.

For the proof, suppose that $g_{\mathfrak{q}} = g_{\mathfrak{q}'}$ and let $\varphi(\bar{x};\bar{b}) \in \mathfrak{q}$. Then $[\bar{b}]_{\equiv^{\mathrm{ls}}_U} \in g_{\mathfrak{q}} = g_{\mathfrak{q}'}$ implies that there is some tuple $\bar{b}' \equiv^{\mathrm{ls}}_U \bar{b}$ with $\varphi(\bar{x};\bar{b}') \in \mathfrak{q}'$. Fix a tuple $\bar{a}'$ realising $\mathfrak{q}'$. Then $\bar{a}' \mathop{\sqrt{}}^{\mathrm{ls}}_U C$ and

$$\bar{b} \equiv^{\mathrm{ls}}_U \bar{b}' \quad \text{implies} \quad \mathbb{M} \vDash \varphi(\bar{a}';\bar{b}) \leftrightarrow \varphi(\bar{a}';\bar{b}')\,.$$

Consequently, $\varphi(\bar{x};\bar{b}) \in \mathfrak{q}'$, as desired.

To conclude the proof, let $N \supseteq U$ be a model of size $|T| \oplus |U|$. Note that the number of $\equiv_N$-classes of finite tuples is at most $|S^{<\omega}(N)| = 2^{|N|}$. By Proposition 2.5, it follows that there are also at most that many $\equiv^{\mathrm{ls}}_U$-equivalence classes of finite tuples. Hence, there are at most $2^{2^{|N|}} = \kappa$ functions of the form $g_{\mathfrak{q}}$. It follows that there are at most κ $\mathop{\sqrt{}}^{\mathrm{li}}$-free extensions of $\mathfrak{p}$ over C.

(b) For a contradiction, suppose that there is a weakly bounded forking relation $\sqrt{}$ with $\sqrt{} \nsubseteq \mathop{\sqrt{}}^{\mathrm{li}}$. Then there are $\bar{a}, B, U \subseteq \mathbb{M}$ such that

$$\bar{a} \sqrt{}_U B \quad \text{and} \quad \bar{a} \mathop{\not\sqrt{}}^{\mathrm{li}}_U B\,.$$

Let $f : \mathrm{Cn} \to \mathrm{Cn}$ be the function bounding $\sqrt{}$ and let $M \supseteq U \cup B$ be a model that is $(|T| \oplus |U|)^+$-saturated and strongly $(|T| \oplus |U|)^+$-homogeneous. By (EXT), we can find some tuple $\bar{a}' \equiv_{UB} \bar{a}$ with $\bar{a}' \sqrt{}_U M$. By (MON), we have $\bar{a}' \mathop{\not\sqrt{}}^{\mathrm{li}}_U M$. Hence, we can use Proposition 2.20 (10) to find an indiscernible sequence $(\bar{b}_i)_{i<\omega}$ over U with $\bar{b}_0, \bar{b}_1 \subseteq M$ such that $\bar{b}_0 \not\equiv_{U\bar{a}'} \bar{b}_1$. Fix some formula $\varphi(\bar{x};\bar{y})$ such that

$$\mathbb{M} \vDash \neg\varphi(\bar{a}';\bar{b}_0) \wedge \varphi(\bar{a}';\bar{b}_1)\,.$$

Let $I \subseteq \omega$ be an infinite set of indices such that

$$\mathbb{M} \vDash \varphi(\bar{a}';\bar{b}_i) \leftrightarrow \varphi(\bar{a}';\bar{b}_k) \quad \text{for all } i, k \in I\,,$$

and let $(\bar{c}_j)_{j\in J}$ be an extension of $(\bar{b}_i)_{i\in I\cup\{0,1\}}$ of size $|J| > f(|T| \oplus |U|)$ that is indiscernible over U and such that the order J is strongly $\aleph_0$-homogeneous. Fix a tuple $\bar{a}'' \equiv_{UM} \bar{a}'$ with $\bar{a}'' \sqrt{}_U M\bar{c}[J]$. For every

$j \in J$, fix an order automorphism $\sigma_j : J \to J$ such that $\sigma_j(0) = j$ and let $\pi_j \in \operatorname{Aut} \mathbb{M}_U$ be an automorphism with

$$\pi_j(\bar{c}_k) = \bar{c}_{\sigma_j(k)}\,, \quad \text{for all } k \in J\,.$$

Setting $\bar{a}_j := \pi_j(\bar{a}'')$ it follows by invariance that

$$\bar{a}_j \sqrt{}_U\, \bar{c}[J] \quad \text{and} \quad \bar{a}_j \not\equiv_{U\bar{c}[J]} \bar{a}_k\,, \quad \text{for } j \neq k\,.$$

Hence, $\operatorname{mult}_{\sqrt{}}(\operatorname{tp}(\bar{a}/U)) \geq |J| > f(|T| \oplus |U|)$. A contradiction. ☐

Corollary 2.25. *Let T be a complete first-order theory. The following statements are equivalent.*

(1) $\sqrt[\mathrm{f}]{} = \sqrt[\mathrm{li}]{}$.

(2) $\sqrt[\mathrm{f}]{}$ *is weakly bounded.*

(3) *If β is an indiscernible sequence over some set U and $A \sqrt[\mathrm{f}]{}_U\, \beta$, then β is indiscernible over $U \cup A$.*

Proof. (1) ⇒ (2) follows by Proposition 2.24 (a).

(2) ⇒ (1) The inclusion $\sqrt[\mathrm{li}]{} \subseteq \sqrt[\mathrm{f}]{}$ follows by Corollary 2.22, while $\sqrt[\mathrm{f}]{} \subseteq \sqrt[\mathrm{li}]{}$ follows by Proposition 2.24 (b).

(1) ⇒ (3) follows by Proposition 2.18.

(3) ⇒ (1) The inclusion $\sqrt[\mathrm{li}]{} \subseteq \sqrt[\mathrm{f}]{}$ follows by Corollary 2.22, while $\sqrt[\mathrm{f}]{} \subseteq \sqrt[\mathrm{li}]{}$ follows by Proposition 2.18. ☐

Theorem 2.26. *If a theory T does not have the independence property, then $\sqrt[\mathrm{li}]{} = \sqrt[\mathrm{f}]{}$.*

Proof. The inclusion $\sqrt[\mathrm{li}]{} \subseteq \sqrt[\mathrm{f}]{}$ was proved in Corollary 2.22. For the converse, it is sufficient, by Lemma F2.4.7, to prove that

$$\bar{a} \sqrt[\mathrm{f}]{}_U\, M \quad \text{implies} \quad \bar{a} \sqrt[\mathrm{li}]{}_U\, M\,,$$

for all models $\mathfrak{M}$ that are $(|T| \oplus |U|)^+$-saturated and strongly $(|T| \oplus |U|)^+$-homogeneous.

Hence, let $\bar{a} \sqrt[f]{}_U M$. We check condition (10) of Proposition 2.20. Let $(\bar{b}_i)_{i<\omega}$ be an indiscernible sequence over U with $\bar{b}_0, \bar{b}_1 \subseteq M$. Then $\bar{a} \sqrt[f]{}_U M$ implies that $\bar{a} \sqrt[d]{}_U \bar{b}_0 \bar{b}_1$. By Lemma F3.1.3, there exists a tuple $\bar{a}' \equiv_{U\bar{b}_0\bar{b}_1} \bar{a}$ such that the sequence $(\bar{b}_{2i}\bar{b}_{2i+1})_{i<\omega}$ is indiscernible over $U \cup \bar{a}'$. For a contradiction, suppose that $\bar{b}_0 \not\equiv_{U\bar{a}} \bar{b}_1$. Then $\bar{b}_0 \not\equiv_{U\bar{a}'} \bar{b}_1$ and there is some formula $\varphi(\bar{x})$ over $U \cup \bar{a}'$ such that

$$\mathbb{M} \vDash \varphi(\bar{b}_0) \wedge \neg\varphi(\bar{b}_1)\,.$$

By indiscernibility of $(\bar{b}_{2i}\bar{b}_{2i+1})_{i<\omega}$ over $U \cup \bar{a}'$, it follows that

$$\mathbb{M} \vDash \varphi(\bar{b}_i) \quad \text{iff} \quad i \text{ is even.}$$

Hence, Proposition E5.4.2 implies that T has the independence property. A contradiction. □

Proposition 2.27. *A simple theory T does not have the independence property if, and only if, $\sqrt[li]{} = \sqrt[f]{}$.*

Proof. (⇒) follows by Theorem 2.26.

(⇐) Suppose that T is a simple theory with the independence property. We have to show that $\sqrt[li]{} \neq \sqrt[f]{}$. We can use Proposition E5.4.2 to find an indiscernible sequence $(\bar{a}_n)_{n<\omega}$ and a formula $\varphi(\bar{x};\bar{b})$ with parameters $\bar{b} \subseteq \mathbb{M}$ such that

$$\mathbb{M} \vDash \varphi(\bar{a}_n;\bar{b}) \quad \text{iff} \quad n \text{ is even.}$$

Using Proposition E5.3.6 we fix an indiscernible sequence $(\bar{a}'_n\bar{a}''_n)_{n<\omega+\omega}$ over $\bar{b}$ with

$$\mathrm{Av}((\bar{a}'_n\bar{a}''_n)_{n<\omega+\omega}/\bar{b}) \supseteq \mathrm{Av}((\bar{a}_{2n}\bar{a}_{2n+1})_{n<\omega}/\bar{b}).$$

Note that this implies that the interleaved sequence $\bar{a}'_0, \bar{a}''_0, \bar{a}'_1, \bar{a}''_1, \dots$ is indiscernible. In particular, we have

$$\bar{a}'_\omega \approx^{\mathrm{ls}}_U \bar{a}''_\omega \quad \text{where} \quad U := \bar{a}'[<\omega]\bar{a}''[<\omega]\,.$$

Let $A := \bar a'[<\omega + \omega]\bar a''[<\omega + \omega]$. Indiscernibility implies that $A \overset{\mathrm{u}}{\downarrow}_U \bar b$. Since $\overset{\mathrm{u}}{\downarrow} \subseteq \overset{\mathrm{f}}{\downarrow}$, it follows that $A \overset{\mathrm{f}}{\downarrow}_U \bar b$ and, by symmetry, $\bar b \overset{\mathrm{f}}{\downarrow}_U A$. But

$$\bar a'_\omega \not\equiv_{\bar b} \bar a''_\omega \quad \text{implies} \quad \bar a'_\omega \not\approx^{\mathrm{ls}}_{U\bar b} \bar a''_\omega\,.$$

Hence, $\bar b \overset{\mathrm{q}}{\not\downarrow}_U A$, which implies that $\bar b \overset{\mathrm{li}}{\not\downarrow}_U A$. Consequently, $\overset{\mathrm{f}}{\downarrow} \neq \overset{\mathrm{li}}{\downarrow}$. □

Theorem 2.28. *Let T be a complete first-order theory. The following statements are equivalent:*

(1) *T is stable.*

(2) *T is simple and it does not have the independence property.*

(3) *T is simple and $\overset{\mathrm{li}}{\downarrow} = \overset{\mathrm{f}}{\downarrow}$.*

(4) *$\overset{\mathrm{li}}{\downarrow}$ is symmetric.*

(5) *$\overset{\mathrm{li}}{\downarrow}$ is right local.*

Proof. (2) ⇔ (3) was already proved in Proposition 2.27.

(1) ⇒ (2) If T is stable, it is simple by Corollary F3.2.19 and it does not have the independence property by Proposition E5.4.11.

(2) ⇒ (1) Let T be a simple theory without the independence property. We have shown in Proposition F3.2.21 that T also does not have the strict order property. Consequently, it follows by Proposition E5.4.11 that T is stable.

(3) ⇒ (4) If T is simple, $\overset{\mathrm{f}}{\downarrow}$ is symmetric. Hence, so is $\overset{\mathrm{li}}{\downarrow} = \overset{\mathrm{f}}{\downarrow}$.

(4) ⇒ (5) Since $\overset{\mathrm{li}}{\downarrow}$ is a forking relation, this implication follows by Theorem F2.4.17.

(5) ⇒ (3) If $\overset{\mathrm{li}}{\downarrow}$ is right local, so is $\overset{\mathrm{f}}{\downarrow} \supseteq \overset{\mathrm{li}}{\downarrow}$. Consequently, T is simple. Furthermore, Theorem F2.4.17 implies that $\overset{\mathrm{li}}{\downarrow}$ is symmetric. Therefore, it follows by Theorem F3.1.9 that $\overset{\mathrm{f}}{\downarrow} \subseteq \overset{\mathrm{d}}{\downarrow} \subseteq \overset{\mathrm{li}}{\downarrow}$. □

3. $\sqrt[i]{}$-*Morley sequences*

In this section we study $\sqrt[i]{}$-Morley sequences in theories without the independence property.

Cofinal types

We start by noting that finiteness of the alternation number can be used to define a kind of 'limit type' of a sequences.

Definition 3.1. The *cofinal type* of a sequence $\alpha = (\bar{a}_i)_{i \in I}$ is the set

$$\mathrm{CF}(\alpha) := \big\{ \varphi(\bar{x}) \bigm| \varphi \text{ a formula over } \mathbb{M} \text{ such that } [\![\varphi(\bar{a}_i)]\!]_{i \in I} \text{ is cofinal in } I \big\}\,.$$

Lemma 3.2. *Let T be a theory without the independence property and let α be an indiscernible sequence. Then* $\mathrm{CF}(\alpha)$ *is a complete type over* $\mathbb{M}$ *which is finitely satisfiable in α.*

Proof. Suppose that $\alpha = (\bar{a}_i)_{i \in I}$. For completeness, consider a formula $\varphi(\bar{x})$ over $\mathbb{M}$. Since $\mathrm{alt}_\varphi(\alpha) < \infty$, there exists some index $k \in I$ such that

$$\mathbb{M} \vDash \varphi(\bar{a}_i) \leftrightarrow \varphi(\bar{a}_j)\,, \quad \text{for all } i, j \geq k\,.$$

Consequently,

$$\varphi \in \mathrm{CF}(\alpha) \quad \text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}_k) \quad \text{iff} \quad \neg\varphi \notin \mathrm{CF}(\alpha)\,.$$

To show that $\mathrm{CF}(\alpha)$ is consistent, consider finitely many formulae $\varphi_0, \dots, \varphi_n \in \mathrm{CF}(\alpha)$. There exists some index $k \in I$ such that

$$\mathbb{M} \vDash \varphi_j(\bar{a}_i)\,, \quad \text{for all } i \geq k \text{ and all } j \leq n\,.$$

In particular,

$$\mathbb{M} \vDash \varphi_0(\bar{a}_k) \wedge \dots \wedge \varphi_n(\bar{a}_k)\,.$$

Hence, $\{\varphi_0, \dots, \varphi_n\}$ is satisfiable. As the tuple satisfying this set belongs to α, it further follows that $\mathrm{CF}(\alpha)$ is finitely satisfiable in α. □

Cofinal types can be used to construct $\overset{\text{i}}{\surd}$-Morley sequences as follows.

Lemma 3.3. *Let T be a theory without the independence property and $\alpha = (\bar{a}_i)_{i \in I}$ an indiscernible sequence over U where the order I has no first element. Let $\alpha^{\mathrm{op}} := (\bar{a}_i)_{i \in I^{\mathrm{op}}}$ be the sequence with reverse ordering and let $\beta = (\bar{b}_j)_{j \in J}$ be generated by* $\mathrm{CF}(\alpha^{\mathrm{op}})$ *over $UC\alpha$.*

(a) *β is a $\overset{\text{i}}{\surd}$-Morley sequence over $UC\alpha$.*

(b) *$\beta\alpha$ is indiscernible over U.*

Proof. We start by proving that, for every formula φ over $UC\alpha$ and every tuple $\bar{j} \in [J]^n$, there are arbitrarily small indices $\bar{\imath} \in [I]^n$ such that

$$\mathbb{M} \vDash \varphi(\bar{b}[\bar{j}]) \leftrightarrow \varphi(\bar{a}[\bar{\imath}])\,.$$

We proceed by induction on n. For $n = 0$ there is nothing to do. Hence, suppose that we have proved the claim already for $n < \omega$ and that

$$\mathbb{M} \vDash \varphi(\bar{b}[\bar{j}], \bar{b}_l)\,,$$

where $\bar{j} \in [J]^n$ and $l \in J$ are indices with $\bar{j} < l$. Since $\bar{b}_l$ realises the type $\mathrm{CF}(\alpha^{\mathrm{op}}) \upharpoonright UC\alpha\bar{b}[<l]$, we have $\varphi(\bar{b}[\bar{j}], \bar{x}) \in \mathrm{CF}(\alpha^{\mathrm{op}})$. Consequently, there are arbitrarily small $k \in I$ such that

$$\mathbb{M} \vDash \varphi(\bar{b}[\bar{j}], \bar{a}_k)\,.$$

By inductive hypothesis, we can find arbitrarily small $\bar{\imath} < k$ such that

$$\mathbb{M} \vDash \varphi(\bar{a}[\bar{\imath}], \bar{a}_k)\,.$$

Having proved the claim, it follows by Corollary E5.4.3 that

$$\mathbb{M} \vDash \varphi(\bar{b}[\bar{j}]) \leftrightarrow \varphi(\bar{b}[\bar{j}'])\,, \quad \text{for all formulae } \varphi \text{ over } UC\alpha \text{ and all indices } \bar{j}, \bar{j}' \in [J]^n.$$

Hence, β is indiscernible over $UC\alpha$. As α is indiscernible over U, it further follows that

$$\mathbb{M} \vDash \varphi(\bar{b}[\bar{\jmath}], \bar{a}[\bar{k}]) \leftrightarrow \varphi(\bar{a}[\bar{\imath}], \bar{a}[\bar{k}]),$$

for all formulae φ over U and all indices $\bar{\imath} \in [I]^n$, $\bar{k} \in [I]^m$, $\bar{\jmath} \in [J]^n$ with $\bar{\imath} < \bar{k}$. This implies that $\beta\alpha$ is indiscernible over U.

To show that β is a $\sqrt[\mathrm{i}]{}$-Morley sequence, it remains to prove that

$$\bar{b}_j \sqrt[\mathrm{i}]{}_{UC\alpha}\ \bar{b}[<j], \quad \text{for all } j \in J.$$

We have shown in Lemma 3.2 that $\mathrm{CF}(\alpha^{\mathrm{op}})$ is a global type that is finitely satisfiable in α. In particular, it is invariant over $UC\alpha$. Hence, the type $\mathrm{CF}(\alpha^{\mathrm{op}}) \restriction UC\alpha\bar{b}[<j]$ realised by $\bar{b}_j$ has a global extension $\mathrm{CF}(\alpha^{\mathrm{op}})$ that is invariant over $UC\alpha$. □

As a concluding remark let us note that being generated by a type $\mathfrak{p}$ only depends on the average type of the sequence.

Lemma 3.4. *Let $\alpha = (\bar{a}_i)_{i\in I}$ and $\beta = (\bar{a}_j)_{j\in J}$ be infinite indiscernible sequences over U and $\mathfrak{p} \in S^{\bar{s}}(U\alpha\beta)$ a type that is invariant over U.*

(a) *If α is generated by $\mathfrak{p}$ over U and* $\mathrm{Av}(\alpha/U) = \mathrm{Av}(\beta/U)$*, then β is also generated by $\mathfrak{p}$ over U.*

(b) *If α and β are generated by $\mathfrak{p}$ over U, then* $\mathrm{Av}(\alpha/U) = \mathrm{Av}(\beta/U)$.

Proof. (a) Let $\varphi(\bar{x}; \bar{y})$ be a formula over U such that $\mathbb{M} \vDash \varphi(\bar{b}_j; \bar{b}[\bar{k}])$, for some $\bar{k} < j$ in J. Let $\bar{l}i$ be a tuple in I with the same order type as $\bar{k}j$. Then $\mathrm{Av}(\alpha/U) = \mathrm{Av}(\beta/U)$ implies that $\mathbb{M} \vDash \varphi(\bar{a}_i; \bar{a}[\bar{l}])$. Consequently, $\varphi(\bar{x}; \bar{a}[\bar{l}]) \in \mathfrak{p} \restriction U\bar{a}[<i]$. Since $\bar{a}[\bar{l}] \equiv_U \bar{b}[\bar{k}]$, it follows by invariance of $\mathfrak{p}$ that $\varphi(\bar{x}; \bar{b}[\bar{k}]) \in \mathfrak{p}$.

(b) We prove by induction on n that

$$\bar{a}[\bar{\imath}] \equiv_U \bar{b}[\bar{\jmath}], \quad \text{for all } \bar{\imath} \in [I]^n \text{ and } \bar{\jmath} \in [J]^n.$$

For $n = 0$, there is nothing to do. Hence, suppose that we have proved the claim already for tuples of length n and consider tuples $\bar{\imath} \in [I]^{n+1}$ and

$\bar{\jmath} \in [J]^{n+1}$. Set $\bar{\imath}' := i_0 \dots i_{n-1}$ and $\bar{\jmath}' := j_0 \dots j_{n-1}$ and let $\varphi(\bar{x}_0, \dots, \bar{x}_n)$ be a formula over U. By inductive hypothesis and invariance of $\mathfrak{p}$, it follows that

$$\begin{aligned} \mathbb{M} \vDash \varphi(\bar{a}[\bar{\imath}'], \bar{a}_{i_n}) \quad &\text{iff} \quad \varphi(\bar{a}[\bar{\imath}'], \bar{x}) \in \mathfrak{p} \\ &\text{iff} \quad \varphi(\bar{b}[\bar{\jmath}'], \bar{x}) \in \mathfrak{p} \\ &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{b}[\bar{\jmath}'], \bar{b}_{j_n}) . \end{aligned}$$

□

The confluence property

Our next aim is to prove a combinatorial characterisation of $\sqrt[i]{}$-Morley sequences in terms of the so-called *confluence property.*

Definition 3.5. Let U be a set of parameters.

(a) Let $\alpha = (\alpha_k)_{k \in K}$ be a family of indiscernible sequences $\alpha_k = (\bar{a}^k_i)_{i \in I_k}$ over U. We say that α is *confluent over* U if there exists some tuple $\bar{c}$ such that, for every $k \in K$, the extended sequence $\alpha_k \bar{c}$ is still indiscernible over U.

(b) A complete type $\Phi((\bar{x}_i)_{i<\omega})$ over U has the *confluence property* if every family $\alpha = (\alpha_k)_{k \in K}$ of indiscernible sequences $\alpha_k = (\bar{a}^k_i)_{i \in I_k}$ over U with

$$\mathrm{Av}(\alpha_k / U) = \Phi , \quad \text{for all } k \in K ,$$

is confluent over U.

(c) We say that a sequence $\alpha = (\bar{a}_i)_{i \in I}$ has the *confluence property* over a set U if it is indiscernible over U and $\mathrm{Av}(\alpha/U)$ has the confluence property.

We start by showing how to find sequences with the confluence property.

Lemma 3.6. *Every infinite sequence* $\alpha = (\bar{a}_i)_{i \in I}$ *such that*

$$\bar{a}_j \equiv_{U \bar{a}[<i]} \bar{a}_i \quad \textit{and} \quad \bar{a}_i \sqrt[i]{}_U \, \bar{a}[<i] , \quad \textit{for all } i \leq j \textit{ in } I ,$$

has the confluence property over U.

Proof. Indiscernibility follows by Lemma F2.4.14. For the confluence property of $\mathrm{Av}(\alpha/U)$, we choose a $(|T| \oplus |U|)^+$-saturated model $\mathfrak{M}$ of T containing U and we use Proposition E5.3.6 to find an indiscernible sequence $\alpha' = (\bar{a}'_n)_{n<\omega}$ over U of length ω with $\mathrm{Av}(\alpha'/U) = \mathrm{Av}(\alpha/U)$. By invariance of $\sqrt[\mathrm{i}]{}$, we have

$$\bar{a}'_n \mathrel{\sqrt[\mathrm{i}]{}}_U \bar{a}'[<n]\,, \quad \text{for all } n < \omega\,.$$

Since $\sqrt[\mathrm{i}]{}$ is a forking relation, we can choose, by induction on $n < \omega$, tuples

$$\bar{b}_n \equiv_{U\bar{a}'[<n]} \bar{a}'_n \quad \text{such that} \quad \bar{b}_n \mathrel{\sqrt[\mathrm{i}]{}}_U M\bar{a}'[<n]\bar{b}[<n]\,.$$

By Lemma F2.4.14, we have $(\bar{b}_n)_{n<\omega} \equiv_U (\bar{a}'_n)_{n<\omega}$. Hence, $\beta = (\bar{b}_n)_{n<\omega}$ is an indiscernible sequence over U with

$$\mathrm{Av}(\beta/U) = \mathrm{Av}(\alpha'/U) = \mathrm{Av}(\alpha/U)\,.$$

To show that this average type has the confluence property over U, consider a family of indiscernible sequences $\beta_k = (\bar{b}^k_i)_{i\in I_k}$, for $k \in K$, over U with $\mathrm{Av}(\beta_k/U) = \mathrm{Av}(\beta/U)$. Since $\bar{b}_0 \mathrel{\sqrt[\mathrm{s}]{}}_U M$, it follows by Lemma 2.19 that there is some tuple $\bar{c} \equiv_M \bar{b}_0$ such that

$$\bar{c} \mathrel{\sqrt[\mathrm{s}]{}}_U M\beta \cup \bigcup_{k\in K} \beta_k\,.$$

We claim that every sequence $\beta_k\bar{c}$ is indiscernible over U. Note that $\bar{c} \mathrel{\sqrt[\mathrm{s}]{}}_U \beta_k$. By Lemma F2.4.14, it is therefore sufficient to prove that

$$\bar{c} \equiv_{U\bar{b}^k[<i]} \bar{b}^k_i\,, \quad \text{for all } i \in I_k\,.$$

According to Lemma 2.19, $\mathrm{tp}(\bar{b}^k_i/M)$ has a unique $\sqrt[\mathrm{s}]{}$-free extension over $M \cup \bar{b}^k[<i]$. Consequently,

$$\bar{c} \mathrel{\sqrt[\mathrm{s}]{}}_M \bar{b}^k[<i]\,, \quad \bar{b}^k_i \mathrel{\sqrt[\mathrm{s}]{}}_M \bar{b}^k[<i]\,, \quad \text{and} \quad \bar{c} \equiv_M \bar{b}_0 \equiv_M \bar{b}^k_i$$

implies that $\bar{c} \equiv_{M\bar{b}^k[<i]} \bar{b}^k_i$. □

In particular, every $\sqrt[i]{}$-Morley sequence has the confluence property. The converse statement also holds. The proof is split into several steps. We start by showing that every sequence α with the confluence property is generated by some invariant type. This type is the so-called *eventual type* of α.

Definition 3.7. The *eventual type* of a sequence $\alpha = (\bar{a}_i)_{i \in I}$ is the set

$$\operatorname{Ev}(\alpha/U) := \{ \varphi(\bar{x}) \mid \varphi(\bar{x}) \in \operatorname{CF}(\alpha\beta) \text{ for some maximally } \varphi\text{-alternating extension } \alpha\beta \text{ of } \alpha \text{ over } U \}.$$

Example. We consider the theory of open dense linear orders. By quantifier-elimination, every strictly increasing sequence $\alpha = (a_i)_{i \in I}$ in $\mathbb{M}$ is indiscernible. Furthermore, such a sequence α is maximally $(x > c)$-alternating, for $c \in \mathbb{M}$, if $a_i > c$, for some $i \in I$. It follows that the eventual type $\operatorname{Ev}(\alpha/\varnothing)$ contains all formulae of the form $x > c$ with $c \in \mathbb{M}$.

Lemma 3.8. *Let $\varphi(\bar{x})$ be a formula over $\mathbb{M}$ and $\alpha = (\bar{a}_i)_{i \in I}$ an infinite indiscernible sequence over U.*

(a) *If α is maximally φ-alternating over U, then*

$$\varphi(\bar{x}) \in \operatorname{CF}(\alpha) \quad \text{iff} \quad \varphi(\bar{x}) \in \operatorname{CF}(\alpha\beta),$$

for every extension $\alpha\beta$ of α that is indiscernible over U.

(b) *If α has the confluence property over U, then*

$$\varphi(\bar{x}) \in \operatorname{CF}(\alpha\beta) \quad \text{iff} \quad \varphi(\bar{x}) \in \operatorname{CF}(\alpha\gamma).$$

for all maximally φ-alternating extensions $\alpha\beta$ and $\alpha\gamma$ of α.

Proof. (a) Set $n := \operatorname{alt}_\varphi(\alpha)$ and let $\bar{k} \in [I]^{n+1}$ be a sequence of indices such that

$$\mathbb{M} \vDash \varphi(\bar{a}_{k_i}) \leftrightarrow \neg\varphi(\bar{a}_{k_{i+1}}), \quad \text{for all } i < n.$$

Then

$$\varphi(\bar{x}) \in \mathrm{CF}(\alpha) \quad \text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}_{k_n}) .$$

For a contradiction, suppose that there is an extension $\alpha\beta = (\bar{a}_i)_{i \in I+J}$ that is indiscernible over U such that

$$\varphi(\bar{x}) \in \mathrm{CF}(\alpha/\mathbb{M}) \quad \text{iff} \quad \varphi(\bar{x}) \notin \mathrm{CF}(\alpha\beta/\mathbb{M}) .$$

Then there is some index $j \in J$ such that

$$\mathbb{M} \vDash \varphi(\bar{a}_j) \leftrightarrow \neg\varphi(\bar{a}_{k_n}) .$$

Consequently, the tuple $\bar{k}j \in [I + J]^{n+2}$ witnesses that $\mathrm{alt}_\varphi(\alpha\beta) > n$. Hence, α is not maximally φ-alternating. A contradiction.

(b) As $\alpha\beta$ and $\alpha\gamma$ have the same average type over U as α and this type has the confluence property, we can find some tuple $\bar{c}$ such that $\alpha\beta\bar{c}$ and $\alpha\gamma\bar{c}$ are indiscernible over U. Since $\alpha\beta$ and $\alpha\gamma$ are maximally φ-alternating, it follows by (a) that

$$\begin{aligned} \varphi(\bar{x}) \in \mathrm{CF}(\alpha\beta) \quad &\text{iff} \quad \varphi(\bar{x}) \in \mathrm{CF}(\alpha\beta\bar{c}) \\ &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{c}) \\ &\text{iff} \quad \varphi(\bar{x}) \in \mathrm{CF}(\alpha\gamma\bar{c}) \\ &\text{iff} \quad \varphi(\bar{x}) \in \mathrm{CF}(\alpha\gamma) . \end{aligned}$$

□

Lemma 3.9. *Let T be a theory without the independence property and let $\alpha = (\bar{a}_i)_{i \in I}$ be an infinite sequence with the confluence property over U.*

(a) $\mathfrak{p} := \mathrm{Ev}(\alpha/U)$ *is a complete type over* $\mathbb{M}$.

(b) $\mathfrak{p}$ *is invariant over U.*

(c) α *is generated by* $\mathfrak{p}$ *over U.*

Proof. (a) Let $\varphi(\bar{x})$ be a formula over $\mathbb{M}$. By Corollary 1.3 there exists a maximally φ-alternating extension $\alpha\beta$ of α. Then $\alpha\beta$ is also maximally

$\neg\varphi$-alternating and it follows by Lemma 3.8 (b) that

$$\begin{aligned} \varphi(\bar{x}) \in \mathrm{Ev}(\alpha/U) \quad &\text{iff} \quad \varphi(\bar{x}) \in \mathrm{CF}(\alpha\beta) \\ &\text{iff} \quad \neg\varphi(\bar{x}) \notin \mathrm{CF}(\alpha\beta) \\ &\text{iff} \quad \neg\varphi(\bar{x}) \notin \mathrm{Ev}(\alpha/U)\,. \end{aligned}$$

Hence, it remains to prove that $\mathrm{Ev}(\alpha/U)$ is satisfiable. Consider finitely many formulae $\varphi_0(\bar{x}), \dots, \varphi_n(\bar{x}) \in \mathrm{Ev}(\alpha/U)$. By Corollary 1.3 there exists an extension $\alpha\beta$ of α that is maximally φ_i-alternating over U, for all $i \leq n$. Suppose that $\beta = (\bar{b}_j)_{j\in J}$. Then

$$\varphi_i(\bar{x}) \in \mathrm{Ev}(\alpha/U) \quad \text{implies} \quad \varphi_i(\bar{x}) \in \mathrm{CF}(\alpha\beta)\,, \quad \text{for all } i \leq n\,,$$

and there exists some index $k \in J$ such that

$$\mathbb{M} \vDash \varphi_i(\bar{b}_j)\,, \quad \text{for all } j \geq k \text{ and } i \leq n\,.$$

This implies that $\mathbb{M} \vDash \varphi_0(\bar{b}_k) \wedge \dots \wedge \varphi_n(\bar{b}_k)$. Hence, $\{\varphi_0, \dots, \varphi_n\}$ is satisfiable.

(b) Consider tuples $\bar{b} \equiv_U \bar{b}'$ and a formula $\varphi(\bar{x}; \bar{y})$ over U. To show that

$$\varphi(\bar{x}; \bar{b}) \in \mathrm{Ev}(\alpha/U) \quad \text{iff} \quad \varphi(\bar{x}; \bar{b}') \in \mathrm{Ev}(\alpha/U)$$

we use Corollary 1.3 to find an extension $\alpha\beta$ of α that is maximally $\varphi(\bar{x}; \bar{b})$-alternating and maximally $\varphi(\bar{x}; \bar{b}')$-alternating over U. Choose a sequence $\alpha'\beta'$ such that

$$\alpha\beta\bar{b} \equiv_U \alpha'\beta'\bar{b}'\,.$$

Then $\alpha'\beta'$ is maximally $\varphi(\bar{x}; \bar{b}')$-alternating. As the type $\mathrm{Av}(\alpha\beta/U) = \mathrm{Av}(\alpha'\beta'/U)$ has the confluence property over U, there is some tuple $\bar{c}$ such that $\alpha\beta\bar{c}$ and $\alpha'\beta'\bar{c}$ are both indiscernible over U. It follows by

Lemma 3.8 (a) that

$$\begin{aligned}
\varphi(\bar{x};\bar{b}) \in \mathrm{Ev}(\alpha/U) \quad &\text{iff} \quad \varphi(\bar{x};\bar{b}) \in \mathrm{CF}(\alpha\beta) \\
&\text{iff} \quad \varphi(\bar{x};\bar{b}') \in \mathrm{CF}(\alpha'\beta') \\
&\text{iff} \quad \varphi(\bar{x};\bar{b}') \in \mathrm{CF}(\alpha'\beta'\bar{c}) \\
&\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{c};\bar{b}') \\
&\text{iff} \quad \varphi(\bar{x};\bar{b}') \in \mathrm{CF}(\alpha\beta\bar{c}) \\
&\text{iff} \quad \varphi(\bar{x};\bar{b}') \in \mathrm{CF}(\alpha\beta) \\
&\text{iff} \quad \varphi(\bar{x};\bar{b}') \in \mathrm{Ev}(\alpha/U)\,.
\end{aligned}$$

(c) To show that $\bar{a}_k$ realises the type $\mathfrak{p} \upharpoonright U\bar{a}[<k]$, we consider a formula $\varphi(\bar{x};\bar{y}_0,\dots,\bar{y}_{n-1})$ over U and a tuple $\bar{\imath} \in [I]^n$ of indices with $\bar{\imath} < k$. Fix a maximally $\varphi(\bar{x};\bar{a}[\bar{\imath}])$-alternating extension $\alpha\beta$ of α over U and let $\bar{c}$ be a tuple such that $\alpha\beta\bar{c}$ is indiscernible over U. Then Lemma 3.8 implies that

$$\begin{aligned}
\varphi(\bar{x};\bar{a}[\bar{\imath}]) \in \mathfrak{p} \upharpoonright U\bar{a}[<k] \quad &\text{iff} \quad \varphi(\bar{x};\bar{a}[\bar{\imath}]) \in \mathrm{CF}(\alpha\beta) \\
&\text{iff} \quad \varphi(\bar{x};\bar{a}[\bar{\imath}]) \in \mathrm{CF}(\alpha\beta\bar{c}) \\
&\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{c};\bar{a}[\bar{\imath}]) \\
&\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}_k;\bar{a}[\bar{\imath}])\,,
\end{aligned}$$

where the last step follows by indiscernibility. □

Combining the above results, we obtain the following characterisation of $\sqrt[\mathrm{i}]{}$-Morley sequences in theories without the independence property.

Theorem 3.10. *Let T be a theory without the independence property, $\alpha = (\bar{a}_i)_{i\in I}$ an infinite sequence, and $\mathfrak{p}$ a type. The following statements are equivalent:*

(1) *α is a $\sqrt[\mathrm{i}]{}$-Morley sequence for $\mathfrak{p} \upharpoonright U$ over U and* $\mathfrak{p} = \mathrm{Ev}(\alpha/U)$.

(2) *α has the confluence property over U and* $\mathfrak{p} = \mathrm{Ev}(\alpha/U)$.

(3) $\mathfrak{p}$ *is a global type that is invariant over* U *and* α *is generated by* $\mathfrak{p}$ *over* U.

Proof. The implication (1) ⇒ (2) follows by Lemma 3.6, and (2) ⇒ (3) was already proved in Lemma 3.9.

(3) ⇒ (1) For $i \leq j$ in I, we have

$$\mathrm{tp}(\bar{a}_j/U\bar{a}[<i]) = \mathfrak{p} \restriction U\bar{a}[<i] = \mathrm{tp}(\bar{a}_i/U\bar{a}[<i])\,.$$

Furthermore, $\mathrm{tp}(\bar{a}_i/U\bar{a}[<i])$ extends to $\mathfrak{p}$, a complete type over $\mathbb{M}$ that is invariant over U. Consequently, we have $\bar{a}_i \overset{\mathrm{i}}{\surd}_U \bar{a}[<i]$ and it follows by Lemma F2.4.14 that α is indiscernible over U.

We have shown that α is a $\overset{\mathrm{i}}{\surd}$-Morley sequence for $\mathfrak{p} \restriction U$ over U. It therefore remains to prove that $\mathfrak{p} = \mathrm{Ev}(\alpha/U)$. Let $\varphi(\bar{x};\bar{c}) \in \mathrm{Ev}(\alpha/U)$ be a formula with parameters $\bar{c} \subseteq \mathbb{M}$. Then $\varphi(\bar{x};\bar{c}) \in \mathrm{CF}(\alpha\beta)$, for some maximally $\varphi(\bar{x};\bar{c})$-alternating extension $\alpha\beta$ of α over U. Let $\bar{b}$ be a tuple realising $\mathfrak{p} \restriction U\alpha\beta\bar{c}$. Applying Lemma 3.4 to the sequences α and $\alpha\beta$, it follows that $\alpha\beta$ is generated by $\mathfrak{p}$ over U. By choice of $\bar{b}$, so is $\alpha\beta\bar{b}$. Consequently, Lemma F2.4.14 implies that the sequence $\alpha\beta\bar{b}$ is indiscernible over U. As $\alpha\beta$ is maximally $\varphi(\bar{x};\bar{c})$-alternating, we therefore have $\varphi(\bar{x};\bar{c}) \in \mathrm{CF}(\alpha\beta\bar{b})$, which implies that $\mathbb{M} \vDash \varphi(\bar{b};\bar{c})$. By choice of $\bar{b}$, it follows that $\varphi(\bar{x};\bar{c}) \in \mathfrak{p} \restriction U\alpha\beta\bar{c} \subseteq \mathfrak{p}$. □

Corollary 3.11. *Let* α *and* β *be infinite* $\overset{\mathrm{i}}{\surd}$*-Morley sequences over* U*. The following statements are equivalent:*

(1) $\mathrm{Av}(\alpha/U) = \mathrm{Av}(\beta/U)$

(2) $\mathrm{Ev}(\alpha/U) = \mathrm{Ev}(\beta/U)$

(3) *There is some complete type* $\mathfrak{p}$ *over* $\mathbb{M}$ *that is invariant over* U *such that* α *and* β *are both generated by* $\mathfrak{p}$.

Proof. (2) ⇒ (3) By Theorem 3.10, both sequences are generated by the type $\mathrm{Ev}(\alpha/U) = \mathrm{Ev}(\beta/U)$, which is complete and invariant over U.

(3) ⇒ (2) If α and β are both generated by $\mathfrak{p}$, it follows by Theorem 3.10 that $\mathrm{Ev}(\alpha/U) = \mathfrak{p} = \mathrm{Ev}(\beta/U)$.

(1) ⇒ (3) By Theorem 3.10, α is generated by $\mathfrak{p} := \mathrm{Ev}(\alpha/U)$. Hence, Lemma 3.4 implies that so is β.

(3) ⇒ (1) follows by Lemma 3.4. □

As a consequence we can derive the following bound on the number of invariant global types.

Proposition 3.12. *Let T be a theory without the independence property and let $\mathfrak{M}$ be a model of T. There exists a bijection between types $\mathfrak{p} \in S^{<\omega}(\mathbb{M})$ that are invariant over M and average types $\mathrm{Av}(\alpha/M)$ of infinite $\sqrt[\mathrm{i}]{}$-Morley sequences α over M.*

Proof. We map a type $\mathfrak{p} \in S^{<\omega}(\mathbb{M})$ that is invariant over M to the average type

$$\Phi_{\mathfrak{p}} := \mathrm{Av}(\alpha/M)\,,$$

where α is any infinite sequence generated by $\mathfrak{p}$ over M. According to Theorem 3.10, the resulting sequence is a $\sqrt[\mathrm{i}]{}$-Morley sequence. Furthermore, if α and β are both generated by $\mathfrak{p}$ over M, it follows by Corollary 3.11 that $\mathrm{Av}(\alpha/M) = \mathrm{Av}(\beta/M)$. Consequently, $\Phi_{\mathfrak{p}}$ does not depend on the choice of α.

The inverse of the function $\mathfrak{p} \mapsto \Phi_{\mathfrak{p}}$ maps an average type Φ of an infinite $\sqrt[\mathrm{i}]{}$-Morley sequence α over M to the type $\mathfrak{p}_\Phi := \mathrm{Ev}(\alpha/M)$. Again it follows by Corollary 3.11 that the type $\mathfrak{p}_\Phi$ does not depend on the choice of α.

It remains to prove that the functions $\mathfrak{p} \mapsto \Phi_{\mathfrak{p}}$ and $\Phi \mapsto \mathfrak{p}_\Phi$ are inverse to each other. Let $\mathfrak{p} \in S^{<\omega}(\mathbb{M})$ be a type that is invariant over M and let α be an infinite sequence that is generated by $\mathfrak{p}$ over M. Then it follows by Theorem 3.10 that $\mathfrak{p}_{\Phi_{\mathfrak{p}}} = \mathrm{Ev}(\alpha/M) = \mathfrak{p}$.

Conversely, consider an average type Φ of some infinite $\sqrt[\mathrm{i}]{}$-Morley sequence α and let $\mathfrak{p}_\Phi := \mathrm{Ev}(\alpha/M)$. By Theorem 3.10, α is generated by $\mathfrak{p}_\Phi$, which implies that $\Phi_{\mathfrak{p}_\Phi} = \mathrm{Av}(\alpha/M) = \Phi$. □

As an application, we derive the following characterisation of theories without the independence property.

Theorem 3.13. *Let T be a complete first-order theory. The following statements are equivalent:*

(1) *T does not have the independence property.*

(2) *$\overset{\mathrm{f}}{\surd}$ is weakly bounded by $f(\kappa) = 2^\kappa$.*

(3) *There is some cardinal $\kappa \geq |T|$ such that, for every type $\mathfrak{p} \in S^{<\omega}(M)$ where M is a model of size $|M| = \kappa$, there are less than 2^{2^κ} $\overset{\mathrm{u}}{\surd}$-free extensions of $\mathfrak{p}$ over any given set $C \supseteq M$.*

(4) *For every $\kappa \geq |T|$, every set U of size $|U| = \kappa$, every type $\mathfrak{p} \in S^{<\omega}(U)$, and every set C, there are at most 2^κ $\overset{\mathrm{u}}{\surd}$-free extensions of $\mathfrak{p}$ over $U \cup C$.*

Proof. (4) $\Rightarrow$ (3) is trivial.

(2) $\Rightarrow$ (4) Let $\kappa \geq |T|$ and let U be a set of size $|U| = \kappa$. Consider a type $\mathfrak{p} \in S^{<\omega}(U)$ and some set $C \subseteq \mathbb{M}$. Let $(\mathfrak{q}_i)_{i<\lambda}$ be an enumeration of all $\overset{\mathrm{u}}{\surd}$-free extensions of $\mathfrak{p}$ over $U \cup C$. Since $\overset{\mathrm{u}}{\surd} \subseteq \overset{\mathrm{f}}{\surd}$, it follows that each $\mathfrak{q}_i$ is also a $\overset{\mathrm{f}}{\surd}$-free extension of $\mathfrak{p}$. By (2), there are at most $2^{|T| \oplus |U|}$ such extensions. Hence, $\lambda \leq 2^{|T|\oplus|U|} = 2^\kappa$.

(1) $\Rightarrow$ (2) Let $U, C \subseteq \mathbb{M}$ be sets and let $(\mathfrak{p}_i)_{i<\lambda}$ be an enumeration without repetitions of all types over $U \cup C$ that do not fork over U. We have to show that $\lambda \leq 2^{|T|\oplus|U|}$. Let $\mathfrak{M}$ be a model of T containing U of size $|M| \leq |T| \oplus |U|$ and let $\mathfrak{N}$ be a model containing $M \cup C$ that is $(|T|\oplus|U|)^+$-saturated and strongly $(|T|\oplus|U|)^+$-homogeneous. By (EXT), we can fix, for every $i < \lambda$, some type $\mathfrak{q}_i \supseteq \mathfrak{p}_i$ over N that does not fork over U. Note that $\mathfrak{p}_i \neq \mathfrak{p}_k$ implies that $\mathfrak{q}_i \neq \mathfrak{q}_k$, for $i \neq k$. Since T does not have the independence property, it follows by Theorem 2.26 that $\overset{\mathrm{f}}{\surd} = \overset{\mathrm{li}}{\surd}$. Hence, each $\mathfrak{q}_i$ is $\overset{\mathrm{li}}{\surd}$-free over U and, thus, also over M. Consequently, we can use Proposition 2.20 to show that $\mathfrak{q}_i$ is $\overset{\mathrm{i}}{\surd}$-free over M. Note that there are at most $2^{|T|\oplus|M|} = 2^{|T|\oplus|U|}$ average types $\mathrm{Av}(\alpha/M)$ of $\overset{\mathrm{i}}{\surd}$-Morley sequences α over M. By Corollary 3.11, this means that there also are at most that many eventual type $\mathrm{Ev}(\alpha/M)$ of such sequences α. Therefore we can use Theorem 3.10 to show that there are at most that many types over N that are $\overset{\mathrm{i}}{\surd}$-free over M. This implies that $\lambda \leq 2^{|T|\oplus|U|}$.

(3) ⇒ (1) Suppose that there is some formula $\varphi(\bar{x};\bar{y})$ with the independence property. Then there are families $(\bar{a}_i)_{i<\omega}$ and $(\bar{b}_s)_{s\subseteq\omega}$ such that

$$\mathbb{M} \vDash \varphi(\bar{a}_i, \bar{b}_s) \quad \text{iff} \quad i \in s\,.$$

Let $\mathfrak{M}$ be a model of T of size $|M| = \kappa$ that contains α and β. We have seen in Theorem B2.4.13 that there are 2^{2^κ} ultrafilters over the set $A := \{\, \bar{a}_i \mid i < \kappa \,\}$. For every ultrafilter $\mathfrak{u}$ over A, set

$$\mathfrak{p}_\mathfrak{u} := \mathrm{Av}(\mathfrak{u}/MC)\,.$$

By Lemma F2.3.10, $\mathfrak{p}_\mathfrak{u}$ is a $\sqrt[\mathrm{u}]{\ }$-free extension of $\mathfrak{p}_\mathfrak{u} \upharpoonright M$. Furthermore, if $\mathfrak{u} \neq \mathfrak{v}$ are distinct ultrafilters, we can fix some set $B \in \mathfrak{u} \setminus \mathfrak{v}$ and an index $s \subseteq \omega$ such that

$$\mathfrak{M} \vDash \varphi(\bar{a}_i;\bar{b}_s) \quad \text{iff} \quad \bar{a}_i \in B\,.$$

Consequently, $\varphi(\bar{x};\bar{b}_s) \in \mathfrak{p}_\mathfrak{u} \setminus \mathfrak{p}_\mathfrak{v}$, which implies that $\mathfrak{p}_\mathfrak{u} \neq \mathfrak{p}_\mathfrak{v}$. It follows that there are at least 2^{2^κ} types over $M \cup C$ that are $\sqrt[\mathrm{u}]{\ }$-free over M. □

4. Dp-rank

Mutually indiscernible sequences

We can characterise theories without the independence property also in terms of a rank that is based on mutually indiscernible sequences.

Definition 4.1. A family $(\alpha_k)_{k\in K}$ of sequences is *mutually indiscernible* over a set U if each sequence α_k is indiscernible over $U \cup \alpha[K \setminus \{k\}]$.

Before giving the definition of the dp-rank, we collect some technical properties of mutually indiscernible sequences. Let us start with ways to construct such families. The first observation is trivial.

Lemma 4.2. *Let $\alpha := (\bar{a}_i)_{i \in I}$ be an indiscernible sequence over U and let $\sim$ be a convex equivalence relation on I. The family $(\alpha|_E)_{E \in I/\sim}$ is mutually indiscernible over U.*

Lemma 4.3. *Let $(\alpha_k)_{k<\gamma}$ be a family of sequences and U a set of parameters. If $(\beta_k)_{k<\gamma}$ is a family such that each β_k is an indiscernible sequence over $U\alpha[>k]\beta[<k]$ with*

$$\mathrm{Av}(\beta_k / U\alpha[>k]\beta[<k]) \supseteq \mathrm{Av}(\alpha_k / U\alpha[>k]\beta[<k]),$$

then $(\beta_k)_{k<\gamma}$ is mutually indiscernible over U.

Proof. Suppose that $\alpha_k = (\bar{a}^k_i)_{i \in I_k}$ and $\beta_k = (\bar{b}^k_i)_{i \in J_k}$, for $k < \gamma$. To show that $(\beta_k)_{k<\gamma}$ is mutually indiscernible over U, we fix some index $k < \gamma$ and we prove by induction on $k < l \leq \gamma$ that β_k is indiscernible over $U\alpha[\geq l]\beta[\downarrow l \setminus \{k\}]$. The result then follows for $l = \gamma$.

For $l = k + 1$, the claim holds by choice of β_k. For the inductive step, suppose that we have already shown that β_k is indiscernible over the set $U\alpha[\geq l]\beta[\downarrow l \setminus \{k\}]$. To show that it is also indiscernible over

$$U\alpha[\geq(l+1)]\beta[\downarrow(l+1) \setminus \{k\}],$$

consider a formula $\varphi(\bar{x}_0, \dots, \bar{x}_{n-1}; \bar{c}, \bar{d})$ with parameters

$$\bar{c} \subseteq \beta^l \quad \text{and} \quad \bar{d} \subseteq U\alpha[\geq(l+1)]\beta[\downarrow l \setminus \{k\}].$$

We have to show that

$$\mathbb{M} \models \varphi(\bar{b}^k[\bar{\imath}]; \bar{c}, \bar{d}) \leftrightarrow \varphi(\bar{b}^k[\bar{\jmath}]; \bar{c}, \bar{d}), \quad \text{for all } \bar{\imath}, \bar{\jmath} \in [J_k]^n.$$

W.l.o.g. we may assume that $\bar{c} = \bar{b}^l[\bar{s}]$, for some $\bar{s} \in [J_l]^m$. Fix indices $\bar{\imath}, \bar{\jmath} \in [J_k]^n$. By inductive hypothesis, the sequence β_k is indiscernible over $U\alpha[\geq l]\beta[\downarrow l \setminus \{k\}]$. Therefore, we have

$$\mathbb{M} \models \varphi(\bar{b}^k[\bar{\imath}]; \bar{a}^l[\bar{t}], \bar{d}) \leftrightarrow \neg\varphi(\bar{b}^k[\bar{\jmath}]; \bar{a}^l[\bar{t}], \bar{d}),$$

for all and all $\bar{t} \in [I_l]^m$. This implies that the formula

$$\varphi(\bar{b}^k[\bar{\imath}]; \bar{x}, \bar{d}) \leftrightarrow \varphi(\bar{b}^k[\bar{\jmath}]; \bar{x}, \bar{d})$$

belongs to

$$\mathrm{Av}(\alpha_l / U\alpha[>l]\beta[<l]) \subseteq \mathrm{Av}(\beta_l / U\alpha[>l]\beta[<l]) .$$

Consequently,

$$\mathbb{M} \vDash \varphi(\bar{b}^k[\bar{\imath}]; \bar{b}^l[\bar{s}], \bar{d}) \leftrightarrow \varphi(\bar{b}^k[\bar{\jmath}]; \bar{b}^l[\bar{s}], \bar{d}) , \qquad \square$$

Let us note the following property of sequences 'diagonally crossing' a family of mutually indiscernible sequences.

Lemma 4.4. *Let $\alpha = (\alpha_k)_{k \in K}$ be a family of mutually indiscernible sequences $\alpha_k = (\bar{a}^k_i)_{i \in I_k}$ over U.*

(a) $(\bar{a}^k_{\eta(k)})_{k \in K} \equiv_U (\bar{a}^k_{\zeta(k)})_{k \in K}$, *for all $\eta, \zeta \in \prod_{k \in K} I_k$.*

(b) *If the index set K is ordered and the sequence $\alpha = (\alpha_k)_{k \in K}$ is indiscernible over U, then each sequence of the form $(\bar{a}^k_{\eta(k)})_{k \in K}$ with $\eta \in \prod_{k \in K} I_k$ is also indiscernible over U.*

Proof. (a) We prove by induction on $n < \omega$ that

$$\bar{a}^{k_0}_{\eta(k_0)} \dots \bar{a}^{k_{n-1}}_{\eta(k_{n-1})} \equiv_{U\alpha[K \setminus \bar{k}]} \bar{a}^{k_0}_{\zeta(k_0)} \dots \bar{a}^{k_{n-1}}_{\zeta(k_{n-1})} , \quad \text{for all } \bar{k} \in [K]^n.$$

For $n = 0$, there is nothing to do. For the inductive step, suppose that we have proved the claim already for n and let $\bar{k} \in [K]^{n+1}$. By mutual indiscernibility, we have

$$\bar{a}^{k_n}_{\eta(k_n)} \equiv_{U\alpha[K \setminus \{k_n\}]} \bar{a}^{k_n}_{\zeta(k_n)} .$$

Therefore, it follows by inductive hypothesis that

$$\begin{aligned}\bar{a}^{k_0}_{\eta(k_0)} \dots \bar{a}^{k_{n-1}}_{\eta(k_{n-1})} \bar{a}^{k_n}_{\eta(k_n)} &\equiv_{U\alpha[K \setminus \bar{k}]} \bar{a}^{k_0}_{\zeta(k_0)} \dots \bar{a}^{k_{n-1}}_{\zeta(k_{n-1})} \bar{a}^{k_n}_{\eta(k_n)} \\ &\equiv_{U\alpha[K \setminus \bar{k}]} \bar{a}^{k_0}_{\zeta(k_0)} \dots \bar{a}^{k_{n-1}}_{\zeta(k_{n-1})} \bar{a}^{k_n}_{\zeta(k_n)} .\end{aligned}$$

(b) Note that indiscernibility of α implies that all index orders I_k are isomorphic. Hence, we may w.l.o.g. assume that $I_k = I$, for some fixed order I. Fix an element $i \in I$. Indiscernibility of α over U implies that the restriction $(\bar{a}^k_i)_{k \in K}$ is also indiscernible over U. By (a) it follows that so is every sequence of the form $(\bar{a}^k_{\eta(k)})_{k \in K}$ with $\eta \in I^K$. ☐

We obtain the following generalisation of Lemma E5.3.11.

Corollary 4.5. *Suppose that $(\alpha_k)_{k \in K}$ is a family of mutually indiscernible sequences over U. For every set C, there exists a set $C' \equiv_U C$ such that $(\alpha_k)_{k \in K}$ is mutually indiscernible over $U \cup C'$.*

Proof. Suppose that $K = \kappa$ is a cardinal and let $\alpha_k = (\bar{a}^k_i)_{i \in I_k}$. By induction on $k < \kappa$, we use Proposition E5.3.6 to choose an indiscernible sequence $\beta_k = (\bar{b}^k_i)_{i \in I_k}$ over $U \cup C \cup \alpha[>k]\beta[<k]$ such that

$$\operatorname{Av}(\beta_k / U\alpha[>k]\beta[<k]) \supseteq \operatorname{Av}(\alpha_k / U\alpha[>k]\beta[<k]) .$$

Then it follows by Lemma 4.3 that the family $(\beta_k)_{k \in K}$ is mutually indiscernible over $U \cup C$. As each α_k is indiscernible over $U \cup \alpha[K \setminus \{k\}]$, we have

$$\operatorname{Av}(\beta_k / U\alpha[K \setminus \{k\}]) = \operatorname{Av}(\alpha_k / U\alpha[K \setminus \{k\}]) .$$

This implies that

$$(\beta_k)_{k \in K} \equiv_U (\alpha_k)_{k \in K} .$$

Therefore, there exists an automorphism $\pi \in \operatorname{Aut} \mathbb{M}_U$ mapping one family to the other one. Consequently, $(\alpha_k)_{k \in K}$ is mutually indiscernible over $U \cup \pi[C]$. ☐

Corollary 4.6. *Let $\alpha = (\alpha_k)_{k \in K}$ be a family of mutually indiscernible sequences $\alpha_k = (\bar{a}^k_i)_{i \in I_k}$ over U. For every family of linear orders $J_k \supseteq I_k$, $k \in K$, there exist sequences $\alpha'_k = (\bar{a}^k_j)_{j \in J_k}$ extending α_k such that the family $(\alpha'_k)_{k \in K}$ is mutually indiscernible over U.*

Proof. As in the preceding corollary, we choose by induction on k an indiscernible sequence $\beta_k = (\bar{b}^k_i)_{i\in J_k}$ over $U \cup \alpha[>k]\beta[<k]$ such that

$$\mathrm{Av}(\beta_k/U\alpha[>k]\beta[<k]) \supseteq \mathrm{Av}(\alpha_k/U\alpha[>k]\beta[<k])\,.$$

Then it follows by Lemma 4.3 that the family $(\beta_k)_{k\in K}$ is mutually indiscernible over U. As each α_k is indiscernible over $U \cup \alpha[K \setminus \{k\}]$, we have

$$\mathrm{Av}(\beta_k|_{I_k}/U\alpha[K \setminus \{k\}]) = \mathrm{Av}(\alpha_k/U\alpha[K \setminus \{k\}])\,.$$

Consequently, there exists an automorphism $\pi \in \mathrm{Aut}\,\mathbb{M}_U$ mapping each $\beta_k|_{I_k}$ to α_k. The family $(\pi(\beta_k))_{k\in K}$ is the desired extension of α. □

Proposition 4.7. *Let T be a theory without the independence property and let $(\alpha_k)_{k\in K}$ be a family of mutually indiscernible sequences over U. For every set C, there exists a subset $K_0 \subseteq K$ of size $|K_0| \leq |T| \oplus |C|$ such that $(\alpha_k)_{k\in K\setminus K_0}$ is mutually indiscernible over $U \cup C$.*

Proof. Suppose that $\alpha_k = (\bar{a}^k_i)_{i\in I_k}$ where each $\bar{a}^k_i = (a^k_{i,j})_{j<\gamma_k}$ is a γ_k-tuple. Let $\mathfrak{M}$ be a model containing U and all sequences α_k, and define

$$\begin{aligned}
P &:= U \cup \{\, a^k_{i,j} \mid k \in K,\ i \in I_k,\ j < \gamma_k \,\},\\
E &:= \{\, \langle a^k_{i,j}, a^k_{i,j'}\rangle \mid k \in K,\ i \in I_k,\ j, j' < \gamma_k \,\},\\
F &:= \{\, \langle a^k_{i,j}, a^k_{i',j'}\rangle \mid k \in K,\ i, i' \in I_k,\ j, j' < \gamma_k \,\},\\
R &:= \{\, \langle a^k_{i,j}, a^k_{i',j}\rangle \mid k \in K,\ i < i' \text{ in } I_k,\ j < \gamma_k \,\}.
\end{aligned}$$

Fix an $|M|^+$-saturated elementary extension

$$\langle \mathfrak{M}_+, P_+, U_+, E_+, F_+, R_+\rangle \succeq \langle \mathfrak{M}, P, U, E, F, R\rangle\,.$$

Using the relations E_+, F_+, and R_+ we see that there are a set $K_+ \supseteq K$, linear orders I^+_k, ordinals γ^+_k, and a family

$$\big(b^k_{i,j}\big)_{k\in K_+, i\in I^+_k,\ j<\gamma^+_k}$$

of elements such that, setting $\bar{b}^k_i := (b^k_{i,j})_{j<\gamma^+_k}$ and $\beta_k := (\bar{b}^k_i)_{i\in I^+_k}$, we have

- $P_+ = U_+ \cup \beta[K_+]$,
- $I_k^+ \supseteq I_k$, $\gamma_k^+ \geq \gamma_k$, and $b_{i,j}^k = a_{i,j}^k$, for $k \in K$, $i \in I_k$, $j < \gamma_k$,
- the family $(\beta_k)_{k \in K_+}$ is mutually indiscernible over U_+.

By Lemma 1.12, we can find a set $W \subseteq P_+$ of size $|W| \leq |T| \oplus |C|$ such that

$$\bar{a} \equiv_W \bar{a}' \quad \text{implies} \quad \bar{a} \equiv_C \bar{a}', \quad \text{for all } \bar{a}, \bar{a}' \subseteq P.$$

We choose a set $K_0 \subseteq K$ of size $|K_0| \leq |W| \leq |T| \oplus |C|$ such that $W \subseteq \beta[K_0]$. We claim that the family $(\alpha_k)_{k \in K \setminus K_0}$ is mutually indiscernible over $U \cup C$. Fix $k \in K' := K \setminus K_0$ and let $\bar{\imath}, \bar{\jmath} \in [I_k]^m$. We have to show that

$$\bar{a}^k[\bar{\imath}] \equiv_{UC\alpha[K' \setminus \{k\}]} \bar{a}^k[\bar{\jmath}].$$

Let $\bar{d} \subseteq U \cup \alpha[K' \setminus \{k\}]$ be finite. Since the sequence β_k is indiscernible over $U \cup \beta[K \setminus \{k\}] \supseteq \bar{d}\beta[K_0]$, we have

$$\bar{b}^k[\bar{\imath}] \equiv_{\bar{d}\beta[K_0]} \bar{b}^k[\bar{\jmath}], \quad \text{which implies that} \quad \bar{a}^k[\bar{\imath}]\bar{d} \equiv_W \bar{a}^k[\bar{\jmath}]\bar{d}.$$

By choice of W, it follows that $\bar{a}^k[\bar{\imath}]\bar{d} \equiv_C \bar{a}^k[\bar{\jmath}]\bar{d}$. We have shown that

$$\bar{a}^k[\bar{\imath}] \equiv_{C\bar{d}} \bar{a}^k[\bar{\jmath}], \quad \text{for all finite } \bar{d} \subseteq U \cup \alpha[K' \setminus \{k\}].$$

Consequently, $\bar{a}^k[\bar{\imath}] \equiv_{UC\alpha[K' \setminus \{k\}]} \bar{a}^k[\bar{\jmath}]$. □

Dp-rank

After these preparations we can introduce the dp-rank.

Definition 4.8. Let $\Phi(\bar{x})$ be a set of formulae over $\mathbb{M}$ and $U \subseteq \mathbb{M}$ a set of parameters.

(a) The *dp-rank* $\mathrm{rk}_{\mathrm{dp}}(\Phi/U)$ of Φ over U is the least cardinal κ such that, for every tuple $\bar{b}$ realising Φ and every family $(\alpha_i)_{i<\kappa}$ of infinite mutually indiscernible sequences over U, there is some index $i < \kappa$ such that α_i is indiscernible over $U\bar{b}$. If such a cardinal does not exist, we set $\mathrm{rk}_{\mathrm{dp}}(\Phi/U) := \infty$.

(b) For a tuple $\bar{a} \subseteq \mathbb{M}$, we set

$$\mathrm{rk}_{\mathrm{dp}}(\bar{a}/U) := \mathrm{rk}_{\mathrm{dp}}(\mathrm{tp}(\bar{a}/U)/U) .$$

Remark. Note that $\mathrm{rk}_{\mathrm{dp}}(\Phi/U) = 0$ if, and only if, Φ is inconsistent.

Example. Let us consider the theory of $\langle \mathbb{Q}, \leq \rangle$. By quantifier-elimination it follows that a family $\alpha = (\alpha_k)_{k \in K}$ of sequences is mutually indiscernible over a set U if, and only if, all tuples in α_k have the same order type over the set $U \cup \alpha[K \setminus \{k\}]$.

Consider a partial type $\Phi(\bar{x})$ with n free variables $\bar{x}$. We claim that

$$\mathrm{rk}_{\mathrm{dp}}(\Phi/\varnothing) \leq n + 1 .$$

Let $\bar{b}$ be an n-tuple realising Φ and $\alpha = (\alpha_k)_{k \leq n+1}$ a family of infinite mutually indiscernible sequences. For simplicity, let us assume that each α_k is a sequence of singletons. For $i \neq j$, it follows that either $\alpha_i < \alpha_j$ or $\alpha_j < \alpha_i$. Furthermore, for every $i < n$, there is at most one index k such that α_k contains both elements below and above b_i. Therefore, we can find some index $k \leq n + 1$ such that

$$\alpha_k < b_i \quad \text{or} \quad b_i < \alpha_k , \quad \text{for all } i < n .$$

This implies that α_k is indiscernible over $\bar{b}$.

We start by stating some basic monotonicity properties of the dp-rank.

Lemma 4.9. *Let Φ be a partial type over U. Then*

$$\mathrm{rk}_{\mathrm{dp}}(\Phi/U) = \mathrm{rk}_{\mathrm{dp}}(\Phi/UC) , \quad \textit{for every set } C .$$

Proof. Let $\kappa := \mathrm{rk}_{\mathrm{dp}}(\Phi/U)$ and consider a tuple $\bar{b}$ realising Φ and a family $(\alpha_k)_{k<\kappa}$ of infinite mutually indiscernible sequences over $U \cup C$.

Suppose that $\alpha_k = (\bar{a}^k_i)_{i \in I_k}$ and let $\bar{c}$ be an enumeration of C. Setting $\alpha'_k := (\bar{a}^k_i \bar{c})_{i \in I_k}$, we obtain a family $(\alpha'_k)_{k<\kappa}$ of infinite mutually indiscernible sequences over U. By choice of κ, there exists some index $k < \kappa$ such that α'_k is indiscernible over $U \cup \bar{b}$. Consequently, α_k is indiscernible over $U \cup \bar{b}\bar{c}$. Hence, $\mathrm{rk}_{\mathrm{dp}}(\Phi/UC) \leq \kappa$.

For the converse inequality, let $\lambda < \kappa$. Then there exists a tuple $\bar{b}$ realising Φ and a family $(\alpha_k)_{k<\lambda}$ of infinite mutually indiscernible sequences over U such that no α_k is indiscernible over $U \cup \bar{b}$. By Corollary 4.5, there exists an automorphism $\pi \in \mathrm{Aut}\,\mathbb{M}_U$ such that the family $(\pi(\alpha_k))_{k<\lambda}$ is mutually indiscernible over $U \cup C$. It follows that the tuple $\pi(\bar{b})$ realises Φ and no sequence $\pi(\alpha_k)$ is indiscernible over $U \cup C \cup \pi(\bar{b})$. Hence, $\mathrm{rk}_{\mathrm{dp}}(\Phi/UC) > \lambda$. □

Corollary 4.10.

(a) $\Phi \subseteq \Psi$ *implies* $\mathrm{rk}_{\mathrm{dp}}(\Phi/U) \geq \mathrm{rk}_{\mathrm{dp}}(\Psi/U)$.

(b) $U \subseteq V$ *implies* $\mathrm{rk}_{\mathrm{dp}}(\bar{a}/U) \geq \mathrm{rk}_{\mathrm{dp}}(\bar{a}/V)$.

Proof. (a) follows immediately from the definition. For (b), note that Lemma 4.9 and (a) implies that

$$\mathrm{rk}_{\mathrm{dp}}(\bar{a}/U) = \mathrm{rk}_{\mathrm{dp}}(\mathrm{tp}(\bar{a}/U)/V) \geq \mathrm{rk}_{\mathrm{dp}}(\bar{a}/V). \qquad \square$$

The next proposition collects several alternative characterisations of the dp-rank.

Proposition 4.11. *Let $\Phi(\bar{x})$ be a partial type over U and $\kappa > 0$ a cardinal. The following statements are equivalent:*

(1) $\mathrm{rk}_{\mathrm{dp}}(\Phi/U) \leq \kappa$

(2) *For every tuple $\bar{b}$ realising Φ and every family $(\alpha_k)_{k \in K}$ of infinite mutually indiscernible sequences over U, there is a set $K_0 \subseteq K$ of size $|K_0| < \kappa$ such that, for every $k \in K \setminus K_0$, all elements of α_k have the same type over $U\bar{b}$.*

(3) *For every tuple $\bar{b}$ realising Φ and every family $(\alpha_k)_{k \in K}$ of infinite mutually indiscernible sequences over U, there is a set $K_0 \subseteq K$ of size $|K_0| < \kappa$ such that the subfamily $(\alpha_k)_{k \in K \setminus K_0}$ is mutually indiscernible over $U\bar{b}$.*

Proof. (3) ⇒ (2) is trivial.

(2) ⇒ (1) Suppose that there exist a tuple $\bar{b}$ realising $\Phi(\bar{x})$ and a family $(\alpha_k)_{k<\kappa}$ of infinite mutually indiscernible sequences $\alpha_k = (\bar{a}^k_i)_{i \in I_k}$ over U such that no α_k is indiscernible over $U\bar{b}$. By Corollary 4.6, we may assume that every index order I_k is dense. For each $k < \kappa$, there are indices $\bar{\imath}, \bar{\jmath} \in [I_k]^{<\omega}$ such that

$$\bar{a}^k[\bar{\imath}] \not\equiv_{U\bar{b}} \bar{a}^k[\bar{\jmath}]\,.$$

Using Lemma E5.3.12 we obtain indices $\bar{u}^k < s^k < t^k < \bar{v}^k$ in I_k such that

$$\bar{a}^k[\bar{u}^k s^k \bar{v}^k] \not\equiv_{U\bar{b}} \bar{a}^k[\bar{u}^k t^k \bar{v}^k]\,.$$

It follows that the family $(\alpha'_k)_{k<\kappa}$ with $\alpha'_k := (\bar{a}^k[l\bar{u}^k\bar{v}^k])_{\bar{u}^k<l<\bar{v}^k}$ violates (2).

(1) ⇒ (3) First, we consider the case where κ is infinite. Suppose that there exist a tuple $\bar{b}$ realising Φ and a family $(\alpha_k)_{k \in K}$ of infinite mutually indiscernible sequences over U such that, for every $K_0 \subseteq K$ of size $|K_0| < \kappa$, the subfamily $(\alpha_k)_{k \in K \setminus K_0}$ is not mutually indiscernible over $U \cup \bar{b}$. By induction on $i < \kappa$, we choose an index $k_i \in K$ and a finite subset $s_i \subseteq K$ as follows. Suppose that we have already defined k_j and s_j, for all $j < i$. Set $S := k[<i] \cup s[<i]$. Then $|S| < \kappa$ and, by assumption, we can find an index $k_i \in K \setminus S$ such that the sequence α_{k_i} is not indiscernible over $U \cup \bar{b} \cup \alpha[K \setminus (S \cup \{k_i\})]$. Therefore, we can find a finite subset $s_i \subseteq K \setminus (S \cup \{k_i\})$ such that α_{k_i} is not indiscernible over $U \cup \bar{b} \cup \alpha[s_i]$.

Having defined $(k_i)_{i<\kappa}$ and $(s_i)_{i<\kappa}$, we set

$$C := \bigcup_{i<\kappa} \alpha[s_i]\,.$$

Then the family $(\alpha_{k_i})_{i<\kappa}$ is mutually indiscernible over $U \cup C$, but no sequence α_{k_i} is indiscernible over $U \cup C \cup \bar{b}$. Consequently, it follows by Lemma 4.9 that $\mathrm{rk_{dp}}(\Phi/U) = \mathrm{rk_{dp}}(\Phi/UC) > \kappa$.

It remains to consider the case where $\kappa = n + 1$ is finite. Let $(\alpha_k)_{k<\lambda}$ be a family of infinite mutually indiscernible sequences over U and let $\bar{b}$ be a tuple realising Φ. We construct the desired subset $K_0 \subseteq \lambda$ by induction on λ.

If $\lambda \leq n$, we can take $K_0 := \lambda$. Hence, suppose that $\lambda = n + m + 1 < \omega$ and that we have already proved the claim for families of size $n + m$. Extending the sequences α_k if necessary, we may assume that they do not have a last element. By induction on $k < \lambda$, we choose a sequence β_k indexed by $\mathbb{Z}$ such that the sequence β^{op} with the reversed ordering is generated by the type $\mathfrak{p}_k := \mathrm{CF}(\alpha_k)$ over $U\bar{b}\alpha[<\lambda]\beta[<k]$. By Lemma 3.3, the family $(\alpha_k^+)_{k<\lambda}$ with $\alpha_k^+ := \alpha_k\beta_k$ is mutually indiscernible over U. As $(\alpha_k)_{k<\lambda}$ is mutually indiscernible over $U\beta[<\lambda]$ and

$$\mathrm{rk_{dp}}(\Phi/U\beta[<\lambda]) = \mathrm{rk_{dp}}(\Phi/U) \leq n + 1 \leq \lambda\,,$$

we can find an index $k_0 < \lambda$ such that α_{k_0} is indiscernible over $U\beta[<\lambda]\bar{b}$. Furthermore, since $(\alpha_k^+)_{k\in\lambda\smallsetminus\{k_0\}}$ is mutually indiscernible over $U\alpha_{k_0}$, we can use the inductive hypothesis to find a set $H \subseteq \lambda \smallsetminus \{k_0\}$ of size $|H| \leq n$ such that $(\alpha_k^+)_{k\in\lambda\smallsetminus(H\cup\{k_0\})}$ is mutually indiscernible over $U\alpha_{k_0}\bar{b}$. If the sequence α_{k_0} is indiscernible over $U\bar{b}\alpha[\lambda \smallsetminus (H \cup \{k_0\})]$, then $(\alpha_k)_{k\in\lambda\smallsetminus H}$ is mutually indiscernible over $U\bar{b}$ and we are done.

For a contradiction, suppose otherwise. Then there is some finite set $C \subseteq U\bar{b}\alpha[\lambda \smallsetminus (H \cup \{k_0\})]$ such that α_{k_0} is not indiscernible over C. Let $\bar{c}_k$ be an enumeration of $C \cap \alpha_k$ and set $C_0 := C \cap (U \cup \bar{b})$. Since $(\alpha_k^+)_{k\in\lambda\smallsetminus(H\cup\{k_0\})}$ is mutually indiscernible over $U\bar{b}\alpha_{k_0}$, we can find, for every $k \in \lambda \smallsetminus (H \cup \{k_0\})$, a tuple $\bar{d}_k \subseteq \beta_k$ such that

$$\bar{d}_k \equiv_{U\bar{b}\alpha_{k_0}\alpha^+[\lambda\smallsetminus(H\cup\{k,k_0\})]} \bar{c}_k\,.$$

It follows that α_{k_0} is not indiscernible over $C_0 \cup \bigcup_k \bar{d}_k \subseteq U\bar{b}\beta[<\lambda]$. This contradicts our choice of k_0.

It remains to consider the case where λ is an infinite cardinal. For every ordinal $\gamma < \lambda$, we can use the inductive hypothesis to find a set $H_\gamma \subseteq \gamma$ of size $|H_\gamma| \leq n$ such that the family $(\alpha_k)_{k \in \gamma \setminus H_\gamma}$ is mutually indiscernible over $U\bar{b}$. We will construct finite sets $K_0, \dots, K_{n-1} \subseteq \lambda$ and indices $s_0, \dots, s_{n-1} < \lambda$ as follows. Suppose that we have already chosen $K_0, \dots, K_{i-1}$ and $s_0, \dots, s_{i-1}$ such that

$$\{s_0, \dots, s_{i-1}\} \subseteq H_\gamma \,, \quad \text{for arbitrarily large } \gamma \,.$$

If the family $(\alpha_k)_{k \in \lambda \setminus \{s_0, \dots, s_{i-1}\}}$ is mutually indiscernible over $U\bar{b}$, we are done. Otherwise, there exists a finite set $K_i \subseteq \lambda \setminus \{s_0, \dots, s_{i-1}\}$ such that $(\alpha_k)_{k \in K_i}$ is not mutually indiscernible over $U\bar{b}$. By choice of the sets H_γ, we have $K_i \cap H_\gamma \neq \emptyset$, for all $\gamma < \lambda$. As the set K_i is finite, there is therefore some index $s_i \in K_i$ such that

$$\{s_0, \dots, s_{i-1}, s_i\} \subseteq H_\gamma \,, \quad \text{for arbitrarily large } \gamma \,.$$

Having constructed $s_0, \dots, s_{n-1}$ as above, it follows that there are arbitrarily large γ such that $H_\gamma = \{s_0, \dots, s_{n-1}\}$. Hence, there are arbitrarily large $\gamma < \lambda$ such that the family $(\alpha_k)_{k \in \gamma \setminus \{s_0, \dots, s_{n-1}\}}$ is mutually indiscernible over $U\bar{b}$. This implies that $(\alpha_k)_{k \in \lambda \setminus \{s_0, \dots, s_{n-1}\}}$ is also mutually indiscernible over $U\bar{b}$. □

We can use this characterisation to give a straightforward proof that the dp-rank is sub-additive.

Proposition 4.12. $\mathrm{rk}_{\mathrm{dp}}(\bar{a}\bar{b}/U) \oplus 1 \leq \mathrm{rk}_{\mathrm{dp}}(\bar{a}/U) \oplus \mathrm{rk}_{\mathrm{dp}}(\bar{b}/U\bar{a})$.

Proof. Let $\kappa := \mathrm{rk}_{\mathrm{dp}}(\bar{a}/U)$ and $\lambda := \mathrm{rk}_{\mathrm{dp}}(\bar{b}/U\bar{a})$. To show that

$$\mathrm{rk}_{\mathrm{dp}}(\bar{a}\bar{b}/U) \oplus 1 \leq \kappa \oplus \lambda \,,$$

consider a tuple $\bar{a}'\bar{b}' \equiv_U \bar{a}\bar{b}$ and a family $(\alpha_k)_{k \in K}$ of infinite mutually indiscernible sequences over U. According to Proposition 4.11 (3), it is sufficient to find a subset $K' \subseteq K$ of size $|K'| \oplus 1 < \kappa \oplus \lambda$ such that $(\alpha_k)_{k \in K \setminus K'}$ is mutually indiscernible over $U\bar{a}'\bar{b}'$.

Note that invariance implies that $\mathrm{rk_{dp}}(\bar b'/U\bar a') = \mathrm{rk_{dp}}(\bar b/U\bar a)$. We use the characterisation in Proposition 4.11 (3) two times: first, to find a subset $K_0 \subseteq K$ of size $|K_0| < \kappa$ such that $(\alpha_k)_{k \in K \setminus K_0}$ is mutually indiscernible over $U \cup \bar a'$; and then, to find a subset $K_1 \subseteq K \setminus K_0$ of size $|K_1| < \lambda$ such that $(\alpha_k)_{k \in K \setminus (K_0 \cup K_1)}$ is mutually indiscernible over $U \cup \bar a' \bar b'$. Since $|K_0 \cup K_1| \oplus 1 < \kappa \oplus \lambda$, the claim follows. □

The dp-rank is well-behaved in theories without the independence properties. In particular, it always exists.

Theorem 4.13. *Let T be a complete first-order theory. The following statements are equivalent:*

(1) *T does not have the independence property.*

(2) $\mathrm{rk_{dp}}(\Phi/U) \leq |T|^+ \oplus |\bar x|^+$, *for every partial type $\Phi(\bar x)$ with variables $\bar x$ and every set U.*

(3) $\mathrm{rk_{dp}}(\Phi/U) < \infty$, *for every partial type $\Phi(\bar x)$ and every set U.*

Proof. (1) ⇒ (2) Let $\bar b$ be a tuple realising Φ and $(\alpha_k)_{k<\kappa}$ a family of infinite mutually indiscernible sequences over U of size $\kappa := |T|^+ \oplus |\bar x|^+$. By Proposition 4.7, there exists a set $K_0 \subseteq \kappa$ of size $|K_0| \leq |T| \oplus |\bar b| < \kappa$ such that the family $(\alpha_k)_{k \in \kappa \setminus K_0}$ is mutually indiscernible over $U \cup \bar b$. Fix $k \in \kappa \setminus K_0 \neq \varnothing$. Then α_k is indiscernible over $U \cup \bar b$.

(2) ⇒ (3) is trivial.

(3) ⇒ (1) Let κ be an infinite cardinal and let $I := \omega \times \kappa$, ordered lexicographically. Suppose that there exists a formula $\varphi(\bar x; \bar y)$ with the independence property. By compactness, there exists a tuple $\bar b$ and an indiscernible sequence $(\bar a_i)_{i \in I}$ such that

$$\mathbb{M} \vDash \varphi(\bar a_i; \bar b) \quad \text{iff} \quad i \in \{0\} \times \kappa\,.$$

By Lemma 4.2, the sequences $\alpha_i := (\bar a_{\langle i,k \rangle})_{k<\kappa}$ are mutually indiscernible over $\varnothing$, but none of them is indiscernible over $\bar b$. This implies that $\mathrm{rk_{dp}}(\bar b/\varnothing) > \kappa$. □

F5. Theories without the array property

1. The array property

In this chapter we consider a property of formulae that generalises both the tree property and the independence property. It is based on families of tuples with a two-dimensional index set.

Definition 1.1. Let γ, δ be ordinals and $\alpha = (\bar{a}_{ij})_{i<\gamma, j<\delta}$ a family of tuples.

(a) The *i-th row* of α is the sequence $\alpha^i := (\bar{a}_{ij})_{j<\delta}$, its *j-th column* is $\alpha_j := (\bar{a}_{ij})_{i<\gamma}$, and its *diagonal* is $(\bar{a}_{ii})_{i<\min\{\gamma,\delta\}}$.

(b) For $I \subseteq \gamma$ and $J \subseteq \delta$, we set

$$\bar{a}[I;J] := \bigcup_{i \in I, j \in J} \bar{a}_{ij}\,.$$

(c) α is *biindiscernible* over a set U if the sequence $(\alpha^i)_{i<\gamma}$ of rows and the sequence $(\alpha_j)_{j<\delta}$ of columns are both indiscernible over U. We call α *strongly indiscernible* over U if, in addition, the sequence $(\alpha^i)_{i<\gamma}$ of rows is mutually indiscernible over U.

We start with presenting two methods to construct strongly indiscernible families.

Lemma 1.2. *Let $\alpha = (\bar{a}_{ij})_{i<\gamma, j<\delta}$ be a family such that the sequence of rows $(\alpha^i)_{i<\gamma}$ is both mutually indiscernible over U and indiscernible over U. Then α is strongly indiscernible.*

Proof. It remains to prove that the sequence of columns $(\alpha_j)_{j<\delta}$ is indiscernible over U. Fix indices $\bar{l} \in [\gamma]^m$ and $\bar{\imath}, \bar{\jmath} \in [\delta]^n$. We claim that

$$\bar{a}[\bar{l};\bar{\imath}] \equiv_U \bar{a}[\bar{l};\bar{\jmath}] .$$

Let $s < m$. Since α^{l_s} is indiscernible over $U \cup \bar{a}[\gamma \setminus \{l_s\};\delta]$, we have

$$\bar{a}[l_s;\bar{\imath}] \equiv_{U\bar{a}[\gamma \setminus \{l_s\};\delta]} \bar{a}[l_s;\bar{\jmath}] ,$$

which implies that

$$\begin{aligned} &\bar{a}[l_0 \dots l_{s-1};\bar{\imath}]\bar{a}[l_s;\bar{\imath}]\bar{a}[l_{s+1} \dots l_{m-1};\bar{\jmath}] \\ \equiv_U\ &\bar{a}[l_0 \dots l_{s-1};\bar{\imath}]\bar{a}[l_s;\bar{\jmath}]\bar{a}[l_{s+1} \dots l_{m-1};\bar{\jmath}] . \end{aligned}$$

By transitivity, it follows that $\bar{a}[\bar{l};\bar{\imath}] \equiv_U \bar{a}[\bar{l};\bar{\jmath}]$. □

The next remark generalises Lemma F4.4.2.

Lemma 1.3. *Let $\beta = (\bar{b}_i)_{i<\delta\gamma}$ be an indiscernible sequence over U and define*

$$\alpha = (\bar{a}_{ij})_{i<\gamma, j<\delta} \quad \text{by} \quad \bar{a}_{ij} := \bar{b}_{\delta i+j} .$$

Then α is strongly indiscernible over U.

Proof. Note that the i-th row

$$\alpha^i = (\bar{a}_{ij})_{j<\delta} = (\bar{b}_{\delta i+j})_{j<\delta}$$

is indiscernible over

$$U \cup \bar{b}[<\delta i] \cup \bar{b}[\geq\delta(i+1)] = U \cup \bigcup_{l \neq i} \alpha^l .$$

By Lemma 1.2, it is therefore sufficient to show that the sequence of rows $(\alpha^i)_{i<\gamma}$ is indiscernible over U. Fix indices $\bar{\imath}, \bar{\jmath} \in [\gamma]^m$ and $\bar{l} \in [\delta]^n$. Then

$$(\bar{b}_{\delta i_s + l_t})_{s<m, t<n} \equiv_U (\bar{b}_{\delta j_s + l_t})_{s<m, t<n}$$

implies that $\bar{a}[\bar{\imath};\bar{l}] \equiv_U \bar{a}[\bar{\jmath};\bar{l}]$. □

Using two-dimensional families we can introduce the array property, which generalises the independence property and the tree property.

Definition 1.4. Let $\varphi(\bar{x};\bar{y})$ be a formula and $k<\omega$.

(a) We say that $\varphi(\bar{x};\bar{y})$ is *consistent over* a family $\beta=(\bar{b}_i)_{i\in I}$ of tuples if the set $\{\,\varphi(\bar{x};\bar{b}_i)\mid i\in I\,\}$ is consistent. Similarly, we say that φ is *inconsistent* or *k-inconsistent* over β, it the above set is, respectively, inconsistent or k-inconsistent.

(b) A *k-array* for φ is a family $\alpha=(\bar{a}_{ij})_{i,j<\omega}$ of tuples such that

- φ is k-inconsistent over each row $\alpha^i=(\bar{a}_{ij})_{j<\omega}$, $i<\omega$, and
- for every function $\eta:\omega\to\omega$, φ is consistent over the sequence $(\bar{a}_{i\eta(i)})_{i<\omega}$.

(c) We say that φ has the *array property,* or the *tree property of the second kind,* if, for some $k<\omega$, there exists a k-array for φ. A theory T has the *array property* if some formula does.

Let us first note that we can choose a k-array always to be strongly indiscernible.

Lemma 1.5. *A formula $\varphi(\bar{x};\bar{y})$ has a k-array if, and only if, it has a strongly indiscernible k-array.*

Proof. $(\Leftarrow)$ is trivial. For $(\Rightarrow)$, suppose that the formula φ has a k-array $\alpha=(\bar{a}_{ij})_{i,j<\omega}$ with rows $(\alpha^i)_{i<\omega}$. By induction on i, we use Proposition E5.3.6 to choose an indiscernible sequence $\beta^i=(\bar{b}_{ij})_{j<\omega}$ over $\alpha[>i]\beta[<i]$ such that

$$\mathrm{Av}(\beta^i/\alpha[>i]\beta[<i])\supseteq\mathrm{Av}(\alpha^i/\alpha[>i]\beta[<i])\,.$$

According to Lemma F4.4.3, the family $(\beta^i)_{i<\omega}$ is mutually indiscernible. Furthermore, the k-inconsistency of $\{\,\varphi(\bar{x};\bar{a}_{ij})\mid j<\omega\,\}$ implies the k-inconsistency of $\{\,\varphi(\bar{x};\bar{b}_{ij})\mid j<\omega\,\}$.

To show that all sets of the form $\{\,\varphi(\bar{x};\bar{b}_{i\eta(i)})\mid i<\omega\,\}$ are consistent, it is sufficient by compactness to prove that, for every $n<\omega$ and every

$\eta : [n] \to \omega$, there exists some tuple $\bar{c}$ with

$$\mathbb{M} \vDash \bigwedge_{i<n} \varphi(\bar{c}; \bar{b}_{i\eta(i)}) .$$

To do so, we prove by induction on $m \leq n$, that, for every function $\eta : [n] \to \omega$, there is some tuple $\bar{c}$ with

$$\mathbb{M} \vDash \bigwedge_{i<m} \varphi(\bar{c}; \bar{b}_{i\eta(i)}) \wedge \bigwedge_{m \leq i<n} \varphi(\bar{c}; \bar{a}_{i\eta(i)}) .$$

For $m = 0$, the existence of $\bar{c}$ follows by choice of the $\bar{a}_{ij}$. For the inductive step, suppose that, for every $\eta : [n] \to \omega$, we have already found a tuple $\bar{c}$ such that

$$\mathbb{M} \vDash \psi_\eta(\bar{c}; \bar{a}_{m\eta(m)}) ,$$

where

$$\psi_\eta(\bar{x}; \bar{y}) := \bigwedge_{i<m} \varphi(\bar{x}; \bar{b}_{i\eta(i)}) \wedge \varphi(\bar{x}; \bar{y}) \wedge \bigwedge_{m<i<n} \varphi(\bar{c}; \bar{a}_{i\eta(i)}) .$$

For a given $j < \omega$, we consider the function $\eta' : [n] \to \omega$ with $\eta'(m) := j$ and $\eta'(i) := \eta(i)$, for $i \neq m$. Then $\psi_{\eta'} = \psi_\eta$ and the inductive hypothesis implies that

$$\mathbb{M} \vDash \exists \bar{x} \psi_\eta(\bar{x}; \bar{a}_{mj}) , \quad \text{for every } j < \omega .$$

Hence,

$$\exists \bar{x} \psi_\eta(\bar{x}; \bar{y}) \in \mathrm{Av}(\alpha_m / \alpha[>m]\beta[<m]) \subseteq \mathrm{Av}(\beta_m / \alpha[>m]\beta[<m]) .$$

Consequently, there is some tuple $\bar{c}$ such that

$$\mathbb{M} \vDash \psi_\eta(\bar{c}; \bar{b}_{m\eta(m)}) .$$

We have shown that the family $\beta = (\beta^i)_{i<\omega}$ has all of the desired properties except possibly for biindiscernibility. To conclude the proof, we can use Proposition E5.3.6 to choose an indiscernible sequence $\beta' = (\beta'^i)_{i<\omega}$ such that

$$\mathrm{Av}(\beta'/\varnothing) \supseteq \mathrm{Av}(\beta/\varnothing)\,.$$

By Lemma 1.2, it follows that β' is strongly indiscernible. □

Next we show that the class of theories without the array property generalises both the simple theories and those without the independence property. We start by proving this implication for formulae.

Proposition 1.6. *Every formula with the array property has the tree property and the independence property.*

Proof. Suppose that φ has a k-array $(\bar a_{ij})_{i,j<\omega}$. We start by showing that φ has the tree property. We set

$$\bar c_{\langle\rangle} := \bar a_{00} \quad \text{and} \quad \bar c_w := \bar a_{n w_{n-1}}\,, \qquad \text{for } w \in \omega^n,\ n > 0\,.$$

Then the family $(\bar c_w)_{w\in\omega^{<\omega}}$ is a witness for the tree property of φ since

- for every $\eta \in \omega^\omega$, the set

  $$\{\, \varphi(\bar x; \bar c_w) \mid w \prec \eta \,\} = \{\varphi(\bar x; \bar a_{00})\} \cup \{\, \varphi(\bar x; \bar a_{(n+1)\eta(n)}) \mid n < \omega \,\}$$

  is consistent and
- for every $w \in \omega^{<\omega}$ of length $n := |w|$, the set

  $$\{\, \varphi(\bar x; \bar c_{wi}) \mid i < \omega \,\} = \{\, \varphi(\bar x; \bar a_{(n+1)i}) \mid i < \omega \,\}$$

  is k-inconsistent.

It remains to check the independence property. By Lemma 1.5, we may assume that α is strongly indiscernible. Let m be the maximal number such that, for some infinite subset $I \subseteq \omega$, there exists a tuple $\bar c$ with

$$\mathbb{M} \vDash \varphi(\bar c; \bar a_{ij})\,, \quad \text{for all } i \in I \text{ and } j < m\,.$$

As φ is k-inconsistent over every column, we have $m < k$. Furthermore, it follows by maximality of m that there exists an infinite subset $J \subseteq I$ such that

$$\mathbb{M} \vDash \neg\varphi(\bar{c}; \bar{a}_{im}), \quad \text{for all } i \in J.$$

Choose a strictly increasing function $g : \omega \to J$ and define $\eta : \omega \to \omega$ by

$$\eta(i) := \begin{cases} 0 & \text{if } i \text{ is even,} \\ m & \text{if } i \text{ is odd.} \end{cases}$$

It follows that

$$\mathbb{M} \vDash \varphi(\bar{c}; \bar{a}_{g(i)\eta(i)}) \quad \text{iff} \quad i \text{ is even.}$$

Since, according to Lemma F4.4.4, the sequence $(\bar{a}_{g(i)\eta(i)})_{i<\omega}$ is indiscernible, it follows by Proposition E5.4.2 that φ has the independence property. □

Thus, theories without the array property generalise both simple theories and theories without the independence property.

Corollary 1.7. *Let T be a complete first-order theory with the array property. Then T is not simple and it has the independence property.*

Our next goal is an alternative characterisation of the array property.

Definition 1.8. Let $\alpha = (\bar{a}_{ij})_{i<\gamma, j<\delta}$ be a family of tuples.

(a) The *transpose* of α is $\alpha^T := (\bar{a}_{ji})_{i<\delta, j<\gamma}$.

(b) The *column k-condensation* of α is the family $\alpha^{(k)} := (\bar{a}'_{ij})_{i<\gamma, j<\delta}$ with

$$\bar{a}'_{ij} := \bar{a}[k * i; j] \quad \text{where} \quad k * i := \langle ki, ki+1, \dots ki+k-1 \rangle.$$

For $\bar{\imath} \in [\gamma]^n$, we similarly set

$$k * \bar{\imath} := (k * i_0) \dots (k * i_{n-1}).$$

(c) For a formula $\varphi(\bar{x};\bar{y})$, we set

$$\varphi^{(k)}(\bar{x};\bar{y}_0\dots\bar{y}_{k-1}) := \bigwedge_{i<k}\varphi(\bar{x};\bar{y}_i)\,.$$

Remark. Note that a formula φ is consistent over a column α_j if, and only if, $\varphi^{(k)}$ is consistent over the condensed column $\alpha_j^{(k)}$.

Lemma 1.9. *Let $\alpha = (\bar{a}_{ij})_{i<\gamma,j<\delta}$ be a family of tuples and $k<\omega$.*

(a) *If α is biindiscernible over U, then so are α^T and $\alpha^{(k)}$.*

(b) *If α is strongly indiscernible over U, then so is $\alpha^{(k)}$.*

Proof. (a) Clearly, if α is biindiscernible over U, so is α^T. To see that the column k-condensation $\alpha^{(k)} = (\bar{b}_{ij})_{i<\gamma,j<\delta}$ is also biindiscernible over U, note that, for all tuples of indices $\bar{\imath}, \bar{\jmath}, \bar{l}$,

$$\begin{aligned} &\bar{a}[k*\bar{l};\bar{\imath}] \equiv_U \bar{a}[k*\bar{l};\bar{\jmath}] \quad \text{implies} \quad \bar{b}[\bar{l};\bar{\imath}] \equiv_U \bar{b}[\bar{l};\bar{\jmath}]\,,\\ \text{and}\quad &\bar{a}[k*\bar{\imath};\bar{l}] \equiv_U \bar{a}[k*\bar{\jmath};\bar{l}] \quad \text{implies} \quad \bar{b}[\bar{\imath};\bar{l}] \equiv_U \bar{b}[\bar{\jmath};\bar{l}]\,.\end{aligned}$$

(b) Suppose that α is strongly indiscernible over U. It follows by (a) that the column k-condensation $\beta := \alpha^{(k)} = (\bar{b}_{ij})_{i<\gamma,j<\delta}$ is biindiscernible over U. To prove that the family $(\beta^i)_{i<\gamma}$ of rows is mutually indiscernible over U, consider indices $\bar{\imath}, \bar{\jmath} \in [\delta]^n$ and set

$$B_l := U \cup \bar{b}[\gamma \smallsetminus \{l\};\delta]\,.$$

Then $B_l = U \cup \bar{a}[\gamma \smallsetminus k*l;\delta]$ and

$$\bar{a}[k*l;\bar{\imath}] \equiv_{UB_l} \bar{a}[k*l;\bar{\jmath}] \quad \text{implies} \quad \bar{b}[l;\bar{\imath}] \equiv_{UB_l} \bar{b}[l;\bar{\jmath}]\,.$$

Hence, β^l is indiscernible over $U \cup B_l$. □

Lemma 1.10. *Let T be a theory without the array property, $\varphi(\bar{x};\bar{y})$ a formula, and $\alpha = (\bar{a}_{ij})_{i,j<\omega}$ a biindiscernible family.*

(a) *Suppose that α is strongly indiscernible. If φ is consistent over the 0-th column $\alpha_0 = (\bar{a}_{i0})_{i<\omega}$, it is consistent over all of α.*

(b) *If φ is consistent over the diagonal $(\bar{a}_{ii})_{i<\omega}$ of α, the formula $\varphi^{(k)}$ is consistent over the diagonal $(\bar{b}_{ii})_{i<\omega}$ of the column k-condensation $\alpha^{(k)} = (\bar{b}_{ij})_{i,j<\omega}$.*

Proof. (a) By compactness, it is sufficient to prove that, for every $k < \omega$, φ is consistent over $(\bar{a}_{ij})_{i<k,j<\omega}$. Fix $k < \omega$. By Lemma 1.9, the column k-condensation $\alpha^{(k)} = (\bar{b}_{ij})_{i,j<\omega}$ is also strongly indiscernible. Furthermore, as φ is consistent over $(\bar{a}_{i0})_{i<\omega}$ and $\bar{a}[\omega; 0] = \bar{b}[\omega; 0]$, it follows that $\varphi^{(k)}$ is consistent over $(\bar{b}_{i0})_{i<\omega}$. By Lemma F4.4.4, this implies that $\varphi^{(k)}$ is consistent over $(\bar{b}_{i,\eta(i)})_{i<\omega}$, for every $\eta : \omega \to \omega$. As $\varphi^{(k)}$ does not have the array property, there therefore exists some $i < \omega$ such that $\varphi^{(k)}$ is consistent over $(\bar{b}_{ij})_{j<\omega}$. By indiscernibility, it follows that it is also consistent over $(\bar{b}_{0j})_{j<\omega}$. This implies that φ is consistent over $(\bar{a}_{ij})_{i<k,j<\omega}$.

(b) We can use Corollary E5.3.10 to extend the sequence $(\alpha^i)_{i<\omega}$ of rows to an indiscernible sequence $(\alpha^i)_{i<\omega^2}$ of length ω^2. Suppose that $\alpha^i = (\bar{a}_{ij})_{j<\omega}$ and set $\bar{c}_{ij} := \bar{a}_{\omega i+j,i}$. By mutual indiscernibility of $(\alpha^i)_i$, we have

$$(\bar{c}_{ij})_{i,j<\omega} = (\bar{a}_{\omega i+j,i})_{i,j<\omega} \equiv (\bar{a}_{\omega i+j,0})_{i,j<\omega} .$$

Furthermore, according to Lemma 1.3, the latter family is strongly indiscernible. Hence, so is $(\bar{c}_{ij})_{i,j<\omega}$. Furthermore, by biindiscernibility of α, we have

$$(\bar{c}_{i0})_{i<\omega} = (\bar{a}_{\omega i,i})_{i<\omega} \equiv (\bar{a}_{ii})_{i<\omega} .$$

Consequently, the consistency of φ over $(\bar{a}_{ii})_{i<\omega}$ implies the consistency of φ over $(\bar{c}_{i0})_{i<\omega}$. It therefore follows by (a) that φ is consistent over $(\bar{c}_{ij})_{i,j<\omega}$. Finally, by biindiscernibility of α, we have

$$(\bar{c}_{ij})_{i<\omega,j<k} = (\bar{a}_{\omega i+j,i})_{i<\omega,j<k} \equiv (\bar{a}_{ki+j,i})_{i<\omega,j<k} .$$

Consequently, φ is consistent over $(\bar{a}_{ki+j,i})_{i<\omega,j<k}$, which implies that $\varphi^{(k)}$ is consistent over $(\bar{b}_{ii})_{i<\omega}$. ☐

Proposition 1.11. *A theory T does not have the array property if, and only if, for every biindiscernible family $\alpha = (\bar{a}_{ij})_{i,j<\omega}$, the consistency of a formula $\varphi(\bar{x};\bar{y})$ over the diagonal $(\bar{a}_{ii})_{i<\omega}$ implies the consistency of φ over α.*

Proof. ($\Leftarrow$) Suppose that some formula φ has a k-array. By Lemma 1.5, we can choose this k-array to be biindiscernible. It follows that φ is consistent over the diagonal of α, but not over α itself.

($\Rightarrow$) Suppose that T does not have the array property and let α be a biindiscernible family such that φ is consistent over the diagonal of α. By compactness, it is sufficient to prove that, for every $k < \omega$, φ is consistent over $(\bar{a}_{ij})_{i,j<k}$. By Lemma 1.10, $\varphi^{(k)}$ is consistent over the diagonal of $\alpha^{(k)}$. Since $\beta := (\alpha^{(k)})^T$ has the same diagonal, it follows by another application of Lemma 1.10 that $(\varphi^{(k)})^{(k)}$ is consistent over the diagonal of $\beta^{(k)} = (\bar{b}_{ij})_{i,j<\omega}$. In particular, $(\varphi^{(k)})^{(k)}(\bar{x};\bar{b}_{00})$ is consistent. Since $\bar{b}_{00} = (\bar{a}_{ij})_{i,j<k}$ the claim follows. □

As an application, let us show that, in theories without the array property, we can characterise dividing in the following way.

Definition 1.12. A formula $\varphi(\bar{x};\bar{b})$ *array-divides* over a set U if there exists a biindiscernible family $\beta = (\bar{b}_{ij})_{i,j<\omega}$ over U such that $\bar{b}_{00} = \bar{b}$ and φ is inconsistent over β.

Lemma 1.13. *Every formula that divides over U also array-divides over U.*

Proof. Suppose that $\varphi(\bar{x};\bar{b})$ divides over U. Then there exists an indiscernible sequence $\beta = (\bar{b}_i)_{i<\omega}$ over U such that $\bar{b}_0 = \bar{b}$ and φ is k-inconsistent over β. By Corollary E5.3.10, we can extend β to an indiscernible sequence $\beta = (\bar{b}_i)_{i<\omega^2}$ over U of length ω^2. Set $\alpha := (\bar{a}_{ij})_{i,j<\omega}$ with $\bar{a}_{ij} := \bar{b}_{\omega i+j}$. By Lemma 1.3, it follows that α is biindiscernible over U. Furthermore, $\bar{a}_{00} = \bar{b}$ and φ is inconsistent over α. Hence, $\varphi(\bar{x};\bar{b})$ array-divides over U. □

Corollary 1.14. *Let T be a theory without the array property. A formula $\varphi(\bar{x};\bar{b})$ divides over U if, and only if, it array-divides over U.*

Proof. We have proved the implication ($\Rightarrow$) already in Lemma 1.13. For ($\Leftarrow$), suppose that $\varphi(\bar{x};\bar{b})$ does not divide over U. To show that it does not array-divide over U, consider a biindiscernible family $\beta = (\bar{b}_{ij})_{i,j<\omega}$ over U such that $\bar{b}_{00} = \bar{b}$. Since the diagonal $(\bar{b}_{ii})_{i<\omega}$ is indiscernible over U, the fact that $\varphi(\bar{x};\bar{b}_{00})$ does not divide over U implies that φ is consistent over $(\bar{b}_{ii})_{i<\omega}$. By Proposition 1.11, it follows that φ is consistent over β. $\square$

2. *Forking and dividing*

Extension bases

Our first question regarding theories without the array property is over which base sets forking and dividing coincide. For this to be the case, the forking relation should have all the properties of the dividing relation. Therefore, we start by collecting some of them.

Definition 2.1. Let $\sqrt[0]{}$ and $\sqrt[1]{}$ be preforking relations and $U \subseteq \mathbb{M}$. We say that *$\sqrt[0]{}$-forking implies $\sqrt[1]{}$-forking over U* if every formula that $\sqrt[0]{}$-forks over U also $\sqrt[1]{}$-forks over U. Similarly, we say that $\sqrt[0]{}$ and $\sqrt[1]{}$ *coincide over U* if we have implications in both directions.

Definition 2.2. Let $\sqrt{}$ be an independence relation and $U \subseteq \mathbb{M}$ a set.

(a) We say that $\sqrt{}$ has *left extension over a set U* if it satisfies the following axiom:

(LEXT) *Left Extension.* If $A_0 \sqrt{}_U B$ and $A_0 \subseteq A_1$ then there is some B' with

$$B' \equiv_{UA_0} B \quad \text{and} \quad A_1 \sqrt{}_U B' .$$

(b) U is a *$\sqrt{}$-base* if $A \sqrt{}_U U$, for all $A \subseteq \mathbb{M}$.

(c) U is a *$\sqrt{}$-extension base* if U is a $\sqrt{}$-base and $\sqrt{}$ has left extension over U.

Let us first note that $\surd$-bases do exist.

Lemma 2.3. (a) *Every set is a $\surd$-base if $\surd$ is one of the relations $\overset{\mathrm{ls}}{\surd}$, $\overset{\mathrm{s}}{\surd}$, or $\overset{\mathrm{d}}{\surd}$.*

(b) *$\overset{\mathrm{u}}{\surd}$ has left extension over every set.*

(c) *Every model is a $\overset{\mathrm{u}}{\surd}$-extension base.*

(d) *Every model is a $\surd$-base for all preforking relations $\surd$.*

Proof. (a) It follows immediately form the definition that $A \overset{\mathrm{s}}{\surd}_U U$, for all sets A and U. As we have seen in Corollary F4.2.22 that $\overset{\mathrm{s}}{\surd} \subseteq \overset{\mathrm{ls}}{\surd}$ it follows that $A \overset{\mathrm{ls}}{\surd}_U U$ as well. For $\overset{\mathrm{d}}{\surd}$, the claim follows immediately from the characterisation in Lemma F3.1.3.

(b) Suppose that $A \overset{\mathrm{u}}{\surd}_U \bar{b}$ and let $C \subseteq \mathbb{M}$. We have to show that there is some tuple $\bar{b}' \equiv_{UA} \bar{b}$ with $AC \overset{\mathrm{u}}{\surd}_U \bar{b}'$. In other words, we have to show that the set

$$\begin{aligned} \Phi(\bar{x}) := \operatorname{tp}(\bar{b}/UA) \\ \cup \{\, \varphi(\bar{x};\bar{c}) \mid \bar{c} \subseteq UAC \text{ and } \varphi(\bar{x};\bar{y}) \text{ a formula over } U \\ \text{such that } \mathbb{M} \vDash \varphi(\bar{b};\bar{d}) \text{ for all } \bar{d} \subseteq U \,\} \end{aligned}$$

is satisfiable. For a contradiction, suppose that Φ is inconsistent. Then we can find a formula $\psi(\bar{x};\bar{a}) \in \operatorname{tp}(\bar{b}/UA)$, finitely many formulae $\varphi_i(\bar{x};\bar{y}_i)$ over U, and parameters $\bar{c}_i \subseteq UAC$ such that

$$\psi(\bar{x};\bar{a}) \vDash \bigvee_{i<n} \neg\varphi_i(\bar{x};\bar{c}_i) \quad \text{and} \quad \mathbb{M} \vDash \varphi_i(\bar{b};\bar{d}) \text{ for all } \bar{d} \subseteq U\,.$$

W.l.o.g. we may assume that the parameters $\bar{c}_i$ are all of the form $\bar{c}_i = \bar{a}\bar{c}$, for some tuple $\bar{c} \subseteq UAC$ that is disjoint from $\bar{a}$. Hence,

$$\psi(\bar{x};\bar{a}) \vDash \bigvee_{i<n} \neg\varphi_i(\bar{x};\bar{a},\bar{c})$$

and it follows by the Coincidence Lemma that

$$\psi(\bar{x};\bar{y}) \vDash \forall \bar{z} \bigvee_{i<n} \neg\varphi_i(\bar{x};\bar{y},\bar{z})\,.$$

Since $A \mathrel{\overset{\mathrm{u}}{\surd}}_U \bar{b}$, there is some tuple $\bar{a}' \subseteq U$ such that $\mathbb{M} \vDash \psi(\bar{b};\bar{a}')$. Fix some tuple $\bar{d} \subseteq U$. Then it follows by the above implication that

$$\mathbb{M} \vDash \bigvee_{i<n} \neg\varphi_i(\bar{b};\bar{a}',\bar{d})\,.$$

Hence, there is some index i with $\mathbb{M} \vDash \neg\varphi_i(\bar{b};\bar{a}',\bar{d})$. As $\bar{a}'\bar{d} \subseteq U$, this contradicts our choice of φ_i.

(c) We have already seen in Lemma F2.3.15 that each model is a $\overset{\mathrm{u}}{\surd}$-base. Hence, the claim follows by (b).

(d) It follows by (c) that every model $\mathfrak{M}$ is a $\overset{\mathrm{u}}{\surd}$-base. Furthermore, we have shown in Theorem F2.3.13 that $\overset{\mathrm{u}}{\surd} \subseteq \surd$. Hence, $\mathfrak{M}$ is also a $\surd$-base. □

The reason we are interested in extension bases is the following result.

Lemma 2.4. *If forking equals dividing over U, then U is a $\overset{\mathrm{f}}{\surd}$-extension base.*

Proof. As forking equals dividing over U, it is sufficient to show that U is a $\overset{\mathrm{d}}{\surd}$-extension base. We have already shown in Lemma 2.3 that U is a $\overset{\mathrm{d}}{\surd}$-base. It therefore remains to show that $\overset{\mathrm{d}}{\surd}$ has left extension over U.

Suppose that $\bar{a} \mathrel{\overset{\mathrm{d}}{\surd}}_U \bar{b}$ and let $\bar{c} \subseteq \mathbb{M}$. To find some tuple $\bar{b}' \equiv_{U\bar{a}} \bar{b}$ with $\bar{a}\bar{c} \mathrel{\overset{\mathrm{d}}{\surd}}_U \bar{b}'$, we set $\mathfrak{p} := \mathrm{tp}(\bar{b}/U\bar{a})$ and

$$\Phi(\bar{x}) := \mathfrak{p}(\bar{x}) \cup \{\, \neg\varphi(\bar{x},\bar{a},\bar{c}) \mid \varphi(\bar{b},\bar{y},\bar{z}) \text{ divides over } U \,\}\,.$$

Clearly, every tuple $\bar{b}'$ realising $\Phi(\bar{x})$ has the desired properties. Hence, it remains to prove that Φ is consistent.

For a contradiction, suppose otherwise. Then

$$\mathfrak{p} \vDash \bigvee_{i<n} \varphi_i(\bar{x}, \bar{a}, \bar{c}),$$

where each formula $\varphi_i(\bar{b}, \bar{y}, \bar{z})$ divides over U. In particular, the disjunction

$$\psi(\bar{b}, \bar{y}, \bar{z}) := \bigvee_{i<n} \varphi_i(\bar{b}, \bar{y}, \bar{z})$$

forks over U. By assumption, this implies that ψ also divides over U. Thus, there exists an indiscernible sequence $\beta = (\bar{b}_i)_{i<\omega}$ over U such that $\bar{b}_0 = \bar{b}$ and $\{ \psi(\bar{b}_i, \bar{y}, \bar{z}) \mid i < \omega \}$ is k-inconsistent, for some $k < \omega$. By Lemma E5.3.11, we can find a sequence $\beta' \equiv_{U\bar{b}} \beta$ such that $\beta' = (\bar{b}'_i)_{i<\omega}$ is indiscernible over $U\bar{a}$. As $\mathfrak{p}$ is a type over $U\bar{a}$, it follows that

$$\operatorname{tp}(\bar{b}'_i/U\bar{a}) = \operatorname{tp}(\bar{b}'_0/U\bar{a}) = \operatorname{tp}(\bar{b}/U\bar{a}) = \mathfrak{p}, \quad \text{for all } i < \omega.$$

This implies that $\mathbb{M} \vDash \psi(\bar{b}'_i, \bar{a}, \bar{c})$, for all i. Thus, the tuple $\bar{c}\bar{a}$ satisfies the set $\{ \psi(\bar{b}'_i, \bar{y}, \bar{z}) \mid i < \omega \}$, which is k-inconsistent by choice of β'. A contradiction. □

Quasi-dividing and the Broom Lemma

Before attacking the questions of when forking and dividing coincide, we take a look at a weakening of dividing called *quasi-dividing.*

Definition 2.5. A formula $\varphi(\bar{x}; \bar{b})$ *quasi-divides* over a set U if there are tuples $\bar{b}_0, \dots, \bar{b}_{n-1}$, for some $n < \omega$, such that

$$\bar{b}_i \equiv_U \bar{b} \quad \text{and} \quad \{ \varphi(\bar{x}; \bar{b}_i) \mid i < n \} \text{ is inconsistent.}$$

Lemma 2.6. *Dividing implies quasi-dividing.*

Proof. Suppose that $\varphi(\bar{x}; \bar{b})$ divides over U. Then there is a sequence $(\bar{b}_i)_{i<\omega}$ such that $\bar{b}_i \equiv_U \bar{b}$ and $\{ \varphi(\bar{x}; \bar{b}_i) \mid i < \omega \}$ is k-inconsistent, for some $k < \omega$. Consequently, the tuples $\bar{b}_0, \dots, \bar{b}_{k-1}$ show that $\varphi(\bar{x}; \bar{b})$ quasi-divides over U. □

We start with a technical lemma that, given a forking relation with left extension, constructs something like a Morley sequence for the inverse relation.

Lemma 2.7. *Let $\surd$ be a forking relation with left extension over a set U, $\beta = (\bar b_n)_{n<\omega}$ an indiscernible sequence over $U \cup C$, and $\bar a$ a tuple such that*

$$C \surd_U \bar a\beta \quad \text{and} \quad \bar b_n \surd_U \bar a\bar b[<n], \quad \text{for all } 0 < n < \omega .$$

For every number $k < \omega$, there exists a sequence $\alpha = (\bar a_i)_{i<k}$ such that $\bar a_0 = \bar a$ and, for all $i < k$,

$$\bar a_i \bar b_i \equiv_{UC} \bar a\bar b_0 \quad \text{and} \quad C\bar a_{k-1}\bar b_{k-1} \dots \bar a_{i+1}\bar b_{i+1} \surd_U \bar a_i\bar b_i .$$

Proof. We prove the claim by induction on k. For $k = 0$, there is nothing to do. For the inductive step, suppose that we have already found a sequence $\alpha' = (\bar a'_i)_{i<k}$ of length k. We will construct one of length $k+1$. Let $\sigma \in \operatorname{Aut}\mathbb{M}_{UC}$ be an automorphism such that $\sigma(\bar b_n) = \bar b_{n+1}$, for all $n < \omega$. Note that $C \surd_U \bar a\bar b_0 \dots \bar b_k$ and $\bar b_i \surd_U \bar a\bar b_0 \dots \bar b_{i-1}$ implies, by Lemma F2.2.4 and induction on $i < k$, that

$$C\bar b_k \dots \bar b_{k-i+1} \surd_U \bar a\bar b_0 \dots \bar b_{k-i} .$$

For $i = k$, we obtain

$$C\bar b_k \dots \bar b_1 \surd_U \bar a\bar b_0 .$$

By (LEXT), we can therefore find tuples $\bar a'\bar b' \equiv_{UC\bar b_k \dots \bar b_1} \bar a\bar b_0$ such that

$$C\bar b_k \dots \bar b_1 \sigma(\bar a'_{k-1}) \dots \sigma(\bar a'_0) \surd_U \bar a'\bar b' .$$

Let $\pi \in \operatorname{Aut}\mathbb{M}_{UC\bar b_k \dots \bar b_1}$ be an automorphism with $\pi(\bar a'\bar b') = \bar a\bar b_0$ and set

$$\bar a_0 := \bar a \quad \text{and} \quad \bar a_{i+1} := \pi(\sigma(\bar a'_i)), \quad \text{for } i < k .$$

Then invariance implies that

$$C\bar{b}_k \dots \bar{b}_1 \bar{a}_k \dots \bar{a}_1 \mathrel{\sqrt{}_U} \bar{a}\bar{b}_0 .$$

We claim that the sequence $\alpha := (\bar{a}_i)_{i<k+1}$ obtained in this way has the desired properties.

Clearly, we have $\bar{a}_0 = \bar{a}$. Furthermore, since $\pi(\bar{b}_i) = \bar{b}_i$ for $0 < i \leq k$, we have

$$\bar{a}_{i+1}\bar{b}_{i+1} = \pi(\sigma(\bar{a}'_i))\bar{b}_{i+1} \equiv_{UC} \sigma(\bar{a}'_i)\sigma(\bar{b}_i) \equiv_{UC} \bar{a}'_i\bar{b}_i \equiv_{UC} \bar{a}\bar{b}_0 .$$

For the last condition, note that, for $i < k$,

$$\begin{aligned} &C\bar{a}'_{k-1}\bar{b}_{k-1} \dots \bar{a}'_{i+1}\bar{b}_{i+1} \mathrel{\sqrt{}_U} \bar{a}'_i\bar{b}_i \\ \Rightarrow\ &C\pi(\sigma(\bar{a}'_{k-1}\bar{b}_{k-1} \dots \bar{a}'_{i+1}\bar{b}_{i+1})) \mathrel{\sqrt{}_U} \pi(\sigma(\bar{a}'_i\bar{b}_i)) \\ \Rightarrow\ &C\bar{a}_k\bar{b}_k \dots \bar{a}_{i+2}\bar{b}_{i+2} \mathrel{\sqrt{}_U} \bar{a}_{i+1}\bar{b}_{i+1} . \end{aligned}$$

Furthermore, we have already seen above that

$$C\bar{a}_k\bar{b}_k \dots \bar{a}_1\bar{b}_1 \mathrel{\sqrt{}_U} \bar{a}_0\bar{b}_0 . \qquad \square$$

The following result is our main technical lemma. Note that, in the case where ψ = false, it states that a formula that forks in a particular way also quasi-divides.

Lemma 2.8 (Broom Lemma). *Let $\sqrt{} \subseteq \overset{\text{li}}{\sqrt{}}$ be a forking relation with left extension over some set U. Suppose that*

$$\vartheta(\bar{x};\bar{a}) \vDash \psi(\bar{x};\bar{c}) \vee \bigvee_{i<n} \varphi_i(\bar{x};\bar{b}^i)$$

and there are indiscernible sequences $\beta_i = (\bar{b}^i_j)_{j<\omega}$ over U such that

- $\bar{b}^i_0 = \bar{b}^i$ *and* $\{\, \varphi_i(\bar{x};\bar{b}^i_j) \mid j < \omega \,\}$ *is k-inconsistent, for every $i < n$,*
- $\bar{b}^i_j \mathrel{\sqrt{}_U} \beta[<i]\bar{b}^i[<j]$, *for all $i < n$ and $0 < j < \omega$,*

- $\bar{c} \sqrt{}_U \beta[<n]$.

Then there exist a number $m < \omega$ and tuples $\bar{a}_0, \dots, \bar{a}_{m-1} \subseteq \mathbb{M}$ such that

$$\bigwedge_{i<m} \vartheta(\bar{x}; \bar{a}_i) \vDash \psi(\bar{x}; \bar{c}) \quad \textit{and} \quad \bar{a}_i \equiv_U \bar{a}\,, \quad \textit{for all } i < m\,.$$

Proof. We prove the claim by induction on n. For $n = 0$, there is nothing to do. For the inductive step, suppose that we have already shown the claim for n. We aim to prove it for $n+1$. According to Proposition F4.2.18, $\bar{c} \sqrt{}_U \beta_0 \dots \beta_n$ implies that each sequence β_i is indiscernible over $U \cup \bar{c}$. Consequently, we can use Lemma 2.7 with $\bar{a} := \beta_0 \dots \beta_{n-1}$ and $\beta := \beta_n$ to construct a sequence $\alpha = (\alpha_i)_{i<k}$ such that

- $\alpha_0 = \beta_0 \dots \beta_{n-1}$,
- $\alpha_i \bar{b}^n_i \equiv_{U\bar{c}} \alpha_0 \bar{b}^n_0$, for all $i < k$,
- $\bar{c}\alpha_{k-1}\bar{b}^n_{k-1} \dots \alpha_{i+1}\bar{b}^n_{i+1} \sqrt{}_U \alpha_i \bar{b}^n_i$, for all $i < k$.

For each $j < k$, we choose an automorphism $\pi_j \in \operatorname{Aut} \mathbb{M}_{U\bar{c}}$ such that $\pi_j(\alpha_0 \bar{b}^n_0) = \alpha_j \bar{b}^n_j$. Then

$$\vartheta(\bar{x}; \pi_j(\bar{a})) \vDash \psi(\bar{x}; \bar{c}) \lor \bigvee_{i<n+1} \varphi_i(\bar{x}; \pi_j(\bar{b}^i))\,.$$

Consequently,

$$\bigwedge_{j<k} \vartheta(\bar{x}; \pi_j(\bar{a})) \vDash \bigwedge_{j<k} \Big[\psi(\bar{x}; \bar{c}) \lor \bigvee_{i<n} \varphi_i(\bar{x}; \pi_j(\bar{b}^i)) \lor \varphi_n(\bar{x}; \pi_j(\bar{b}^n)) \Big]\,.$$

This implies that

$$\bigwedge_{j<k} \vartheta(\bar{x}; \pi_j(\bar{a})) \land \neg \Big[\psi(\bar{x}; \bar{c}) \lor \bigvee_{i<n} \bigvee_{j<k} \varphi_i(\bar{x}; \pi_j(\bar{b}^i)) \Big] \vDash \bigwedge_{j<k} \varphi_n(\bar{x}; \pi_j(\bar{b}^n))\,.$$

Since $\{ \varphi_n(\bar{x}; \bar{b}^n_j) \mid j < \omega \}$ is k-inconsistent and $\pi_j(\bar{b}^n) = \bar{b}^n_j$, it follows that the formula

$$\bigwedge_{j<k} \vartheta(\bar{x}; \pi_j(\bar{a})) \wedge \neg\Big[\psi(\bar{x}; \bar{c}) \vee \bigvee_{i<n} \bigvee_{j<k} \varphi_i(\bar{x}; \pi_j(\bar{b}^i))\Big]$$

is inconsistent. Hence,

$$\bigwedge_{j<k} \vartheta(\bar{x}; \pi_j(\bar{a})) \vDash \psi(\bar{x}; \bar{c}) \vee \bigvee_{i<n} \bigvee_{j<k} \varphi_i(\bar{x}; \pi_j(\bar{b}^i)) .$$

For $s \leq k$, set

$$\psi_s(\bar{x}; \bar{c}^s) := \psi(\bar{x}; \bar{c}) \vee \bigvee_{i<n} \bigvee_{s \leq j<k} \varphi_i(\bar{x}; \pi_j(\bar{b}^i)) .$$

By induction on s, we will find tuples $\bar{a}_0, \dots, \bar{a}_{m-1}$ such that

$$\bigwedge_{i<m} \vartheta(\bar{x}; \bar{a}_i) \vDash \psi_s(\bar{x}; \bar{c}^s) \quad \text{and} \quad \bar{a}_i \equiv_U \bar{a} , \quad \text{for all } i < m .$$

Then the statement of the lemma will follow for $s = k$. For $s = 0$, we can take the tuples $\pi_i(\bar{a})$ from above. For the inductive step, suppose that

$$\bigwedge_{i<m} \vartheta(\bar{x}; \bar{a}_i) \vDash \psi_s(\bar{x}; \bar{c}^s) \quad \text{where} \quad \bar{a}_i \equiv_U \bar{a} .$$

Note that

$$\psi_s(\bar{x}; \bar{c}^s) \equiv \psi_{s+1}(\bar{x}; \bar{c}^{s+1}) \vee \bigvee_{i<n} \varphi_i(\bar{x}; \pi_s(\bar{b}^i))$$

and the sequences $\pi_s(\beta_i)$ satisfy

- $\pi_s(\bar{b}^i_0) = \pi_s(\bar{b}^i)$ and $\{ \varphi_i(\bar{x}; \pi_s(\bar{b}^i_j)) \mid j < \omega \}$ is k-inconsistent, for every $i < n$,
- $\pi_s(\bar{b}^i_j) \mathop{\sqrt{}}_U \pi_s(\beta[<i])\pi_s(\bar{b}^i[<j])$, for all $i < n$ and $j < \omega$.

Furthermore, $\bar{b}^0 \dots \bar{b}^{n-1} \subseteq \beta_0 \dots \beta_{n-1} = \alpha_0$ implies

$$\pi_j(\bar{b}^0) \dots \pi_j(\bar{b}^{n-1}) \subseteq \pi_j(\alpha_0) = \alpha_j .$$

Consequently, we have $\bar{c}^{s+1} \subseteq \bar{c}\alpha_{k-1} \dots \alpha_{s+1}$ and

$$\bar{c}\alpha_{k-1} \dots \alpha_{s+1} \mathop{\surd}_U \alpha_s \quad \text{implies} \quad \bar{c}^{s+1} \mathop{\surd}_U \pi_s(\beta[<n]) .$$

Therefore, we can use the inductive hypothesis on n to obtain a number $m' < \omega$ and tuples $\bar{a}_{ij}$, for $i < m$ and $j < m'$, such that $\bar{a}_{ij} \equiv_U \bar{a}_i \equiv_U \bar{a}$ and

$$\bigwedge_{j<m'} \bigwedge_{i<m} \vartheta(\bar{x}; \bar{a}_{ij}) \vDash \psi_{s+1}(\bar{x}; \bar{c}^{s+1}) .$$

□

Remark. Note that we do *not* require that $\bar{b}^i_0 \mathop{\surd}_U \beta[<i]$. This will be essential in the applications below.

Recall that the Lemma of Kim states that, in a simple theory, every $\overset{\mathrm{f}}{\surd}$-Morley sequence is a witness for dividing. The next result contains a similar statement for certain $\surd$-Morley sequences.

Lemma 2.9. *Let $\surd \subseteq \overset{\mathrm{li}}{\surd}$ be a forking relation, U a $\surd$-extension base, and $\varphi(\bar{x}; \bar{y})$ a formula without the array property. For every tuple $\bar{b}$ such that $\varphi(\bar{x}; \bar{b})$ divides over U, there exists a model $\mathfrak{M}$ containing U and a global type $\mathfrak{p}$ extending* $\mathrm{tp}(\bar{b}/M)$ *such that $\mathfrak{p}$ is $\surd$-free over U and every sequence generated by $\mathfrak{p}$ over M witnesses that $\varphi(\bar{x}; \bar{b})$ divides over U.*

Proof. Since $\varphi(\bar{x}; \bar{b})$ divides over U, there exists a number $k < \omega$ and an indiscernible sequence $\beta = (\bar{b}_i)_{i<\omega}$ over U such that $\bar{b}_0 = \bar{b}$ and the set $\{\, \varphi(\bar{x}; \bar{b}_i) \mid i < \omega \,\}$ is k-inconsistent. Let $\mathfrak{N}$ be a $(|T| \oplus |U|)^+$-saturated and strongly $(|T| \oplus |U|)^+$-homogeneous model containing U. We can use Lemma E5.3.9 to extend β to an indiscernible sequence $\beta' = (\bar{b}_i)_{i<\lambda}$ over U of length $\lambda := (2^{|T| \oplus |N|})^+$. As $\beta' \mathop{\surd}_U U$, we find a sequence $\beta'' = (\bar{b}''_i)_{i<\lambda}$ such that $\beta'' \equiv_U \beta'$ and $\beta'' \mathop{\surd}_U N$.

As there are at most $2^{|T| \oplus |N|} < \lambda$ types over N, there exists an infinite subset $I \subseteq \lambda$ such that every tuple $\bar{b}''_i$ with $i \in I$ has the same type over N.

Let $\mathfrak{q}_0$ be this type and let $\mathfrak{M} \preceq \mathfrak{N}$ be some model containing U of size $|M| \leq |T| \oplus |U|$. Choose a strictly increasing function $g : \omega \to I$ and set $\alpha := (\bar{b}''_{g(i)})_{i<\omega}$.

Let $\mathfrak{q}$ be the type of α over N. Since $\beta'' \sqrt{}_U N$ and $\sqrt{} \subseteq \overset{\text{li}}{\sqrt{}}$, it follows that $\mathfrak{q}_0$ and $\mathfrak{q}$ are $\overset{\text{li}}{\sqrt{}}$-free over U. By Proposition F4.2.20 (5), this implies that they are $\overset{\text{s}}{\sqrt{}}$-free over M. By saturation of $\mathfrak{N}$, there exists a sequence $(\alpha_i)_{i<\omega}$ in N that is generated by $\mathfrak{q}$ over M. By Lemma F2.4.14, $(\alpha_i)_{i<\omega}$ is indiscernible over M. Suppose that $\alpha_i = (\bar{a}^i_n)_{n<\omega}$.

Let $i, j, k < \omega$. As $\mathfrak{q}$ is $\overset{\text{s}}{\sqrt{}}$-free over M it follows by transitivity that

$$\alpha[>k] \overset{\text{s}}{\sqrt{}}_M \alpha[\leq k] .$$

Since $\bar{a}^k_i$ and $\bar{a}^k_j$ both realise $\mathfrak{q}_0 \upharpoonright M\alpha[<k]$, we furthermore have

$$\bar{a}^k_i \equiv_{M\alpha[<k]} \bar{a}^k_j .$$

Consequently, $\alpha[>k] \overset{\text{s}}{\sqrt{}}_{M\alpha[<k]} \alpha_k$ implies that

$$\bar{a}^k_i \equiv_{M\alpha[<k]\alpha[>k]} \bar{a}^k_j .$$

As in Lemma F4.4.4, it follows that

$$(\bar{a}^k_{\eta(k)})_{k<\omega} \equiv_M (\bar{a}^k_0)_{k<\omega} , \quad \text{for all } \eta : \omega \to \omega .$$

By Proposition F2.4.3, $\mathfrak{q}_0$ has some global extension $\mathfrak{q}_1$ that is $\sqrt{}$-free over U. Fix a tuple $\bar{b}'$ realising $\mathfrak{q}_1 \upharpoonright M$. Then $\bar{b}' \equiv_U \bar{b}$ and there exists an automorphism $\pi \in \operatorname{Aut} \mathbb{M}_U$ with $\pi(\bar{b}') = \bar{b}$. Applying π to $\mathfrak{q}_1$ we obtain a global type $\mathfrak{p}$ extending $\operatorname{tp}(\bar{b}/\pi[M])$ that is $\sqrt{}$-free over U. We claim that this type $\mathfrak{p}$ and the model $M' := \pi[M]$ have the desired properties.

As $\mathfrak{q}_1$ is $\sqrt{}$-free over U, so is $\mathfrak{p}$. By base monotony it follows that $\mathfrak{p}$ is $\sqrt{}$-free over M. Hence, consider a sequence $(\bar{c}_i)_{i<\omega}$ generated by $\mathfrak{p}$ over M. As each tuple $\bar{c}_i$ realises $\mathfrak{p} \upharpoonright U = \mathfrak{q}_1 \upharpoonright U$, we have $\bar{c}_i \equiv_U \bar{b}$. Set $\bar{d}_i := \pi^{-1}(\bar{c}_i)$. Then the sequence $(\bar{d}_i)_{i<\omega}$ is generated by $\mathfrak{q}_1$ over M. Since so is the sequence $(\bar{a}^i_0)_{i<\omega}$, it follows by Lemma F2.4.14 that

$$(\bar{d}_i)_{i<\omega} \equiv_M (\bar{a}^i_0)_{i<\omega} .$$

Note that $\alpha_i \equiv_M \alpha$ implies that $\{ \varphi(\bar{x}; \bar{a}^i_n) \mid n < \omega \}$ is k-inconsistent. If the set $\{ \varphi(\bar{x}; \bar{a}^i_0) \mid i < \omega \}$ were consistent, the family $(\bar{a}^i_j)_{i,j<\omega}$ would form a k-array. Since the formula φ does not have the array property, the set $\{ \varphi(\bar{x}; \bar{a}^i_0) \mid i < \omega \}$ is therefore inconsistent. By indiscernibility, it follows that it is l-inconsistent, for some l. Hence, so is the set $\{ \varphi(\bar{x}; \bar{d}_i) \mid i < \omega \}$ and, applying the automorphism π, also the set $\{ \varphi(\bar{x}; \bar{c}_i) \mid i < \omega \}$. □

Using these lemmas we can derive the first step of our proof that forking equals dividing over certain sets.

Lemma 2.10. *Let T be a theory without the array property and $\sqrt{} \subseteq \sqrt[\mathrm{li}]{}$ a forking relation. Then forking implies quasi-dividing over every $\sqrt{}$-extension base U.*

Proof. Consider a formula $\varphi(\bar{x}; \bar{a})$ that forks over U. By Lemma F2.4.4, there are formulae $\psi_i(\bar{x}; \bar{b}_i)$ that divide over U such that $\varphi(\bar{x}; \bar{a}) \vDash \bigvee_{i<n} \psi_i(\bar{x}; \bar{b}_i)$. By Lemma 2.9, there are models $\mathfrak{M}_i$ and global types $\mathfrak{p}_i$, for $i < n$, such that $\mathfrak{p}_i$ extends $\mathrm{tp}(\bar{b}_i/M)$, $\mathfrak{p}_i$ is $\sqrt{}$-free over U, and every sequence generated by $\mathfrak{p}_i$ over M witnesses that $\psi_i(\bar{x}; \bar{b}_i)$ divides over U. For $i < n$, we choose a sequence $\beta_i = (\bar{b}^i_j)_{j<\omega}$ generated by $\mathfrak{p}_i$ as follows. We start with $\bar{b}^i_0 := \bar{b}_i$, which realises $\mathfrak{p}_i \upharpoonright M$. For $j > 0$, we choose a tuple $\bar{b}^i_j$ realising $\mathfrak{p}_i \upharpoonright M\beta[<i]\bar{b}^i[<j]$. It follows that

- $\bar{b}^i_0 = \bar{b}_i$ and the set $\{ \varphi_i(\bar{x}; \bar{b}^i_j) \mid j < \omega \}$ is k_i-inconsistent, for every $i < n$,
- $\bar{b}^i_j \sqrt{}_U \beta[<i]\bar{b}^i[<j]\,,$ for all $i < n$ and $0 < j < \omega$,
- $\varnothing \sqrt{}_U \beta[<n]\,.$

By Lemma 2.8, we can therefore find tuples $\bar{a}_i \equiv_U \bar{a}$, for $i < m$, such that

$$\varphi(\bar{x}; \bar{a}) \vDash \mathrm{false} \vee \bigvee_{i<n} \psi_i(\bar{x}; \bar{b}_i) \quad \text{implies} \quad \bigwedge_{i<m} \varphi(\bar{x}; \bar{a}_i) \vDash \mathrm{false}\,.$$

Consequently, $\varphi(\bar{x}; \bar{a})$ quasi-divides over U. □

Strict Lascar invariance

Above we have found a criterion for the fact that forking implies quasi-dividing over a given set. It remains to find conditions showing that quasi-dividing implies dividing. To do so, we introduce the following combination of the relations $\mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{li}}$ and $\mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{f}}$.

Definition 2.11. For sets $A, B, U \subseteq \mathbb{M}$, we define

$$A \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{fli}}_U B \quad :\text{iff} \quad A \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{li}}_U B \quad \text{and} \quad B \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{f}}_U A\,,$$
$$A \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{sli}}_U B \quad :\text{iff} \quad A \;{}^*\big(\mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{fli}}\big)_U\, B\,.$$

Lemma 2.12. $\bar{a} \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{sli}}_U B$ *if, and only if,* $\mathrm{tp}(\bar{a}/UB)$ *has a global extension* $\mathfrak{p}$ *that is Lascar-invariant over U and such that*

$$BC \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{f}}_U \bar{a}'\,, \quad \text{for all } C \subseteq \mathbb{M} \text{ and all } \bar{a}' \text{ realising } \mathfrak{p} \restriction UBC\,.$$

Proof. ($\Leftarrow$) Let $\mathfrak{p}$ be an extension of $\mathrm{tp}(\bar{a}/UB)$ as above. To show that $\bar{a} \;{}^*\big(\mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{fli}}\big)_U\, B$, we fix some set $C \subseteq \mathbb{M}$. Let $\bar{a}'$ be a tuple realising $\mathfrak{p} \restriction UBC$. Then $\bar{a}' \equiv_{UB} \bar{a}$ and, by choice of $\mathfrak{p}$, we have $\bar{a} \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{li}}_U BC$ and $BC \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{f}}_U \bar{a}$. This implies that $\bar{a} \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{fli}}_U BC$.

($\Rightarrow$) Let $\bar{a} \;{}^*\big(\mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{fli}}\big)_U\, B$. By Proposition F2.4.3, $\mathrm{tp}(\bar{a}/UB)$ has a global extension $\mathfrak{p}$ that is $\mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{fli}}$-free over U. As $\mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{fli}} \subseteq \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{li}}$, it is also Lascar invariant over U. For the second condition, suppose that $C \subseteq \mathbb{M}$ and let $\bar{a}'$ be a realisation of $\mathfrak{p} \restriction UBC$. Then $\bar{a}' \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{fli}}_U BC$ implies $BC \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{f}}_U \bar{a}'$. □

Lemma 2.13. *The relation* $\mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{fli}}$ *satisfies* (INV), (MON), (NOR), *and* (FIN).

Proof. (INV) follows from invariance of $\mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{li}}$ and $\mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{f}}$.

(MON) Suppose that $A \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{fli}}_U B$ and let $A_0 \subseteq A$ and $B_0 \subseteq B$. Then $A \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{li}}_U B$ and $B \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{f}}_U A$ and it follows that $A_0 \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{li}}_U B_0$ and $B_0 \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{f}}_U A_0$. Hence, $A_0 \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{fli}}_U B_0$.

(NOR) Suppose that $A \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{fli}}_U B$. Then $A \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{li}}_U B$ and $B \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{f}}_U A$ and it follows that $AU \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{li}}_U BU$ and $BU \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{f}}_U AU$. Hence, $AU \mathop{\smile\hskip-0.9em ^| \ }^{\mathrm{fli}}_U BU$.

(FIN) Suppose that $A_0 \overset{\mathrm{fli}}{\surd}_U B$, for all finite $A_0 \subseteq A$. Then $A_0 \overset{\mathrm{li}}{\surd}_U B$ and $B \overset{\mathrm{f}}{\surd}_U A_0$, for all finite $A_0 \subseteq A$. This implies that $A \overset{\mathrm{li}}{\surd}_U B$ and $B \overset{\mathrm{f}}{\surd}_U A$. Hence, $A \overset{\mathrm{fli}}{\surd}_U B$. □

Corollary 2.14. *The relation* $\overset{\mathrm{sli}}{\surd}$ *satisfies* (INV), (MON), (NOR), (FIN), *and* (EXT).

Proof. A closer look at the proof of Proposition F2.4.5 reveals that, to establish the axioms (INV), (MON), (NOR), (FIN), and (EXT) for the relation $\overset{*}{\surd}$, we only need to assume that $\surd$ satisfies (INV), (MON), (NOR), and (FIN). □

The reason we are interested in the relation $\overset{\mathrm{sli}}{\surd}$ is the following variant of the Lemma of Kim for theories with the array property.

Lemma 2.15. *Let T be a theory without the array property, $\varphi(\bar{x};\bar{b})$ a formula that divides over U, and $(\bar{b}_n)_{n<\omega}$ a sequence such that*

$$\bar{b}_n \equiv_U \bar{b} \quad \textit{and} \quad \bar{b}_n \overset{\mathrm{sli}}{\surd}_U \bar{b}[<n], \quad \textit{for all } n < \omega .$$

Then $\{ \varphi(\bar{x};\bar{b}_n) \mid n < \omega \}$ *is inconsistent.*

Proof. Applying a suitable automorphism, we may assume that $\bar{b}_0 = \bar{b}$. Since the formula $\varphi(\bar{x};\bar{b})$ divides over U, there exists an indiscernible sequence $\alpha = (\bar{a}_i)_{i<\omega}$ such that $\bar{a}_0 = \bar{b}$ and $\{ \varphi(\bar{x};\bar{a}_i) \mid i < \omega \}$ is k-inconsistent, for some $k < \omega$. By induction on $n < \omega$, we construct a family $(\alpha_j)_{j<n}$ of sequences $\alpha_j = (\bar{a}^j_i)_{i<\omega}$ such that

- each α_j is indiscernible over $U\alpha[<j]\bar{b}_{j+1} \dots \bar{b}_{n-1}$,
- $\alpha_j \equiv_U \alpha$, and $\bar{a}^j_0 = \bar{b}_j$.

For $n = 1$, we can take the sequence $\alpha_0 := \alpha$. For the inductive step, suppose we have already constructed a family $(\alpha'_j)_{j<n}$ of size n. Since

$$\bar{b}_n \overset{\mathrm{sli}}{\surd}_U \bar{b}[<n],$$

we can use (EXT) to find a family $(\alpha''_i)_{i<n}$ such that

$$\alpha''[<n] \equiv_{U\bar{b}[<n]} \alpha'[<n] \quad \text{and} \quad \bar{b}_n \mathop{\smile\!\!\!\!\!|}^{\text{sli}}_U \alpha''[<n] .$$

Since $\bar{b}_n \equiv_U \bar{b}$, there is some indiscernible sequence $\alpha'_n \equiv_U \alpha$ starting with $\bar{b}_n$. Note that $\bar{b}_n \mathop{\smile\!\!\!\!\!|}^{\text{sli}}_U \alpha''[<n]$ implies that $\alpha''[<n] \mathop{\smile\!\!\!\!\!|}^{\text{d}}_U \bar{b}_n$. By Lemma F3.1.3, we can therefore find a sequence $\alpha''_n \equiv_{U\bar{b}_n} \alpha'_n$ such that α''_n is indiscernible over $\alpha''[<n]$. We claim that the family $(\alpha''_i)_{i<n+1}$ has the desired properties.

Let $i < n$. By construction the sequence α''_i is indiscernible over $U\alpha''[<i]\bar{b}_{i+1}\dots\bar{b}_{n-1}$. Furthermore, we have $\bar{b}_n \mathop{\smile\!\!\!\!\!|}^{\text{li}}_U \alpha''[<n]$, which implies that

$$\bar{b}_n \mathop{\smile\!\!\!\!\!|}^{\text{li}}_{U\alpha''[<i]\bar{b}_{i+1}\dots\bar{b}_{n-1}} \alpha''_i .$$

By Proposition F4.2.18, it therefore follows that α''_i is also indiscernible over $U\alpha''[<i]\bar{b}_{i+1}\dots\bar{b}_{n-1}\bar{b}_n$. Finally, the sequence α''_n is indiscernible over $U\alpha''[<n]$ by construction.

Having constructed sequences $(\alpha_j)_{j<n}$ of length n, for every $n < \omega$, it follows by compactness that there also exists an infinite family $(\alpha_j)_{j<\omega}$ with the same properties.

To conclude the proof suppose, towards a contradiction, that the set $\{\, \varphi(\bar{x};\bar{b}_n) \mid n < \omega \,\}$ is consistent. For $\eta : \omega \to \omega$ and $n < \omega$, a straightforward induction on i shows that

$$\bar{a}^0_{\eta(0)} \dots \bar{a}^{n-1}_{\eta(n-1)} \equiv_U \bar{a}^0_{\eta(0)} \dots \bar{a}^{n-i-1}_{\eta(n-i-1)} \bar{a}^{n-i}_0 \dots \bar{a}^{n-1}_0 .$$

This implies that

$$(\bar{a}^i_{\eta(i)})_{i<\omega} \equiv_U (\bar{a}^i_0)_{i<\omega} = (\bar{b}_i)_{i<\omega} .$$

Consequently, $\{\, \varphi(\bar{x};\bar{a}^i_{\eta(i)}) \mid i < \omega \,\}$ is consistent, for every $\eta : \omega \to \omega$. Furthermore, $\alpha_j \equiv_U \alpha$ implies that $\{\, \varphi(\bar{x};\bar{a}^i_n) \mid n < \omega \,\}$ is k-inconsistent, for some k. Consequently, the family $(\bar{a}^j_i)_{i,j<\omega}$ forms a k-array for φ. A contradiction. □

We obtain our first result for forking equalling dividing over $\overset{\text{sli}}{\surd}$-bases.

Proposition 2.16. *Let T be a theory without the array property and U a $\overset{\text{sli}}{\surd}$-base. Then forking equals dividing over U.*

Proof. Suppose that $\varphi(\bar{x};\bar{a})$ forks over U. Then there exist formulae $\psi_i(\bar{x};\bar{b}_i)$ that divide over U such that $\varphi(\bar{x};\bar{a}) \vDash \bigvee_{i<n} \psi_i(\bar{x};\bar{b}_i)$. Set $\bar{c} := \bar{a}\bar{b}_0 \dots \bar{b}_{n-1}$ and let $\mathfrak{p} := \operatorname{tp}(\bar{c}/U)$. Since $\bar{c} \overset{\text{sli}}{\surd}_U U$ there exists a global type $\mathfrak{q}$ extending $\mathfrak{p}$ that is $\overset{\text{fli}}{\surd}$-free over U. Let $\mathfrak{M}$ be a model containing U and let $\gamma = (\bar{c}_i)_{i<\omega}$ be a sequence generated by $\mathfrak{q}$ over M. Note that, by Proposition F4.2.20 (5), $\mathfrak{q}$ is $\overset{\text{s}}{\surd}$-free over M. Hence, it follows by Lemma F2.4.14, that γ is a $\overset{\text{sli}}{\surd}$-Morley sequence. Suppose that $\bar{c}_i = \bar{a}^i\bar{b}^i_0 \dots \bar{b}^i_{n-1}$. We claim that the set $\{\, \varphi(\bar{x};\bar{a}^i) \mid i < \omega \,\}$ is inconsistent. Since γ is indiscernible and $\bar{a}^i \equiv_U \bar{a}$, this implies that $\varphi(\bar{x};\bar{a})$ divides over U.

For a contradiction, suppose that there exists a tuple $\bar{d}$ realising the above set. Then there exists a function $g : \omega \to [n]$ such that

$$\mathbb{M} \vDash \psi_{g(i)}(\bar{d};\bar{b}^i_{g(i)})\,, \quad \text{for all } i < \omega\,.$$

Choose an infinite subset $I \subseteq \omega$ and an index $k < n$ such that $g(i) = k$, for all $i \in I$. It follows that $\{\, \psi_k(\bar{x};\bar{b}^i_k) \mid i < \omega \,\}$ is consistent. This contradicts Lemma 2.15 ☐

It remains to prove that $\surd$-extension bases are also $\overset{\text{sli}}{\surd}$-bases. We start with a technical lemma.

Lemma 2.17. *Let $\surd$ be a forking relation and U a $\surd$-base such that forking implies quasi-dividing over U.*

(a) *Every type $\mathfrak{p}$ over U has a global extension $\mathfrak{q}$ that is $\surd$-free over U and such that*

$$C \overset{\text{f}}{\surd}_U \bar{a}\,, \quad \textit{for all } C \subseteq \mathbb{M} \textit{ and all } \bar{a} \textit{ realising } \mathfrak{q} \restriction UC\,.$$

(b) *Every type* $\mathfrak{p}$ *over* U *has a global extension* $\mathfrak{q}$ *that is* $\sqrt[\mathrm{f}]{}$*-free over* U *and such that*

$$C \sqrt{}_U\, \bar{a}\,, \quad \text{for all } C \subseteq \mathbb{M} \text{ and all } \bar{a} \text{ realising } \mathfrak{q} \upharpoonright UC\,.$$

Proof. (a) Fix a tuple $\bar{a}$ realising $\mathfrak{p}$ and set

$$\begin{aligned} \Phi(\bar{x}) := \mathfrak{p}(\bar{x}) \cup \{\, \neg\varphi(\bar{x};\bar{b}) \mid \bar{b} \subseteq \mathbb{M},\ \varphi(\bar{a};\bar{y})\ \sqrt[\mathrm{f}]{}\text{-forks over } U \,\} \\ \cup \{\, \neg\psi(\bar{x};\bar{b}) \mid \bar{b} \subseteq \mathbb{M},\ \psi(\bar{x};\bar{b})\ \sqrt{}\text{-forks over } U \,\}\,. \end{aligned}$$

By (DEF), every global type containing Φ has the desired properties. Hence, it remains to show that Φ is satisfiable.

For a contradiction, suppose otherwise. Then there exist formulae $\varphi_i(\bar{x};\bar{y}_i)$, $i < m$, and $\psi_i(\bar{x};\bar{z}_i)$, $i < n$, and corresponding parameters $\bar{b}_0, \dots, \bar{b}_{m-1}, \bar{b}'_0, \dots, \bar{b}'_{n-1}$ such that

$$\mathfrak{p} \vDash \bigvee_{i<m} \varphi_i(\bar{x};\bar{b}_i) \vee \bigvee_{i<n} \psi_i(\bar{x};\bar{b}'_i)\,,$$

each $\varphi_i(\bar{a};\bar{y})$ $\sqrt[\mathrm{f}]{}$-forks over U, and each $\psi_i(\bar{x};\bar{b}'_i)$ $\sqrt{}$-forks over U. As the disjunction $\bigvee_{i<m} \varphi_i(\bar{a};\bar{y}_i)$ also $\sqrt[\mathrm{f}]{}$-forks over U, we may assume that $m = 1$.

Since forking implies quasi-dividing over U, there are parameters $\bar{a}_0, \dots, \bar{a}_{k-1}$ such that $\bar{a}_i \equiv_U \bar{a}$ and the set $\{\, \varphi_0(\bar{a}_i;\bar{y}) \mid i < k \,\}$ is inconsistent. Set $\bar{c} := \bar{a}_0 \dots \bar{a}_{k-1}$ and $\mathfrak{r}(\bar{x}_0, \dots, \bar{x}_{k-1}) := \mathrm{tp}(\bar{c}/U)$. Then

$$\mathfrak{r} \upharpoonright \bar{x}_j \vDash \varphi_0(\bar{x}_j;\bar{b}_0) \vee \bigvee_{i<n} \psi_i(\bar{x}_j;\bar{b}'_i)\,.$$

Hence,

$$\mathfrak{r} \vDash \bigwedge_{j<k} \Big[\varphi_0(\bar{x}_j;\bar{b}_0) \vee \bigvee_{i<n} \psi_i(\bar{x}_j;\bar{b}'_i) \Big]\,.$$

Consequently,

$$\mathfrak{r} \vDash \neg \bigwedge_{j<k} \varphi_0(\bar{x}_j;\bar{b}_0) \quad \text{implies that} \quad \mathfrak{r} \vDash \bigvee_{j<k} \bigvee_{i<n} \psi_i(\bar{x}_j;\bar{b}'_i)\,.$$

Since U is a $\sqrt{}$-base, we have $\bar c \sqrt{}_U U$. Hence, there is some tuple $\bar c' \equiv_U \bar c$ such that $\bar c' \sqrt{}_U \bar b'_0 \dots \bar b'_{n-1}$. As $\bar c' = \bar c'_0 \dots \bar c'_{k-1}$ realises $\mathfrak{r}$, there are indices $j < k$ and $i < n$ such that $\mathbb{M} \vDash \psi_i(\bar c'_j; \bar b'_i)$. But this implies that $\bar c'_j \not\sqrt{}_U \bar b'_i$. A contradiction.

(b) The proof is similar to the one above. Fix a tuple $\bar a$ realising $\mathfrak{p}$ and set

$$\begin{aligned}\Phi(\bar x) := \mathfrak{p}(\bar x) \cup \{\, \neg\varphi(\bar x; \bar b) \mid \bar b \subseteq \mathbb{M},\ \varphi(\bar x; \bar b)\ \sqrt[\mathrm{f}]{}\text{-forks over } U \,\} \\ \cup \{\, \neg\psi(\bar x; \bar b) \mid \bar b \subseteq \mathbb{M},\ \psi(\bar a; \bar y) \sqrt{}\text{-forks over } U \,\}.\end{aligned}$$

Suppose that Φ is inconsistent. Then we can find formulae $\varphi_i(\bar x; \bar y_i)$, $i < m$, and $\psi_i(\bar x; \bar z_i)$, $i < n$, and parameters $\bar b_0, \dots, \bar b_{m-1}, \bar b'_0, \dots, \bar b'_{n-1}$ such that

$$\mathfrak{p} \vDash \bigvee_{i<m} \varphi_i(\bar x; \bar b_i) \vee \bigvee_{i<n} \psi_i(\bar x; \bar b'_i),$$

each $\varphi_i(\bar x; \bar b_i)$ $\sqrt[\mathrm{f}]{}$-forks over U, and each $\psi_i(\bar a; \bar z_i)$ $\sqrt{}$-forks over U. As above, we may assume that $m = 1$.

Since forking implies quasi-dividing over U, there are parameters $\bar c_0, \dots, \bar c_{k-1}$ such that $\bar c_j \equiv_U \bar b_0$ and the set $\{\, \varphi_0(\bar x; \bar c_j) \mid j < k \,\}$ is inconsistent. Choose tuples $\bar d_{ji}$ such that

$$\bar c_j \bar d_{j0} \dots \bar d_{j(n-1)} \equiv_U \bar b_0 \bar b'_0 \dots \bar b'_{n-1}, \quad \text{for } j < k.$$

Since the type $\mathfrak{p}$ is over U, it follows by invariance that

$$\mathfrak{p} \vDash \varphi_0(\bar x; \bar c_j) \vee \bigvee_{i<n} \psi_i(\bar x; \bar d_{ji}), \quad \text{for all } j < k.$$

As above, this implies that

$$\mathfrak{p} \vDash \bigvee_{j<k} \bigvee_{i<n} \psi_i(\bar x; \bar d_{ji}).$$

Set $\bar d := (\bar d_{ji})_{j<k, i<n}$. As U is a $\sqrt{}$-base, we have $\bar d \sqrt{}_U U$. Consequently, there is some tuple $\bar d' \equiv_U \bar d$ such that

$$\bar d' \sqrt{}_U \bar a.$$

Since $\bar{a}$ realises $\mathfrak{p}$, there are indices $j < k$ and $i < n$ such that

$$\mathbb{M} \vDash \psi_i(\bar{a}; \bar{d}'_{ji}) .$$

But this implies that $\bar{d}'_{ji} \not\surd_U \bar{a}$. A contradiction. $\square$

Corollary 2.18. *Let T be a theory without the array property and U a $\overset{\mathrm{li}}{\surd}$-base such that forking implies quasi-dividing over U. Then U is a $\overset{\mathrm{sli}}{\surd}$-base.*

Proof. Fix a tuple $\bar{a} \subseteq \mathbb{M}$. We can use Lemma 2.17 to find a global extension $\mathfrak{q}$ of $\mathrm{tp}(\bar{a}/U)$ that is $\overset{\mathrm{li}}{\surd}$-free over U and such that $C \overset{\mathrm{f}}{\surd}_U \bar{a}'$, for all sets $C \subseteq \mathbb{M}$ and all tuples $\bar{a}'$ realising $\mathfrak{q} \upharpoonright UC$. By Lemma 2.12, this implies that $\bar{a} \overset{\mathrm{sli}}{\surd}_U U$. $\square$

Corollary 2.19. *Let T be a theory without the array property and $\surd \subseteq \overset{\mathrm{li}}{\surd}$ a forking relation. Every $\surd$-extension base is a $\overset{\mathrm{sli}}{\surd}$-base.*

Proof. Let U be a $\surd$-extension base. We have proved in Lemma 2.10 that forking implies quasi-dividing over U. Furthermore, since $\surd \subseteq \overset{\mathrm{li}}{\surd}$ and U is a $\surd$-base, it is also a $\overset{\mathrm{li}}{\surd}$-base. Consequently, the claim follows by Corollary 2.18. $\square$

Proposition 2.20. *Let T be a theory without the array property. Then forking equals dividing over every set that is a $\surd$-extension base, for some forking relation $\surd \subseteq \overset{\mathrm{li}}{\surd}$.*

Proof. By Corollary 2.19, every $\surd$-extension base is a $\overset{\mathrm{sli}}{\surd}$-base. Hence, the claim follows by Proposition 2.16. $\square$

Corollary 2.21. *Let T be a theory without the array property. Then forking equals dividing over every model M.*

Proof. We have seen in Lemma 2.3 (c) that every model is a $\overset{\mathrm{u}}{\surd}$-extension base. Consequently, the claim follows by Proposition 2.20. $\square$

Combining the above results we obtain the following characterisation of those sets over which forking equals dividing.

Theorem 2.22 (Chernikov, Kaplan). *Let T be a theory without the array property and $U \subseteq \mathbb{M}$ be a set. The following statements are equivalent.*

(1) *Forking equals dividing over U.*

(2) *U is a $\sqrt[\mathrm{f}]{}$-base.*

(3) *$\sqrt[\mathrm{f}]{}$ has left extension over U.*

Proof. The implications (1) ⇒ (2) and (1) ⇒ (3) follow by Lemma 2.4. Conversely, suppose that (2) or (3) holds. Let $\varphi(\bar{x};\bar{b})$ be a formula that forks over U. To show that $\varphi(\bar{x};\bar{b})$ also divides over U, we fix a model $\mathfrak{M}$ containing U.

If (2) holds, we have $M \sqrt[\mathrm{f}]{}_U\, U$ which, by (EXT), implies that there is some model $M' \equiv_U M$ with $M' \sqrt[\mathrm{f}]{}_U\, \bar{b}$.

If (3) holds, we have $U \sqrt[\mathrm{f}]{}_U\, \bar{b}$ which, by (LEXT), implies that there is some model $M' \equiv_U M$ with $M' \sqrt[\mathrm{f}]{}_U\, \bar{b}$.

Thus, in both cases we have found a model M' such that $M' \sqrt[\mathrm{f}]{}_U\, \bar{b}$. We claim that $\varphi(\bar{x};\bar{b})$ also forks over M'. Since forking equals dividing over models, it then follows that $\varphi(\bar{x};\bar{b})$ divides over M'. In particular, it divides over U.

To prove the claim suppose, for a contradiction, that $\varphi(\bar{x};\bar{b})$ does not fork over M'. Then we have $\bar{a} \sqrt[\mathrm{f}]{}_{M'}\, \bar{b}$, for every tuple $\bar{a}$ satisfying $\varphi(\bar{x};\bar{b})$. By (LTR), this implies that $\bar{a}M' \sqrt[\mathrm{f}]{}_U\, \bar{b}$, which contradicts the fact that $\varphi(\bar{x};\bar{b})$ forks over U. □

Corollary 2.23. *Let T be a theory without the array property.*

(a) *A set U is a $\sqrt[\mathrm{sli}]{}$-base if, and only if, it is a $\sqrt[\mathrm{li}]{}$-base.*

(b) *Forking equals dividing over every $\sqrt[\mathrm{li}]{}$-base.*

Proof. (b) Let U be a $\sqrt[\mathrm{li}]{}$-base. Since $\sqrt[\mathrm{li}]{} \subseteq \sqrt[\mathrm{f}]{}$, it is also a $\sqrt[\mathrm{f}]{}$-base. By Theorem 2.22, it follows that forking equals dividing over U.

(a) The implication ($\Rightarrow$) follows by the inclusion $\overset{\text{sli}}{\surd} \subseteq \overset{\text{li}}{\surd}$. For ($\Leftarrow$), let U be a $\overset{\text{li}}{\surd}$-base. By (b), forking equals dividing over U. Since dividing implies quasi-dividing, it follows that forking implies quasi-dividing over U. By Corollary 2.18, it follows that U is a $\overset{\text{sli}}{\surd}$-base. □

3. *The Independence Theorem*

The Independence Theorem contains a characterisation of simple theories in terms of a certain property of the forking relation. A weaker version of this property also holds for theories without the array property. In this section we will present the weak version, use it to derive the strong one, and show that the latter characterises simple theories.

The chain condition

Before turning to the Independence Theorem itself, we first consider a closely related property called the *chain condition*.

Definition 3.1. A preforking relation $\surd$ satisfies the *chain condition* over a set $U \subseteq \mathbb{M}$ if, for every indiscernible sequence $(\bar{b}_i)_{i \in I}$ over U and every set of formulae $\Phi(\bar{x}; \bar{y})$ such that, for some $i_0 \in I$, the set $\Phi(\bar{x}; \bar{b}_{i_0})$ does not $\surd$-fork over U, the union $\bigcup_{i \in I} \Phi(\bar{x}; \bar{b}_i)$ also does not $\surd$-fork over U.

The chain condition can be characterised is several equivalent ways. The following list is somewhat parallel to the characterisation of dividing in Lemma F3.1.3.

Proposition 3.2. *Let $\surd$ be a forking relation and $U \subseteq \mathbb{M}$ a set of parameters. The following statements are equivalent.*

(1) *$\surd$ satisfies the chain condition over U.*

(2) *If a formula $\varphi(\bar{x}; \bar{b})$ does not $\surd$-fork over U and $\bar{b} \approx^{\text{ls}}_U \bar{b}'$, then $\varphi(\bar{x}; \bar{b}) \wedge \varphi(\bar{x}; \bar{b}')$ also does not $\surd$-fork over U.*

(3) *For every cardinal λ, there exists a cardinal κ such that, for every partial type $\mathfrak{p}$ over U and every family $(\mathfrak{q}_i)_{i<\kappa}$ of partial types of size $|\mathfrak{q}_i| < \lambda$ such that no $\mathfrak{p} \cup \mathfrak{q}_i$ $\surd$-forks over U, there are indices $i < j$ such that $\mathfrak{p} \cup \mathfrak{q}_i \cup \mathfrak{q}_j$ does not $\surd$-fork over U.*

(4) *For every indiscernible sequence $\beta = (\bar{b}_i)_{i<\omega}$ over U and every tuple $\bar{a} \surd_U \bar{b}_0$, there exists a sequence $\beta' \equiv_{U\bar{b}_0} \beta$ such that β' is indiscernible over $U\bar{a}$ and $\bar{a} \surd_U \beta'$.*

Proof. (2) ⇒ (3) By Corollary F4.2.9, there exists a cardinal κ such that, for every sequence $(\bar{b}_i)_{i<\kappa}$ of tuples of size $|\bar{b}_i| < \lambda$, there are indices $i < j$ such that $\bar{b}_i \approx^{\mathrm{ls}}_U \bar{b}_j$. Increasing κ, if necessary, we may ensure that κ is larger than the number of sets of formulae of size less than λ. We claim that this cardinal κ has the desired properties.

Let $\mathfrak{p}$ and $(\mathfrak{q}_i)_{i<\kappa}$ be types as above. Then there exists a subset $I \subseteq \kappa$ of size $|I| = \kappa$, a set $\Phi(\bar{x};\bar{y})$ of formulae (without parameters), and tuples $\bar{b}_i \in \mathbb{M}^{<\lambda}$ such that

$$\mathfrak{q}_i(\bar{x}) = \Phi(\bar{x};\bar{b}_i)\,, \quad \text{for all } i \in I\,.$$

By choice of κ, we can find indices $i < j$ in I such that $\bar{b}_i \approx^{\mathrm{ls}}_U \bar{b}_j$. We claim that the type

$$\mathfrak{p} \cup \mathfrak{q}_i \cup \mathfrak{q}_j = \mathfrak{p}(\bar{x}) \cup \Phi(\bar{x};\bar{b}_i) \cup \Phi(\bar{x};\bar{b}_j)$$

does not $\surd$-fork over U.

For a contradiction, suppose otherwise. By compactness, we can then find finite sets $\Psi_0 \subseteq \mathfrak{p}$ and $\Phi_0 \subseteq \Phi$ such that

$$\Psi_0(\bar{x}) \cup \Phi_0(\bar{x};\bar{b}_i) \cup \Phi_0(\bar{x};\bar{b}_j) \quad \surd\text{-forks over } U\,.$$

Setting

$$\varphi(\bar{x};\bar{y}) := \bigwedge \Psi_0(\bar{x}) \wedge \bigwedge \Phi_0(\bar{x};\bar{y})\,,$$

it follows that the formula $\varphi(\bar{x};\bar{b}_i) \wedge \varphi(\bar{x};\bar{b}_j)$ $\surd$-forks over U. On the other hand, $\mathfrak{p} \cup \mathfrak{q}_i \vDash \varphi(\bar{x};\bar{b}_i)$ implies that $\varphi(\bar{x};\bar{b}_i)$ does not $\surd$-fork over U. As $\bar{b}_i \approx^{\mathrm{ls}}_U \bar{b}_j$, this contradicts (2).

(3) ⇒ (1) Let κ be the cardinal from (3) associated with $\lambda := |\Phi|^+$. Extending the sequence $(\bar{b}_i)_{i \in I}$ we may assume that $|I| \geq \kappa$. For $w \subseteq I$, set

$$\Phi_w := \bigcup_{i \in w} \Phi(\bar{x}; \bar{b}_i)\,.$$

By compactness, it is sufficient to show that there is no finite subset $w \subseteq I$ such that Φ_w $\surd$-forks over U. We proceed by induction on $|w|$. For $w = \{i\}$, the claim holds since $\bar{b}_i \equiv_U \bar{b}_{i_0}$ and $\Phi(\bar{x}; \bar{b}_{i_0})$ does not $\surd$-fork over U. Hence, suppose that $n := |w| > 1$. Let $F := [I]^{n-1}$. By inductive hypothesis, no set Φ_s with $s \in F$ $\surd$-forks over U. Hence, we can use (3) to find indices $s \neq t \in F$ such that $\Phi_s \cup \Phi_t$ does not $\surd$-fork over U. Choosing sets $u, v \in F$ such that $\operatorname{ord}(uv) = \operatorname{ord}(st)$ and $w \subseteq u \cup v$, it follows by indiscernibility that $\Phi_w \subseteq \Phi_u \cup \Phi_v$ does not $\surd$-fork over U.

(1) ⇒ (4) Set $\mathfrak{p}(\bar{x}, \bar{x}') := \operatorname{tp}(\bar{a}\bar{b}_0/U)$. We extend β to an indiscernible sequence $\beta = (\bar{b}_i)_{i<\gamma}$ over U of length $\gamma \geq \beth_{\lambda^+}$ where $\lambda := 2^{|T| \oplus |U| \oplus |\bar{b}_0|}$. By the chain condition, the union $\bigcup_{i<\gamma} \mathfrak{p}(\bar{x}, \bar{b}_i)$ does not $\surd$-fork over U. Hence, there exists a tuple $\bar{a}'$ realising $\bigcup_{i<\gamma} \mathfrak{p}(\bar{x}, \bar{b}_i)$ such that $\bar{a}' \surd_U \beta$. Then $\bar{a}' \equiv_{U\bar{b}_0} \bar{a}$ and we can find a sequence $\beta' = (\bar{b}'_i)_{i<\gamma}$ such that $\bar{a}'\beta \equiv_{U\bar{b}_0} \bar{a}\beta'$. By Theorem E5.3.7 and choice of γ, there exists an indiscernible sequence $\beta'' = (\bar{b}''_n)_{n<\omega}$ over $U\bar{a}\bar{b}_0$ such that, for every $\bar{\imath} \in [\omega]^{<\omega}$, there is some $\bar{\jmath} \in [\gamma]^{<\omega}$ with

$$\bar{b}''[\bar{\imath}] \equiv_{U\bar{a}\bar{b}_0} \bar{b}'[\bar{\jmath}]\,.$$

By finite character, $\bar{a} \surd_U \bar{b}_0\beta'$ implies that $\bar{a} \surd_U \bar{b}_0\beta''$. By choice of β'' we can find, for every $n < \omega$, some tuple $\bar{\jmath} \in [\gamma]^n$ such that

$$\bar{b}_0\bar{b}''_0 \dots \bar{b}''_{n-1} \equiv_{U\bar{a}\bar{b}_0} \bar{b}_0\bar{b}'[\bar{\jmath}] \equiv_{U\bar{b}_0} \bar{b}_0\bar{b}[\bar{\jmath}] \equiv_{U\bar{b}_0} \bar{b}_0\bar{b}_1 \dots \bar{b}_n\,.$$

This implies that $\bar{b}_0\beta'' \equiv_{U\bar{b}_0} \beta$. Hence, the sequence $\beta''' := \bar{b}_0\beta''$ has the desired properties.

(4) ⇒ (2) Suppose that (2) does not hold. Then we can find a formula $\varphi(\bar{x}; \bar{y})$ and an indiscernible sequence $\beta = (\bar{b}_i)_{i<\omega}$ over U such

that the formula $\varphi(\bar{x};\bar{b}_0)$ does not $\surd$-fork over U, but the conjunction $\varphi(\bar{x};\bar{b}_0) \wedge \varphi(\bar{x};\bar{b}_1)$ does. We choose a tuple $\bar{a} \in \varphi(\bar{x};\bar{b}_0)^{\mathbb{M}}$ with $\bar{a} \surd_U \bar{b}_0$. For every sequence $\beta' = (\bar{b}'_i)_{i<\omega} \equiv_{U\bar{b}_0} \beta$ that is indiscernible over $U\bar{a}$, we then have $\mathbb{M} \vDash \varphi(\bar{a};\bar{b}'_i)$, for all i. As the conjunction $\varphi(\bar{x};\bar{b}'_0) \wedge \varphi(\bar{x};\bar{b}'_1)$ $\surd$-forks over U, it follows that $\bar{a} \not\surd_U \beta'$, for each such sequence β'. Therefore, (4) fails as well. □

As several of the characterisations of the chain condition are similar to characterisations of the dividing relation, we obtain the following implication.

Lemma 3.3. *If a preforking relation $\surd$ satisfies the chain condition over a set U then*

$$\bar{a} \surd_U \bar{b} \quad \textit{implies} \quad \bar{a} \overset{\mathrm{d}}{\surd}_U \bar{b}\,.$$

Proof. Suppose that $\bar{a} \surd_U \bar{b}$. To show that $\bar{a} \overset{\mathrm{d}}{\surd}_U \bar{b}$, we use condition (3) from Lemma F3.1.3. Hence, let $(\bar{b}_n)_{n<\omega}$ be an indiscernible sequence over U with $\bar{b}_0 = \bar{b}$. Setting $\Phi(\bar{x},\bar{x}') := \mathrm{tp}(\bar{a}\bar{b}/U)$, it follows by the chain condition that there exists a tuple $\bar{a}'$ realising $\bigcup_{n<\omega} \Phi(\bar{x},\bar{b}_n)$ with $\bar{a}' \surd_U \bar{b}$. In particular, we have

$$\bar{a}' \equiv_{U\bar{b}} \bar{a} \quad \text{and} \quad \bar{b}_i \equiv_{U\bar{a}'} \bar{b}_k\,, \quad \text{for all } i,k < \omega\,. \qquad \square$$

As a first application of the chain condition, let us show that array-dividing equals dividing. Once we have shown that in theories without the array property $\overset{\mathrm{d}}{\surd}$ satisfies the chain condition, the following result will generalise Corollary 1.14.

Proposition 3.4. *Suppose that $\overset{\mathrm{d}}{\surd}$ satisfies the chain condition over a set U. A formula divides over U if, and only if, it array-divides over U.*

Proof. ($\Rightarrow$) was already proved in Lemma 1.13. For ($\Leftarrow$), suppose that $\varphi(\bar{x};\bar{b})$ does not divide over U. To show that it also does not array-divide over U, we consider a family $\beta = (\bar{b}_{ij})_{i,j<\omega}$ that is biindiscernible over U with $\bar{b}_{00} = \bar{b}$. We apply the chain condition to the sequence $\beta^0 = (\bar{b}_{i0})_{i<\omega}$ to show that the set $\{\, \varphi(\bar{x};\bar{b}_{i0}) \mid i < \omega \,\}$ does not divide over U. Applying the chain condition again, this time to the sequence $(\beta^i)_{i<\omega}$ of rows, it follows that the set $\{\, \varphi(\bar{x};\bar{b}_{ij}) \mid i, j < \omega \,\}$ does not divide over U. In particular, this set is consistent. □

Finally, we show that, in theories without the array property, $\sqrt[\mathrm{f}]{}$ satisfies the chain condition. We start by proving this implication over models before generalising it to arbitrary $\sqrt[\mathrm{f}]{}$-bases.

Lemma 3.5. *Let T be a theory without the array property and let $\mathfrak{M}$ be a model of T. Then $\sqrt[\mathrm{f}]{}$ satisfies the chain condition over M.*

Proof. We check condition (2) of Proposition 3.2. Let $\bar{b} \approx^{\mathrm{ls}}_M \bar{b}'$ be tuples and $\varphi(\bar{x};\bar{y})$ a formula such that the conjunction $\varphi(\bar{x};\bar{b}) \wedge \varphi(\bar{x};\bar{b}')$ forks over M. We have to show that $\varphi(\bar{x};\bar{b})$ also forks over M. Set $\kappa := \beth_{\lambda^+}$ where $\lambda := 2^{|T| \oplus |M|}$. Since $\bar{b} \approx^{\mathrm{ls}}_U \bar{b}'$, there exists an indiscernible sequence $\beta' = (\bar{b}'_i)_{i<\kappa}$ over M of length κ such that $\bar{b}'_0 = \bar{b}$ and $\bar{b}'_1 = \bar{b}'$. We have seen in Lemma 2.3 that M is a $\sqrt[\mathrm{u}]{}$-extension base. By Corollary 2.19 this implies that M is a $\sqrt[\mathrm{sli}]{}$-base. Furthermore, we have shown in Corollary 2.14 that $\sqrt[\mathrm{sli}]{}$ satisfies the extension axiom. Hence, we have $\beta' \sqrt[\mathrm{sli}]{}_M M$ and there exists a global type $\mathfrak{p} \supseteq \mathrm{tp}(\beta'/M)$ that is $\sqrt[\mathrm{sli}]{}$-free over M. Let $\beta = (\beta^i)_{i<\omega}$ be a sequence generated by $\mathfrak{p}$ over M where $\beta^i = (\bar{b}_{ij})_{j<\omega}$. By indiscernibility of β^0 and the fact that forking equals dividing over M, it follows for all pairs $j \neq j'$ of indices that the formula $\varphi(\bar{x};\bar{b}_{0j}) \wedge \varphi(\bar{x};\bar{b}_{0j'})$ divides over M. By choice of β and Lemma 2.15, this implies that the set

$$\{\, \varphi(\bar{x};\bar{b}_{ij}) \wedge \varphi(\bar{x};\bar{b}_{ij'}) \mid i < \omega \,\}$$

is inconsistent. We can use Theorem E5.3.7 to find an indiscernible sequence $\alpha = (\alpha^i)_{i<\omega}$ over M such that, for every $\bar{\imath} \in [\omega]^{<\omega}$, there is some $\bar{\jmath} \in [\kappa]^{<\omega}$ with $\alpha[\bar{\imath}] \equiv_M \beta[\bar{\jmath}]$. It follows that the family α is biindiscernible over M and the formula φ is inconsistent over α. Consequently,

$\varphi(\bar{x}; \bar{b}_{00})$ array-divides over M. According to Corollary 1.14 and Theorem 2.22, this implies that $\varphi(\bar{x}; \bar{b}_{00})$ also divides and forks over M. □

Theorem 3.6. *In a theory without the array property, $\sqrt[\mathrm{f}]{\ }$ satisfies the chain condition over every $\sqrt[\mathrm{f}]{\ }$-base.*

Proof. Let U be a $\sqrt[\mathrm{f}]{\ }$-base, $\varphi(\bar{x}; \bar{y})$ a formula, and $\beta = (\bar{b}_i)_{i<\omega}$ an indiscernible sequence over U such that $\varphi(\bar{x}; \bar{b}_0)$ does not fork over U. Fix a model $\mathfrak{M}$ containing U. Then $M \sqrt[\mathrm{f}]{\ }_U\ U$ and it follows by (EXT) that there exists a model $M' \equiv_U M$ such that $M' \sqrt[\mathrm{f}]{\ }_U\ \beta$. According to Theorem 2.22, we have $M' \sqrt[\mathrm{d}]{\ }_U\ \beta$. By Lemma F2.2.4, it therefore follows that a formula over β divides over U if, and only if, it divides over M'. In particular, $\varphi(\bar{x}; \bar{b}_0)$ does not divide over M'. By Lemma 3.5, the formula $\varphi(\bar{x}; \bar{b}_0) \wedge \varphi(\bar{x}; \bar{b}_1)$ does not divide over M'. Hence, it also does not divide over U. The claim follows since forking equals dividing over U. □

Corollary 3.7. *In a theory without the array property, $\sqrt[\mathrm{d}]{\ }$ satisfies the chain condition over every $\sqrt[\mathrm{f}]{\ }$-base.*

Proof. Let U be a $\sqrt[\mathrm{f}]{\ }$-base. According to Theorem 2.22, forking equals dividing over U. Consequently, $\sqrt[\mathrm{d}]{\ }$ has the chain condition over U if, and only if, $\sqrt[\mathrm{f}]{\ }$ does. Hence, the claim follows by the preceding theorem. □

The Independence Theorem

There are two versions of the Independence Theorem: a weak one that holds in all theories without the array property, and a strong one that characterises simple theories.

Definition 3.8. (a) A preforking relation $\sqrt{\ }$ satisfies the *Weak Independence Theorem* over a set $U \subseteq \mathbb{M}$ if it has the following property:

(WIND) If $\bar{a}, \bar{b}, \bar{b}', \bar{c} \subseteq \mathbb{M}$ are tuples satisfying

$$\bar{c} \sqrt{}_U \bar{a}\bar{b}\,, \quad \bar{a} \sqrt{}_U \bar{b}\bar{b}'\,, \quad \text{and} \quad \bar{b} \equiv^{\mathrm{ls}}_U \bar{b}'\,,$$

then there exists a tuple $\bar{c}'$ such that

$$\bar{c}' \sqrt{}_U \bar{a}\bar{b}'\,, \quad \bar{c}' \equiv_{U\bar{a}} \bar{c}\,, \quad \text{and} \quad \bar{b}'\bar{c}' \equiv_U \bar{b}\bar{c}\,.$$

(b) A preforking relation $\sqrt{}$ satisfies the *Independence Theorem* over a set $U \subseteq \mathbb{M}$ if it has the following property:

(IND) If $\bar{a}, \bar{b}, A, B \subseteq \mathbb{M}$ are tuples such that

$$\bar{a} \equiv_U \bar{b}\,, \quad \bar{a} \sqrt{}_U A\,, \quad \bar{b} \sqrt{}_U B\,, \quad \text{and} \quad A \sqrt{}_U B\,,$$

then there exists a tuple $\bar{c}$ such that

$$\bar{c} \equiv_{UA} \bar{a}\,, \quad \bar{c} \equiv_{UB} \bar{b}\,, \quad \text{and} \quad \bar{c} \sqrt{}_U AB\,.$$

We say that $\sqrt{}$ satisfies the Independence Theorem *for a class* $\mathcal{C} \subseteq \wp(\mathbb{M})$, if it satisfies the theorem over every $U \in \mathcal{C}$.

Remark. The statement of the second axiom becomes clearer when we rephrase it in terms of types. Then it reads:

> Let $\mathfrak{p}, \mathfrak{q}, \mathfrak{r}$ be types over, respectively, U, $U \cup A$, and $U \cup B$.
> If $\mathfrak{q}$ and $\mathfrak{r}$ are $\sqrt{}$-free extensions of $\mathfrak{p}$ and $A \sqrt{}_U B$, then $\mathfrak{q} \cup \mathfrak{r}$ is also a $\sqrt{}$-free extension of $\mathfrak{p}$.

We start by proving that the weak version holds in all theories without the array property.

Theorem 3.9. *For a forking relation $\sqrt{}$, the chain condition over a set U implies the Weak Independence Theorem over U.*

Proof. Suppose that $\sqrt{}$ satisfies the chain condition over U and let

$$\bar{c} \sqrt{}_U \bar{a}\bar{b}\,, \quad \bar{a} \sqrt{}_U \bar{b}\bar{b}'\,, \quad \text{and} \quad \bar{b} \equiv^{\mathrm{ls}}_U \bar{b}'\,.$$

We first consider the case where $\bar b \approx^{\mathrm{ls}}_U \bar b'$. By Lemma 3.3, we have $\bar a \mathrel{\sqrt[\mathrm{d}]{\ }_U} \bar b\bar b'$ and, hence, $\bar a \mathrel{\sqrt[\mathrm{d}]{\ }_{U\bar b}} \bar b'$. Therefore, we can use Lemma F4.2.12 to find a tuple $\bar a'$ such that $\bar a\bar b \approx^{\mathrm{ls}}_U \bar a'\bar b'$. Thus, there exists an indiscernible sequence $(\bar a_i\bar b_i)_{i<\omega}$ over U with $\bar a_0\bar b_0\bar a_1\bar b_1 = \bar a\bar b\bar a'\bar b'$. Since we have $\bar c \mathrel{\sqrt{\ }_U} \bar a_0\bar b_0$, it follows by Proposition 3.2 (4) that there is a tuple $\bar c' \equiv_{U\bar a_0\bar b_0} \bar c$ such that $\bar c' \mathrel{\sqrt{\ }_U} \bar a[\omega]\bar b[\omega]$ and $(\bar a_i\bar b_i)_{i<\omega}$ is indiscernible over $U\bar c'$. This implies that

$$\bar c' \mathrel{\sqrt{\ }_U} \bar a\bar b', \quad \bar c' \equiv_{U\bar a} \bar c, \quad \text{and} \quad \bar b'\bar c' \equiv_U \bar b\bar c' \equiv_U \bar b\bar c.$$

It remains to prove the general case. Fix a sequence $\bar b_0 \approx^{\mathrm{ls}}_U \cdots \approx^{\mathrm{ls}}_U \bar b_n$ such that $\bar b_0 = \bar b$ and $\bar b_n = \bar b'$. By (EXT), there is a tuple $\bar a' \equiv_{U\bar b\bar b'} \bar a$ such that $\bar a' \mathrel{\sqrt{\ }_U} \bar b_0 \dots \bar b_n$. Choosing tuples $\bar b'_0, \dots, \bar b'_n$ with

$$\bar a\bar b'_0 \dots \bar b'_n \equiv_{U\bar b\bar b'} \bar a'\bar b_0 \dots \bar b_n$$

it follows that $\bar b'_0 = \bar b$, $\bar b'_n = \bar b'$,

$$\bar b'_0 \approx^{\mathrm{ls}}_U \cdots \approx^{\mathrm{ls}}_U \bar b'_n \quad \text{and} \quad \bar a \mathrel{\sqrt{\ }_U} \bar b'_0 \dots \bar b'_n.$$

By the special case we have proved above, we can inductively find tuples $\bar c_0, \dots, \bar c_n$ such that $\bar c_0 = \bar c$,

$$\bar c_{i+1} \mathrel{\sqrt{\ }_U} \bar a\bar b'_{i+1}, \quad \bar c_{i+1} \equiv_{U\bar a} \bar c_i, \quad \text{and} \quad \bar b'_{i+1}\bar c_{i+1} \equiv_U \bar b'_i\bar c_i.$$

The tuple $\bar c' := \bar c_n$ has the desired properties. □

By Theorem 3.6, we can conclude that, in theories without the array property, $\sqrt[\mathrm{f}]{\ }$ satisfies the chain condition and, thus, the Weak Independence Theorem over $\sqrt[\mathrm{f}]{\ }$-bases.

Corollary 3.10 (Weak Independence Theorem; Ben Yaacov, Chernikov). *In a theory T without the array property, $\sqrt[\mathrm{f}]{\ }$ satisfies the Weak Independence Theorem over every $\sqrt[\mathrm{f}]{\ }$-base.*

Let us turn to the strong version of the Independence Theorem. Our goal is to show that it characterises $\downfree^{\mathrm{f}}$ in simple theories: a symmetric forking relation $\downfree$ satisfies the Independence Theorem if, and only if, $\downfree = \downfree^{\mathrm{f}}$ and the theory in question is simple. We start by proving that forking satisfies (IND) in simple theories.

Theorem 3.11 (Independence Theorem). *In a simple first-order theory $\downfree^{\mathrm{f}}$ satisfies the Independence Theorem for the class of all models.*

Proof. Let $\mathfrak{M}$ be a model and suppose that

$$\bar a \equiv_M \bar b\,, \quad \bar a \downfree^{\mathrm{f}}_M A\,, \quad \bar b \downfree^{\mathrm{f}}_M B\,, \quad \text{and} \quad A \downfree^{\mathrm{f}}_M B\,.$$

As in simple theories every set is a $\downfree^{\mathrm{f}}$-base, we have $\bar a \downfree^{\mathrm{f}}_{MA} MA$. Therefore, we can use Lemma F4.2.13 to find a tuple $\bar a' \equiv_{MA} \bar a$ such that $\bar a' \downfree^{\mathrm{f}}_{MA} B\bar a_0 \bar b$. Then it follows by transitivity that

$$\begin{aligned}
&\bar a' \downfree^{\mathrm{f}}_{MA} B\bar b && \text{and} && \bar a' \downfree^{\mathrm{f}}_M A && \text{implies} && \bar a' \downfree^{\mathrm{f}}_M AB\bar b\,,\\
&\bar a' \downfree^{\mathrm{f}}_M AB && \text{and} && B \downfree^{\mathrm{f}}_M A && \text{implies} && B\bar a' \downfree^{\mathrm{f}}_M A\,,\\
&\bar a' \downfree^{\mathrm{f}}_M B\bar b && \text{and} && \bar b \downfree^{\mathrm{f}}_M B && \text{implies} && \bar a'\bar b \downfree^{\mathrm{f}}_M B\,.
\end{aligned}$$

Furthermore, $\bar a' \equiv_M \bar a \equiv_M \bar b$, which implies that $\bar a' \equiv^{\mathrm{ls}}_M \bar b$. Hence, we can apply Corollary 3.10 to the statement $A \downfree^{\mathrm{f}}_M B\bar a'$ to find a set A' such that

$$A' \downfree^{\mathrm{f}}_M B\bar b\,, \quad A' \equiv_{MB} A\,, \quad \text{and} \quad \bar b A' \equiv_M \bar a' A\,.$$

Let $\bar c$ be a tuple such that $A'B\bar b \equiv_M AB\bar c$. Then

$$A \downfree^{\mathrm{f}}_M \bar c B \quad \text{and} \quad \bar c \downfree^{\mathrm{f}}_M B \quad \text{implies} \quad AB \downfree^{\mathrm{f}}_M \bar c\,.$$

Furthermore, we have

$$\bar c A \equiv_M \bar b A' \equiv_M \bar a' A \equiv_M \bar a A \quad \text{and} \quad \bar c B \equiv_M \bar b B\,. \qquad \square$$

It remains to prove that forking is the only symmetric forking relations satisfying the Independence Theorem.

Definition 3.12. A class $\mathcal{C} \subseteq \wp(\mathbb{M})$ of small sets is *invariant* if

$$C \equiv_\varnothing C' \quad \text{implies} \quad C \in \mathcal{C} \Leftrightarrow C' \in \mathcal{C}.$$

We call $\mathcal{C}$ *dense* if, for every set $A \subseteq \mathbb{M}$, there is some $C \in \mathcal{C}$ with $A \subseteq C$.

Example. Every class containing all models is dense. In particular, the class of all $\sqrt[\mathrm{u}]{}$-extension bases and the class of all $\sqrt[\mathrm{f}]{}$-bases are invariant and dense.

We start with a lemma constructing a Morley sequence. The proof follows the lines of the proofs of Lemmas F2.4.13 and F2.4.15.

Lemma 3.13. *Let $\downarrow$ be a right local forking relation, let $\mathcal{C} \subseteq \wp(\mathbb{M})$ be invariant and dense, and let $(\bar{a}_n)_{n<\omega}$ be an indiscernible sequence over U.*

There exists a set $C \in \mathcal{C}$ containing U and a type $\mathfrak{p} \in S^{\bar{s}}(C)$ extending $\mathrm{tp}(\bar{a}_0/U)$ such that $(\bar{a}_n)_{n<\omega}$ is a $\downarrow$-Morley sequence for $\mathfrak{p}$ over C.

Proof. Let $\kappa := \mathrm{loc}(\downarrow)^+ \oplus |\bar{a}_0|^+$. We can use Lemma E5.3.9 to extend $(\bar{a}_n)_{n<\omega}$ to an indiscernible sequence $(\bar{a}_\alpha)_{\alpha\leq\kappa}$ over U. We construct an increasing chain $(C_\alpha)_{\alpha<\kappa}$ of sets $C_\alpha \in \mathcal{C}$ such that, for every $\alpha < \kappa$,

$$U \cup \bar{a}[<\alpha] \subseteq C_\alpha \quad \text{and} \quad (\bar{a}_i)_{\alpha<i\leq\kappa} \text{ is indiscernible over } C_\alpha.$$

For the inductive step, suppose that C_i has already been defined for all $i < \alpha$. As $\mathcal{C}$ is dense, we can choose some set $C' \in \mathcal{C}$ containing $V_\alpha := U \cup \bar{a}[<\alpha] \cup \bigcup_{i<\alpha} C_i$. Since the sequence $(\bar{a}_i)_{\alpha<i<\kappa}$ is indiscernible over V_α, we can apply Lemma E5.3.11 to obtain a set $C_\alpha \equiv_{V_\alpha} C'$ such that $(\bar{a}_i)_{\alpha<i<\kappa}$ is indiscernible over $V_\alpha \cup C_\alpha$. By invariance, it follows that $C_\alpha \in \mathcal{C}$.

After having constructed the sequence $(C_\alpha)_{\alpha<\kappa}$, we can find a set $W \subseteq \bigcup_{\alpha<\kappa} C_\alpha$ of size $|W| < \mathrm{loc}(\downarrow) \oplus |\bar{a}_0|^+ \leq \kappa$ such that

$$\bar{a}_\kappa \downarrow_W \bigcup_{\alpha<\kappa} C_\alpha.$$

Since κ is regular, there exists an index $\gamma < \kappa$ such that $W \subseteq C_\gamma$. By (MON) and (BMON), it follows that

$$\bar{a}_\kappa \downarrow_{C_\gamma} \bigcup_{\gamma<i<\kappa} \bar{a}_i \,.$$

By (INV), we therefore have

$$\bar{a}_\alpha \downarrow_{C_\gamma} \bigcup_{\gamma<i<\alpha} \bar{a}_i \,, \quad \text{for all } \gamma < \alpha < \kappa \,.$$

Hence, $(\bar{a}_\alpha)_{\gamma<\alpha<\kappa}$ is a $\downarrow$-Morley sequence for $\mathrm{tp}(\bar{a}_\kappa/C_\gamma)$ over C_γ. Fix an automorphism $\pi \in \mathrm{Aut}\,\mathbb{M}_U$ such that $\pi[\bar{a}_{\gamma+n+1}] = \bar{a}_n$, for all $n < \omega$. By invariance, it follows that $(\bar{a}_n)_{n<\omega}$ is a $\downarrow$-Morley sequence for $\mathfrak{p} := \mathrm{tp}(\pi[\bar{a}_\kappa]/\pi[C_\gamma])$ over $C := \pi[C_\gamma]$. ☐

The main argument is contained in a technical lemma which states that the Independence Theorem implies the following weaker variant of the chain condition.

Definition 3.14. A preforking relation $\surd$ satisfies the *chain condition for Morley sequences* over a set $U \subseteq \mathbb{M}$ if, for every $\surd$-Morley sequence $(\bar{b}_i)_{i\in I}$ over U and every set of formulae $\Phi(\bar{x};\bar{y})$ such that, for some $i_0 \in I$, the set $\Phi(\bar{x};\bar{b}_{i_0})$ does not $\surd$-fork over U, the union $\bigcup_{i\in I} \Phi(\bar{x};\bar{b}_i)$ also does not $\surd$-fork over U.

Lemma 3.15. *Let $\surd$ be a forking relation satisfying the Independence Theorem over a set U. Then $\surd$ satisfies the chain condition for Morley sequence over U.*

Proof. Let $(\bar{b}_n)_{n<\omega}$ be a $\surd$-Morley sequence over U and let $\Phi(\bar{x};\bar{y})$ be a set such that $\Phi(\bar{x};\bar{b}_0)$ does not $\surd$-fork over U. We fix a tuple $\bar{a}$ with $\bar{a} \surd_U \bar{b}_0$ and we set $\mathfrak{p}(\bar{x},\bar{x}') := \mathrm{tp}(\bar{a}\bar{b}_0/U)$. We have to show that there exists a tuple $\bar{c}$ realising $\bigcup_{n<\omega} \mathfrak{p}(\bar{x},\bar{b}_n)$ such that $\bar{c} \surd_U \bar{b}[<\omega]$.

To do so, we construct a sequence $(\bar{c}_n)_{n<\omega}$ such that

$$\bar{c}_n \surd_U \bar{b}[\leq n] \quad \text{and} \quad \bar{c}_n \text{ realises } \bigcup_{i\leq n} \mathfrak{p}(\bar{x},\bar{b}_i) \,.$$

We start with $\bar{c}_0 := \bar{a}$. Then $\bar{c}_0$ realises $\mathfrak{p}(\bar{x}, \bar{b}_0)$ and $\bar{c}_0 \sqrt{}_U \bar{b}_0$. For the inductive step, suppose that $\bar{c}_n$ has already been defined. Let $\bar{a}'$ be a realisation of $\mathfrak{p}(\bar{x}, \bar{b}_{n+1})$. Then

$$\bar{a}' \equiv_U \bar{c}_n\,, \quad \bar{a}' \sqrt{}_U \bar{b}_{n+1}\,, \quad \bar{c}_n \sqrt{}_U \bar{b}[\leq n]\,,$$

and $\quad \bar{b}_{n+1} \sqrt{}_U \bar{b}[\leq n]\,,$

which, by the Independence Theorem, implies that there is a tuple $\bar{c}_{n+1}$ such that

$$\bar{c}_{n+1} \equiv_{U\bar{b}_{n+1}} \bar{a}', \quad \bar{c}_{n+1} \equiv_{U\bar{b}[\leq n]} \bar{c}_n\,, \quad \text{and} \quad \bar{c}_{n+1} \sqrt{}_U \bar{b}[\leq n]\bar{b}_{n+1}\,.$$

It follows that $\bar{c}_{n+1}$ realises the types $\operatorname{tp}(\bar{a}'/U\bar{b}_{n+1}) = \mathfrak{p}(\bar{x}, \bar{b}_{n+1})$ and $\operatorname{tp}(\bar{c}_n/U\bar{b}[\leq n]) \supseteq \bigcup_{i\leq n} \mathfrak{p}(\bar{x}, \bar{b}_i)$.

In particular, note that $\bar{c}_{n+1} \equiv_{U\bar{b}[\leq n]} \bar{c}_n$. Hence, having constructed the sequence $(\bar{c}_n)_{n<\omega}$, we can use the Compactness Theorem to find a tuple $\bar{c}$ such that

$$\bar{c} \equiv_{U\bar{b}[\leq n]} \bar{c}_n\,, \quad \text{for all } n < \omega\,.$$

Consequently, $\bar{c}$ realises $\bigcup_{n<\omega} \mathfrak{p}(\bar{x}, \bar{b}_n)$. Furthermore, (INV) and (DEF) implies that $\bar{c} \sqrt{}_U \bar{b}[<\omega]$. □

For symmetric forking relations, we can strengthen Lemma 3.3 as follows.

Theorem 3.16. *If a symmetric forking relation $\downarrow$ satisfies the chain condition for Morley sequences for a class C that is invariant and dense, then $\downarrow = \sqrt[\mathrm{d}]{}$.*

Proof. We have shown in Theorem F3.1.9 that $\sqrt[\mathrm{d}]{} \subseteq \downarrow$, for every symmetric forking relation. Conversely, suppose that $\bar{a} \downarrow_U \bar{b}$. To show that $\bar{a} \sqrt[\mathrm{d}]{}_U \bar{b}$, set $\mathfrak{p}(\bar{x}, \bar{x}') := \operatorname{tp}(\bar{a}\bar{b}/U)$ and let $(\bar{b}_n)_{n<\omega}$ by an indiscernible sequence over U with $\bar{b}_0 = \bar{b}$. By Lemma F3.1.3 (3), it is sufficient to show that there is a tuple realising $\bigcup_{n<\omega} \mathfrak{p}(\bar{x}, \bar{b}_n)$. As $\downarrow$ is right local, we can use Lemma 3.13 to find a set $C \in \mathcal{C}$ containing U such that $(\bar{b}_n)_{n<\omega}$ is a $\downarrow$-Morley sequence over C. Since $\bar{a} \downarrow_U \bar{b}_0$, there is some $\bar{a}' \equiv_{U\bar{b}_0} \bar{a}$ such that $\bar{a}' \downarrow_U C\bar{b}_0$. Set $\mathfrak{p}'(\bar{x}, \bar{x}') := \operatorname{tp}(\bar{a}'\bar{b}_0/C)$. By the chain condition for Morley sequences, the union $\bigcup_{n<\omega} \mathfrak{p}'(\bar{x}, \bar{b}_n)$ does not $\downarrow$-fork over C. In particular, it is consistent. Hence, it follows that there is a tuple realising

$$\bigcup_{n<\omega} \mathfrak{p}(\bar{x}, \bar{b}_n) \subseteq \bigcup_{n<\omega} \mathfrak{p}'(\bar{x}, \bar{b}_n) . \qquad \square$$

We obtain the following characterisation of simple theories.

Theorem 3.17. *Let T be a complete first-order theory. The following statements are equivalent.*

(1) *T is simple.*
(2) *There exists a symmetric forking relation $\downarrow$ satisfying the Independence Theorem for the class of all models.*
(3) *There exists a symmetric forking relation $\downarrow$ satisfying the chain condition for Morley sequences for the class of all models.*
(4) *There exists a symmetric forking relation $\downarrow$ satisfying the chain condition for the class of all models.*

Proof. (4) ⇒ (3) is trivial; (3) ⇒ (1) follows by Theorem 3.16; (1) ⇒ (4) by Lemma 3.5; (1) ⇒ (2) was already proved in Theorem 3.11; and (2) ⇒ (3) follows by Lemma 3.15. □

As an application we consider the theory of the random graph.

Proposition 3.18. *The theory of the random graph is simple.*

Proof. By Theorem 3.16, it is sufficient to prove that the relation

$$A \downarrow^{0}_{U} B \quad :\text{iff} \quad A \cap B \subseteq U$$

is a symmetric forking relation satisfying the Independence Theorem. $\downarrow^0$ obviously satisfies the axioms (INV), (MON), (NOR), (LRF), (BMON), and (SYM).

(LTR) Suppose that $A_2 \downarrow^{\circ}_{A_1} B$ and $A_1 \downarrow^{\circ}_{A_0} B$ where $A_0 \subseteq A_1 \subseteq A_2$. Then $A_2 \cap B \subseteq A_1$ and $A_1 \cap B \subseteq A_0$. Hence,

$$A_2 \cap B \subseteq A_1 \cap B \subseteq A_0 ,$$

which implies that $A_2 \downarrow^{\circ}_{A_0} B$.

(DEF) Suppose that $A \not\downarrow^{\circ}_U B$. Then there is some element $b \in A \cap B \setminus U$. For every element $a \in (x = b)^{\mathbb{M}}$ it follows that $a \not\downarrow^{\circ}_U b$.

(EXT) Suppose that $\bar a \downarrow^{\circ}_U B_0$ and let $B_0 \subseteq B_1$. Using the extension axioms, we can find a tuple $\bar a'$ such that

$$\operatorname{atp}(\bar a'/UB_0) = \operatorname{atp}(\bar a/UB_0) \quad \text{and} \quad (\bar a' \setminus U) \cap B_1 = \varnothing .$$

By ultrahomogeneity, there exists an automorphism $\pi \in \operatorname{Aut} \mathbb{M}_{UB_0}$ mapping $\bar a$ to $\bar a'$. Hence, $\bar a' \equiv_{UB_0} \bar a$ and $\bar a' \downarrow^{\circ}_U B_1$.

(IND) We prove that $\downarrow^{\circ}$ satisfies the Independence Theorem for the class of all subsets of $\mathbb{M}$. Suppose that

$$\bar a \equiv_U \bar b , \quad \bar a \downarrow^{\circ}_U A , \quad \bar b \downarrow^{\circ}_U B , \quad \text{and} \quad A \downarrow^{\circ}_U B .$$

Replacing A and B by, respectively, $A \setminus U$ and $B \setminus U$, we may assume that $A \cap U = \varnothing$ and $B \cap U = \varnothing$. Let

$$\bar d := \bar a \cap U , \quad \bar a' := \bar a \setminus U , \quad \text{and} \quad \bar b' := \bar b \setminus U .$$

Note that $\bar a' \cap (U \cup A) = \varnothing$ and $\bar b' \cap (U \cup B) = \varnothing$. Since U, A, B are disjoint, we can use the extension axioms to find a tuple $\bar c'$ disjoint from $U \cup A \cup B$ such that

$$\operatorname{atp}(\bar c'/UA) = \operatorname{atp}(\bar a'/UA) \quad \text{and} \quad \operatorname{atp}(\bar c'/UB) = \operatorname{atp}(\bar b'/UB) .$$

It follows that

$$\bar c'\bar d \equiv_{UA} \bar a'\bar d , \quad \bar c'\bar d \equiv_{UB} \bar b'\bar d , \quad \text{and} \quad \bar c'\bar d \downarrow^{\circ}_U AB . \qquad \square$$

Part G.

Geometric Model Theory

G1. Stable theories

1. Definable types

A key property of stable theories is the definability of types. The relation $\overset{\mathrm{df}}{\vee\!\!\!/}$ will thus play a major role in this chapter. In the next section, we will study its properties in the context of stable theories. But first, we consider the relation $\overset{\mathrm{df}}{\vee\!\!\!/}$ in an arbitrary theory, where it is usually much less well-behaved: if U and B are small enough it might happen that $\mathrm{tp}(\bar{a}/UB)$ is definable over U just because some formula 'accidentally' is a φ-definition, although it ceases to be a definition for every extension $\mathrm{tp}(\bar{a}/UB')$ with $B' \supseteq B$. Therefore, when investigating a statement of the form $A \overset{\mathrm{df}}{\vee\!\!\!/}_U B$ we usually assume that one of the sets B and U is large. In particular, there is hope that the derived relation $^*(\overset{\mathrm{df}}{\vee\!\!\!/})$ is much better behaved. We start with relating $\overset{\mathrm{df}}{\vee\!\!\!/}$ to the relation $\overset{\mathrm{u}}{\vee\!\!\!/}$.

Lemma 1.1. *Let* $A, A', B, U \subseteq \mathbb{M}$.

(a) $A \overset{\mathrm{df}}{\vee\!\!\!/}_U B$ *and* $B \overset{\mathrm{u}}{\vee\!\!\!/}_U U$ *implies* $B \overset{\mathrm{u}}{\vee\!\!\!/}_U A$.

(b) $A \overset{\mathrm{df}}{\vee\!\!\!/}_U B$, $B \overset{\mathrm{u}}{\vee\!\!\!/}_U A'$, *and* $A \equiv_U A'$ *implies* $A \equiv_{UB} A'$.

Proof. (a) Let $A \overset{\mathrm{df}}{\vee\!\!\!/}_U B$ and $B \overset{\mathrm{u}}{\vee\!\!\!/}_U U$. To show that $B \overset{\mathrm{u}}{\vee\!\!\!/}_U A$, suppose that $\mathbb{M} \vDash \varphi(\bar{a}; \bar{b})$, where $\bar{a} \subseteq A$, $\bar{b} \subseteq B$, and $\varphi(\bar{x}; \bar{y})$ is a formula over U. We have to find a tuple $\bar{c} \subseteq U$ such that $\mathbb{M} \vDash \varphi(\bar{a}; \bar{c})$. Since $\bar{a} \overset{\mathrm{df}}{\vee\!\!\!/}_U \bar{b}$, the type $\mathrm{tp}(\bar{a}/U\bar{b})$ has a φ-definition δ over U. Hence,

$$\mathbb{M} \vDash \varphi(\bar{a}; \bar{b}) \quad \text{implies} \quad \mathbb{M} \vDash \delta(\bar{b}) .$$

Since $\bar{b} \overset{\mathrm{u}}{\vee\!\!\!/}_U U$, there is some $\bar{c} \subseteq U$ with $\mathbb{M} \vDash \delta(\bar{c})$. Consequently, $\mathbb{M} \vDash \varphi(\bar{a}; \bar{c})$.

(b) Let $\bar{a}$ be an enumeration of A and $\bar{a}'$ the corresponding enumeration of A'. For every formula $\varphi(\bar{x};\bar{y})$ over U, we fix a φ-definition $\delta_\varphi(\bar{y})$ of $\operatorname{tp}(\bar{a}/UB)$ over U. It is sufficient to prove that δ_φ is also a φ-definition of $\operatorname{tp}(\bar{a}'/UB)$ since this implies that

$$\mathbb{M} \vDash \varphi(\bar{a};\bar{b}) \quad \text{iff} \quad \mathbb{M} \vDash \delta(\bar{b}) \quad \text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}';\bar{b}),$$

for all $\bar{b} \subseteq U \cup B$.

For a contradiction, suppose that the formula δ_φ is not a φ-definition of $\operatorname{tp}(\bar{a}'/UB)$. Then there exists a tuple $\bar{b} \subseteq U \cup B$ such that

$$\mathbb{M} \vDash \neg(\varphi(\bar{a}';\bar{b}) \leftrightarrow \delta_\varphi(\bar{b})).$$

Since $B \overset{\mathrm{u}}{\surd}_U \bar{a}'$, there is a tuple $\bar{c} \subseteq U$ such that

$$\mathbb{M} \vDash \neg(\varphi(\bar{a}';\bar{c}) \leftrightarrow \delta_\varphi(\bar{c})).$$

As $\bar{a} \equiv_U \bar{a}'$, this implies that

$$\mathbb{M} \vDash \neg(\varphi(\bar{a};\bar{c}) \leftrightarrow \delta_\varphi(\bar{c})).$$

Consequently, δ_φ is no φ-definition of $\operatorname{tp}(\bar{a}/UB)$. A contradiction. □

The relation $\overset{\mathrm{df}}{\surd}$ is particularly well-behaved if the base set is a model.

Lemma 1.2. *Let T be a complete first-order theory, $\mathfrak{M}$ a model of T, and $A \subseteq \mathbb{M}$ a set such that $A \overset{\mathrm{df}}{\surd}_M M$.*

(a) $A \;{}^*(\overset{\mathrm{df}}{\surd})_M\; M$.

(b) *For every set $B \subseteq \mathbb{M}$, there exists a set $A' \equiv_M A$ such that $B \overset{\mathrm{u}}{\surd}_M A'$.*

(c) *If $A', B \subseteq \mathbb{M}$ are sets such that*

$$B \overset{\mathrm{u}}{\surd}_M A, \quad B \overset{\mathrm{u}}{\surd}_M A' \quad \textit{and} \quad A \equiv_M A', \quad \textit{then} \quad A \equiv_{\text{[redacted]}B} A'.$$

Proof. (a) Suppose that $\bar{a} \overset{\mathrm{df}}{\surd}_M M$. To show that $\bar{a} \; {}^{*}(\overset{\mathrm{df}}{\surd})_M \; M$, consider a set $C \supseteq M$. For every formula $\varphi(\bar{x};\bar{y})$ fix a φ-definition $\delta_\varphi(\bar{y})$ of $\mathrm{tp}(\bar{a}/M)$ over M and set

$$\Phi(\bar{x}) := \{\, \varphi(\bar{x};\bar{c}) \mid \bar{c} \subseteq C\,,\ \mathbb{M} \vDash \delta_\varphi(\bar{c}) \,\}\,.$$

Note that $\mathrm{tp}(\bar{a}/M) \subseteq \Phi$. Hence, if $\bar{a}'$ is a tuple satisfying Φ, then $\bar{a}' \equiv_M \bar{a}$ and $\bar{a}' \overset{\mathrm{df}}{\surd}_M C$ since each formula δ_φ is a φ-definition of $\mathrm{tp}(\bar{a}'/C)$.

Thus, it remains to prove that Φ is satisfiable. Consider a finite subset $\Phi_0 = \{\varphi_0(\bar{x};\bar{c}_0), \dots, \varphi_n(\bar{x};\bar{c}_n)\} \subseteq \Phi$. By definition of Φ we have

$$\mathbb{M} \vDash \delta_{\varphi_0}(\bar{c}_0) \wedge \dots \wedge \delta_{\varphi_n}(\bar{c}_n)\,.$$

We have seen in Lemma F2.3.15 that $C \overset{\mathrm{u}}{\surd}_M M$. Hence, we can find tuples $\bar{b}_0, \dots, \bar{b}_n \subseteq M$ such that

$$\mathbb{M} \vDash \delta_{\varphi_0}(\bar{b}_0) \wedge \dots \wedge \delta_{\varphi_n}(\bar{b}_n)\,.$$

This implies that

$$\mathbb{M} \vDash \varphi_0(\bar{a};\bar{b}_0) \wedge \dots \wedge \varphi_n(\bar{a};\bar{b}_n)\,.$$

Consequently, the model $\mathbb{M}$ satisfies Φ_0 if we interpret the variables $\bar{x}$ by $\bar{a}$ and the constants $\bar{c}_i$ by $\bar{b}_i$.

(b) Given $B \subseteq \mathbb{M}$, we can use (a) to find a set $A' \equiv_M A$ such that $A' \overset{\mathrm{df}}{\surd}_M B$. Furthermore, we have $B \overset{\mathrm{u}}{\surd}_M M$, by Lemma F2.3.15. Hence, Lemma 1.1 (a) implies that $B \overset{\mathrm{u}}{\surd}_M A'$.

(c) Let $\bar{a}$ be an enumeration of A and let $\bar{a}'$ be the corresponding enumeration of A'. By (a), we can find a tuple $\bar{a}'' \equiv_M \bar{a}$ such that $\bar{a}'' \overset{\mathrm{df}}{\surd}_M B$. By Lemma 1.1 (b), $B \overset{\mathrm{u}}{\surd}_M \bar{a}$ implies that $\bar{a}'' \equiv_{MB} \bar{a}$. Since $\bar{a}'' \equiv_M \bar{a}'$ and $B \overset{\mathrm{u}}{\surd}_M \bar{a}'$, it follows in the same way that $\bar{a}'' \equiv_{MB} \bar{a}'$. Hence, $\bar{a} \equiv_{MB} \bar{a}'' \equiv_{MB} \bar{a}'$. ☐

Lemma 1.3 (Harrington). *Let $\mathfrak{M}$ be an $\aleph_0$-saturated model, $\mathfrak{p}(\bar{x}), \mathfrak{q}(\bar{y}) \in S^{<\omega}(M)$, and let $\varphi(\bar{x};\bar{y})$ be a formula over M that does not have the order property. If $\delta(\bar{y})$ is a φ-definition of $\mathfrak{p}$ and $\varepsilon(\bar{x})$ a φ-definition of $\mathfrak{q}$, then*

$$\varepsilon(\bar{x}) \in \mathfrak{p}(\bar{x}) \quad \text{iff} \quad \delta(\bar{y}) \in \mathfrak{q}(\bar{y})\,.$$

Proof. Let $\bar{c} \subseteq M$ be the parameters occurring in φ. By induction on n, we construct two sequences $(\bar{a}_n)_{n<\omega}$ and $(\bar{b}_n)_{n<\omega}$ in M as follows. Suppose that we have already defined $\bar{a}_0, \dots, \bar{a}_{n-1}$ and $\bar{b}_0, \dots, \bar{b}_{n-1}$. As $\mathfrak{M}$ is $\aleph_0$-saturated, we can choose a tuple $\bar{b}_n \subseteq M$ realising $\mathfrak{q} \upharpoonright \bar{c}\bar{a}_0 \dots \bar{a}_{n-1}$ and a tuple $\bar{a}_n \subseteq M$ realising $\mathfrak{p} \upharpoonright \bar{c}\bar{b}_0 \dots \bar{b}_n$.

Having constructed $(\bar{a}_n)_{n<\omega}$ and $(\bar{b}_n)_{n<\omega}$ it follows, for $i < k$, that

$$\begin{aligned} \mathfrak{M} \vDash \varphi(\bar{a}_i; \bar{b}_k) \quad &\text{iff} \quad \varphi(\bar{a}_i; \bar{y}) \in \mathfrak{q} \upharpoonright \bar{c}\bar{a}_0 \dots \bar{a}_{k-1} \\ &\text{iff} \quad \mathfrak{M} \vDash \varepsilon(\bar{a}_i) \\ &\text{iff} \quad \varepsilon(\bar{x}) \in \mathfrak{p}, \end{aligned}$$

and, for $i \geq k$,

$$\begin{aligned} \mathfrak{M} \vDash \varphi(\bar{a}_i; \bar{b}_k) \quad &\text{iff} \quad \varphi(\bar{x}; \bar{b}_k) \in \mathfrak{p} \upharpoonright \bar{c}\bar{b}_0 \dots \bar{b}_i \\ &\text{iff} \quad \mathfrak{M} \vDash \delta(\bar{b}_k) \\ &\text{iff} \quad \delta(\bar{x}) \in \mathfrak{q}. \end{aligned}$$

Hence, if $\delta \notin \mathfrak{q}$ and $\varepsilon \in \mathfrak{p}$, then

$$\mathfrak{M} \vDash \varphi(\bar{a}_i; \bar{b}_k) \quad \text{iff} \quad i < k,$$

and φ has the order property. A contradiction. In the same way we obtain a contradiction if we assume that $\delta \in \mathfrak{q}$ and $\varepsilon \notin \mathfrak{p}$. □

We have already seen in Lemma F2.3.3 (a) that $\overset{\mathrm{df}}{\surd} \subseteq \overset{\mathrm{s}}{\surd}$. The converse holds only in special circumstances.

Lemma 1.4. *Suppose that $\mathfrak{M}$ is a κ-saturated and strongly κ-homogeneous model and let $U \subseteq M$ be a set of size $|U| < \kappa$. If $A \subseteq \mathbb{M}$ is a set such that $A \overset{\mathrm{df}}{\surd}_M M$, then*

$$A \overset{\mathrm{df}}{\surd}_U M \quad \text{iff} \quad A \overset{\mathrm{s}}{\surd}_U M.$$

Proof. ($\Rightarrow$) We have seen in Lemma F2.3.3 (a) that $\overset{\mathrm{df}}{\surd} \subseteq \overset{\mathrm{s}}{\surd}$.

($\Leftarrow$) Let $\bar{a} \subseteq A$, let $\varphi(\bar{x};\bar{y})$ be a formula, and let $\delta(\bar{y})$ be a φ-definition of $\operatorname{tp}(\bar{a}/M)$ over M. It is sufficient to show that the relation $\delta^{\mathfrak{M}}$ is definable over U. By definition of $\overset{\mathrm{s}}{\smile\!\!\!\!\!|}$,

$$\bar{b} \equiv_U \bar{b}' \quad \text{implies} \quad \bar{b} \equiv_{U\bar{a}} \bar{b}', \quad \text{for all } \bar{b}, \bar{b}' \subseteq M .$$

Hence, if $\bar{b}, \bar{b}' \subseteq M$ are tuples such that $\bar{b} \equiv_U \bar{b}'$, then

$$\begin{aligned} \mathbb{M} \vDash \delta(\bar{b}) \quad &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a};\bar{b}) \\ &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a};\bar{b}') \quad \text{iff} \quad \mathbb{M} \vDash \delta(\bar{b}') . \end{aligned}$$

Consequently, we have

$$\mathfrak{M} \vDash \delta(\bar{b}) \quad \text{iff} \quad \mathfrak{M} \vDash \delta(\pi(\bar{b})) , \quad \text{for all } \pi \in \operatorname{Aut} \mathfrak{M}_U ,$$

and it follows by Lemma E2.1.10 that $\delta^{\mathfrak{M}}$ is definable over U. □

Another immediate consequence of the inclusion $\overset{\mathrm{df}}{\smile\!\!\!\!\!|} \subseteq \overset{\mathrm{s}}{\smile\!\!\!\!\!|}$ is the corresponding inclusion between the starred relations.

Proposition 1.5. $^*(\overset{\mathrm{df}}{\smile\!\!\!\!\!|}) \subseteq \overset{\mathrm{i}}{\smile\!\!\!\!\!|}$

Proof. We have seen in Lemma F2.3.3 (a) that $\overset{\mathrm{df}}{\smile\!\!\!\!\!|} \subseteq \overset{\mathrm{s}}{\smile\!\!\!\!\!|}$. This implies that $^*(\overset{\mathrm{df}}{\smile\!\!\!\!\!|}) \subseteq {}^*(\overset{\mathrm{s}}{\smile\!\!\!\!\!|}) = \overset{\mathrm{i}}{\smile\!\!\!\!\!|}$. □

We conclude this section with a comparison of $\overset{\mathrm{df}}{\smile\!\!\!\!\!|}$ with $\overset{\mathrm{d}}{\smile\!\!\!\!\!|}$.

Lemma 1.6. *Let $\mathfrak{M}$ be a κ-saturated model and let $U \subseteq M$ be a set of size $|U| < \kappa$. Then*

$$A \overset{\mathrm{df}}{\smile\!\!\!\!\!|}_U M \quad \textit{implies} \quad A \overset{\mathrm{d}}{\smile\!\!\!\!\!|}_U M , \quad \textit{for all } A \subseteq \mathbb{M} .$$

Proof. Suppose that $\bar{a} \overset{\mathrm{df}}{\smile\!\!\!\!\!|}_U M$. To show that $\bar{a} \overset{\mathrm{d}}{\smile\!\!\!\!\!|}_U M$ it is sufficient, by Lemma F3.1.3 and (DEF), to prove that, for every indiscernible sequence

$(\bar{b}_n)_{n<\omega}$ over U with $\bar{b}_0 \subseteq M$ and $|\bar{b}_0| < \aleph_0$, there exists a tuple $\bar{a}' \equiv_{U\bar{b}_0} \bar{a}$ such that

$$\bar{b}_m \equiv_{U\bar{a}'} \bar{b}_n\,, \quad \text{for all } m, n < \omega\,.$$

Hence, let $(\bar{b}_n)_{n<\omega}$ be such a sequence. As M is κ-saturated, we can find tuples $\bar{c}_n \subseteq M$, $n < \omega$, such that

$$(\bar{c}_n)_n \equiv_{U\bar{b}_0} (\bar{b}_n)_n\,.$$

Note that, by Lemma F2.3.3 (a),

$$\bar{a} \overset{\mathrm{df}}{\sqrt{}}_U M \quad \text{implies} \quad \bar{a} \overset{\mathrm{s}}{\sqrt{}}_U M\,.$$

Consequently,

$$\bar{c}_m \equiv_U \bar{c}_n \quad \text{implies} \quad \bar{c}_m \equiv_{U\bar{a}} \bar{c}_n\,, \quad \text{for } m, n < \omega\,.$$

Fixing a tuple $\bar{a}'$ such that

$$\bar{a}(\bar{c}_n)_n \equiv_{U\bar{b}_0} \bar{a}'(\bar{b}_n)_n\,,$$

it follows that

$$\bar{c}_m \equiv_{U\bar{a}} \bar{c}_n \quad \text{implies} \quad \bar{b}_m \equiv_{U\bar{a}'} \bar{b}_n\,, \quad \text{for } m, n < \omega\,. \qquad \square$$

2. Forking in stable theories

In this section we collect properties of preforking relations in stable theories. First, note that we have seen in Corollary F3.2.19 that every stable theory is simple. Hence, in stable theories there is no difference between dividing and forking and both relations are symmetric.

Furthermore, in stable theories the relation $\overset{\mathrm{df}}{\sqrt{}}$ is much better behaved than usual. For instance, we have already seen in Theorem C3.5.17 that $\overset{\mathrm{df}}{\sqrt{}}$ is right reflexive.

Theorem 2.1. *Let T be a stable theory. Then*

$$A \mathrel{\sqrt[\mathrm{df}]{}}_U U\,, \quad \textit{for all } A, U \subseteq \mathbb{M}\,.$$

Let us consider the case where the base set is a model. In this case it will turn out that most preforking relations coincide.

Lemma 2.2. *Let T be a stable theory and $\mathfrak{M}$ a model of T. Then*

$$A \mathop{\downarrow}^{\mathrm{d}}_M B \quad \textit{implies} \quad A \mathrel{\sqrt[\mathrm{u}]{}}_M B\,, \quad \textit{for all } A, B \subseteq \mathbb{M}\,.$$

Proof. Suppose that $\bar{a} \mathop{\downarrow}^{\mathrm{d}}_M \bar{b}$ and $\mathbb{M} \vDash \varphi(\bar{a};\bar{b})$. We have to find a tuple $\bar{a}' \subseteq M$ such that $\mathbb{M} \vDash \varphi(\bar{a}';\bar{b})$.

By Lemma F2.3.15, we have $\bar{a} \mathrel{\sqrt[\mathrm{u}]{}}_M M$. Hence, we can use Proposition F2.4.10 to construct a $\sqrt[\mathrm{u}]{}$-Morley sequence $(\bar{a}^n)_{n<\omega}$ for $\mathrm{tp}(\bar{a}/M)$ over M. Since $\bar{b} \mathop{\downarrow}^{\mathrm{d}}_M \bar{a}$, it follows by Lemma F3.1.3 that there exists a tuple $\bar{b}' \equiv_{M\bar{a}} \bar{b}$ such that $(\bar{a}^n)_{n<\omega}$ is indiscernible over $M \cup \bar{b}'$. Consequently, $\mathbb{M} \vDash \varphi(\bar{a}^n;\bar{b}')$, for all $n < \omega$. By Theorem 2.1, we have

$$\bar{b}' \mathrel{\sqrt[\mathrm{df}]{}}_{M\cup\bigcup_{n<\omega}\bar{a}^n} M \cup \bigcup_{n<\omega} \bar{a}^n\,.$$

Let $\delta(\bar{y})$ be a φ-definition of $\mathrm{tp}(\bar{b}'/M \cup \bigcup_n \bar{a}^n)$ over $M \cup \bigcup_n \bar{a}^n$ and choose $n < \omega$ such that δ is a formula over $M \cup \bar{a}^0 \dots \bar{a}^{n-1}$. Since

$$\bar{a}^n \mathrel{\sqrt[\mathrm{u}]{}}_M M\bar{a}^0 \dots \bar{a}^{n-1} \quad \text{and} \quad \mathbb{M} \vDash \delta(\bar{a}^n)\,,$$

there exists a tuple $\bar{a}' \subseteq M$ such that $\mathbb{M} \vDash \delta(\bar{a}')$. Hence, $\mathbb{M} \vDash \varphi(\bar{a}';\bar{b})$, as desired. □

Corollary 2.3. *Let T be a stable theory and $\mathfrak{M}$ a model of T. Then*

$$A \mathop{\downarrow}^{\mathrm{d}}_M B \quad \textit{implies} \quad A \mathrel{\sqrt[\mathrm{df}]{}}_M B\,, \quad \textit{for all } A, B \subseteq \mathbb{M}\,.$$

Proof. According to Theorem 2.1, we have $A \mathrel{\sqrt[\mathrm{df}]{}}_M M$, which implies that $A \;{}^*(\sqrt[\mathrm{df}]{})_M\; M$, by Lemma 1.2 (a). Hence, we can find a set $A' \equiv_M A$

with $A' \mathop{\surd}^{\mathrm{df}}_M B$. Furthermore, $B \mathop{\downarrow}^{\mathrm{d}}_M A$ implies $B \mathop{\surd}^{\mathrm{u}}_M A$, by Lemma 2.2. Consequently, it follows by Lemma 1.1 (b) that $A' \equiv_{MB} A$. Therefore, $A' \mathop{\surd}^{\mathrm{df}}_M B$ implies that $A \mathop{\surd}^{\mathrm{df}}_M B$. □

The previous results imply that, in a stable theory, the relations $\mathop{\surd}^{\mathrm{u}}$, $\mathop{\surd}^{\mathrm{i}}$, $\mathop{\surd}^{\mathrm{df}}$, $\mathop{\downarrow}^{\mathrm{d}}$, and $\mathop{\downarrow}^{\mathrm{f}}$ are all equivalent, at least over models.

Theorem 2.4. *Let T be a stable theory and $\mathfrak{M}$ a model of T. Then*

$$\begin{aligned} A \mathop{\surd}^{\mathrm{u}}_M B \quad &\text{iff} \quad A \mathop{\surd}^{\mathrm{i}}_M B \quad \text{iff} \quad A \mathop{\surd}^{\mathrm{df}}_M B \\ &\text{iff} \quad A \mathop{\downarrow}^{\mathrm{f}}_M B \quad \text{iff} \quad A \mathop{\downarrow}^{\mathrm{d}}_M B\,. \end{aligned}$$

Proof. We have already seen in Proposition F3.1.12 that

$$A \mathop{\surd}^{\mathrm{u}}_M B \quad \Rightarrow \quad A \mathop{\surd}^{\mathrm{i}}_M B \quad \Rightarrow \quad A \mathop{\downarrow}^{\mathrm{f}}_M B \quad \Rightarrow \quad A \mathop{\downarrow}^{\mathrm{d}}_M B\,.$$

For stable theories, the implication $A \mathop{\downarrow}^{\mathrm{d}}_M B \Rightarrow A \mathop{\surd}^{\mathrm{u}}_M B$ is provided by Lemma 2.2. Furthermore, we can use Corollary 2.3 to show that

$$A \mathop{\downarrow}^{\mathrm{d}}_M B \quad \Rightarrow \quad A \mathop{\surd}^{\mathrm{df}}_M B\,,$$

while Lemma 1.1 (a) implies that

$$A \mathop{\surd}^{\mathrm{df}}_M B \quad \Rightarrow \quad B \mathop{\surd}^{\mathrm{u}}_M A\,.$$

Since we have already proved that, over models, $\mathop{\surd}^{\mathrm{u}}$ is equivalent to $\mathop{\downarrow}^{\mathrm{d}}$, it follows by symmetry of $\mathop{\downarrow}^{\mathrm{d}}$ that

$$A \mathop{\surd}^{\mathrm{df}}_M B \quad \Rightarrow \quad B \mathop{\surd}^{\mathrm{u}}_M A \quad \Leftrightarrow \quad A \mathop{\surd}^{\mathrm{u}}_M B\,.$$ □

There is an even closer connection between $\mathop{\surd}^{\mathrm{i}}$, $^*(\mathop{\surd}^{\mathrm{df}})$, and $\mathop{\downarrow}^{\mathrm{f}}$.

Proposition 2.5. *Let T be a stable theory.*

(a) $\mathop{\surd}^{\mathrm{i}} = {}^*(\mathop{\surd}^{\mathrm{df}})$.

(b) $A \mathop{\downarrow}^{\mathrm{f}}_{U} B$ iff $A \mathop{\vee\!\!\!/}^{\mathrm{i}}_{\mathrm{acl}^{\mathrm{eq}}(U)} B$.

Proof. (a) ($\supseteq$) was already proved in Proposition 1.5. For ($\subseteq$), suppose that $\bar{a} \mathop{\vee\!\!\!/}^{\mathrm{i}}_{U} B$. To show that $\bar{a} \;{}^{*}(\mathop{\vee\!\!\!/}^{\mathrm{df}})_{U}\; B$, we consider a set $C \supseteq B$. We have to find a tuple $\bar{a}' \equiv_{UB} \bar{a}$ with $\bar{a}' \mathop{\vee\!\!\!/}^{\mathrm{df}}_{U} C$. Let $\mathfrak{M}$ be a $|U|^{+}$-saturated and strongly $|U|^{+}$-homogeneous model of T that contains $U \cup C$. Since $\bar{a} \mathop{\vee\!\!\!/}^{\mathrm{i}}_{U} B$, there exists a tuple $\bar{a}' \equiv_{UB} \bar{a}$ with $\bar{a}' \mathop{\vee\!\!\!/}^{\mathrm{s}}_{U} M$. Furthermore, we have seen in Theorem 2.1 that $\bar{a}' \mathop{\vee\!\!\!/}^{\mathrm{df}}_{M} M$. Consequently, it follows by Lemma 1.4 that $\bar{a}' \mathop{\vee\!\!\!/}^{\mathrm{df}}_{U} M$. In particular, $\bar{a}' \mathop{\vee\!\!\!/}^{\mathrm{df}}_{U} C$.

(b) ($\Leftarrow$) According to Proposition F3.1.12, $A \mathop{\vee\!\!\!/}^{\mathrm{i}}_{\mathrm{acl}^{\mathrm{eq}}(U)} B$ implies that $A \mathop{\downarrow}^{\mathrm{f}}_{\mathrm{acl}^{\mathrm{eq}}(U)} B$. Moreover, we have $\mathrm{acl}^{\mathrm{eq}}(U) \mathop{\downarrow}^{\mathrm{f}}_{U} A$, by Corollary F2.2.12 and Lemma F3.1.8. By symmetry and transitivity, it therefore follows that $A \mathop{\downarrow}^{\mathrm{f}}_{U} B$.

($\Rightarrow$) Suppose that $\bar{a} \mathop{\downarrow}^{\mathrm{f}}_{U} B$. We will prove that $\bar{a} \;{}^{*}(\mathop{\vee\!\!\!/}^{\mathrm{df}})_{\mathrm{acl}^{\mathrm{eq}}(U)}\; B$. Then the claim follows by (a). Hence, consider a set $C \supseteq \mathrm{acl}^{\mathrm{eq}}(U) \cup B$. We fix a $(|T| \oplus |U|)^{+}$-saturated and strongly $(|T| \oplus |U|)^{+}$-homogeneous model $\mathfrak{N}$ containing C and a tuple $\bar{a}' \equiv_{UB} \bar{a}$ with $\bar{a}' \mathop{\downarrow}^{\mathrm{f}}_{U} N$. We will prove that

$$\mathrm{Gb}(\mathrm{tp}(\bar{a}'/N)) \subseteq M^{\mathrm{eq}}, \quad \text{for every model } \mathfrak{M} \text{ with } U \subseteq M \subseteq N.$$

By Lemma E2.1.9, this implies that $\mathrm{Gb}(\mathrm{tp}(\bar{a}'/N)) \subseteq \mathrm{acl}^{\mathrm{eq}}(U)$. Consequently, we can use Lemma E2.3.10 to show that $\mathrm{tp}(\bar{a}'/N)$ is definable over $\mathrm{acl}^{\mathrm{eq}}(U)$. In particular, $\bar{a}' \mathop{\vee\!\!\!/}^{\mathrm{df}}_{\mathrm{acl}^{\mathrm{eq}}(U)} C$.

It remains to prove the above claim. Consider a model $\mathfrak{M}$ with $U \subseteq M \subseteq N$. Then $\bar{a}' \mathop{\downarrow}^{\mathrm{f}}_{M} N$ implies that $\bar{a}' \mathop{\vee\!\!\!/}^{\mathrm{df}}_{M} N$, by Theorem 2.4. Therefore, it follows by Lemma E2.3.8 that

$$\mathrm{Gb}(\mathrm{tp}(\bar{a}'/N)) \subseteq \mathrm{dcl}^{\mathrm{eq}}(M) = M^{\mathrm{eq}}. \qquad \square$$

Corollary 2.6. *In a stable theory,*

$$A \mathop{\downarrow}^{\mathrm{f}}_{U} B \quad \textit{implies} \quad A \mathop{\vee\!\!\!/}^{\mathrm{df}}_{\mathrm{acl}^{\mathrm{eq}}(U)} B\,.$$

3. *Stationary types*

In this section we study types with a unique free extension over every set of parameters. Such types are called *stationary.*

Definition 3.1. A type $\mathfrak{p}$ over U is *stationary* if, for every set $C \subseteq \mathbb{M}$, $\mathfrak{p}$ has a unique free extension to a complete type over $U \cup C$.

We start by proving that stationary types exist.

Proposition 3.2. *In a stable theory, every type over a set of the form* $\mathrm{acl}^{\mathrm{eq}}(U)$ *is stationary.*

Proof. Note that a type $\mathfrak{p}(\bar{x})$ is stationary if, and only if, for every finite tuple $\bar{x}' \subseteq \bar{x}$ of variables, the restriction $\mathfrak{p} \upharpoonright \bar{x}'$ is stationary. Hence, it is sufficient to consider types $\mathfrak{p} \in S^{<\omega}(\mathrm{acl}^{\mathrm{eq}}(U))$. Let $C \supseteq \mathrm{acl}^{\mathrm{eq}}(U)$ be a set and suppose that $\bar{a}$ and $\bar{a}'$ are two realisations of $\mathfrak{p}$ with

$$\bar{a} \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{f}}_{\mathrm{acl}^{\mathrm{eq}}(U)} C \quad \text{and} \quad \bar{a}' \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{f}}_{\mathrm{acl}^{\mathrm{eq}}(U)} C .$$

We have to show that $\bar{a} \equiv_C \bar{a}'$. Hence, consider a formula $\varphi(\bar{x}; \bar{c})$ with $\bar{c} \subseteq C$. We choose an $\aleph_0$-saturated model $\mathfrak{M}$ containing $C \cup \bar{a}\bar{a}'$. There are tuples $\bar{a}_* \equiv_C \bar{a}$ and $\bar{a}'_* \equiv_C \bar{a}'$ with

$$\bar{a}_* \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{f}}_{\mathrm{acl}^{\mathrm{eq}}(U)} M \quad \text{and} \quad \bar{a}'_* \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{f}}_{\mathrm{acl}^{\mathrm{eq}}(U)} M .$$

Since $\bar{c} \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{f}}_{\mathrm{acl}^{\mathrm{eq}}(U)} \bar{a}$, there is a tuple $\bar{c}_* \equiv_{\mathrm{acl}^{\mathrm{eq}}(U) \cup \bar{a}} \bar{c}$ with

$$\bar{c}_* \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{f}}_{\mathrm{acl}^{\mathrm{eq}}(U)} M .$$

By Corollary 2.6, the types

$$\mathfrak{q} := \mathrm{tp}(\bar{a}_*/M) , \quad \mathfrak{q}' := \mathrm{tp}(\bar{a}'_*/M) , \quad \text{and} \quad \mathfrak{r} := \mathrm{tp}(\bar{c}_*/M)$$

are definable over $\mathrm{acl}^{\mathrm{eq}}(U)$. Let $\delta(\bar{y})$, $\delta'(\bar{y})$, and $\varepsilon(\bar{x})$ be the corresponding φ-definitions. By Lemma 1.3, it follows that

$$\begin{aligned}
\mathbb{M} \vDash \varphi(\bar{a};\bar{c}) \quad &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}_*;\bar{c}) \\
&\text{iff} \quad \mathbb{M} \vDash \delta(\bar{c}) \\
&\text{iff} \quad \mathbb{M} \vDash \delta(\bar{c}_*) \\
&\text{iff} \quad \delta(\bar{y}) \in \mathfrak{r} \\
&\text{iff} \quad \varepsilon(\bar{y}) \in \mathfrak{q} \\
&\text{iff} \quad \varepsilon(\bar{y}) \in \mathfrak{q}' \\
&\text{iff} \quad \delta'(\bar{y}) \in \mathfrak{r} \\
&\text{iff} \quad \mathbb{M} \vDash \delta'(\bar{c}_*) \\
&\text{iff} \quad \mathbb{M} \vDash \delta'(\bar{c}) \\
&\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}'_*;\bar{c}) \quad \text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}';\bar{c})\,. \qquad \square
\end{aligned}$$

Corollary 3.3. *In a stable theory types over models are stationary.*

Proof. Note that every type over a model M has a unique extension to a type over $M^{\mathrm{eq}} = \mathrm{dcl}^{\mathrm{eq}}(M)$, which is an algebraically closed set. Hence, the claim follows by Proposition 3.2. □

In Proposition 3.7 below we will present a characterisation of stationary types in terms of the relation $\sqrt[\mathrm{i}]{}$. We start with two technical lemmas. In the first one, we prove that all free extensions of a given type are conjugate.

Lemma 3.4. *Let T be a stable theory, $\kappa > |T|$ a cardinal, $\mathfrak{M}$ a strongly κ-homogeneous model of T, and $U \subseteq M$ a set of size $|U| < \kappa$. If*

$$\bar{a} \equiv_U \bar{a}'\,, \quad \bar{a} \mathop{\downarrow}^{\mathrm{f}}_U M\,, \quad \text{and} \quad \bar{a}' \mathop{\downarrow}^{\mathrm{f}}_U M\,,$$

then there exists an automorphism $\pi \in \mathrm{Aut}\,\mathbb{M}_U$ such that $\pi(\bar{a}') = \bar{a}$ and $\pi[M] = M$.

Proof. Since $\bar{a} \equiv_U \bar{a}'$, there exists an automorphism $\pi \in \operatorname{Aut} \mathbb{M}_U$ with $\pi(\bar{a}') = \bar{a}$. By Corollary E2.1.7, we have

$$\pi[\operatorname{acl}^{\mathrm{eq}}(U)] = \operatorname{acl}^{\mathrm{eq}}(U) .$$

As $\mathfrak{M}$ is strongly $|\operatorname{acl}^{\mathrm{eq}}(U)|^+$-homogeneous, we can find an automorphism $\sigma_0 \in \operatorname{Aut} \mathfrak{M}$ with

$$\sigma_0 \upharpoonright \operatorname{acl}^{\mathrm{eq}}(U) = \pi \upharpoonright \operatorname{acl}^{\mathrm{eq}}(U) .$$

Let $\sigma \in \operatorname{Aut} \mathbb{M}$ be an extension of σ_0. For every formula $\varphi(\bar{x}; \bar{c})$ with parameters $\bar{c} \subseteq \operatorname{acl}^{\mathrm{eq}}(U)$, it follows that

$$\begin{aligned} \mathbb{M} \vDash \varphi(\sigma(\bar{a}'); \bar{c}) \quad &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}'; \sigma^{-1}(\bar{c})) \\ &\text{iff} \quad \mathbb{M} \vDash \varphi(\pi(\bar{a}'); \pi(\sigma^{-1}(\bar{c}))) \\ &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}; \bar{c}) . \end{aligned}$$

Hence, $\sigma(\bar{a}') \equiv_{\operatorname{acl}^{\mathrm{eq}}(U)} \bar{a}$. By invariance, we have

$$\sigma(\bar{a}') \mathop{\downarrow}^{\mathrm{f}}_{\operatorname{acl}^{\mathrm{eq}}(U)} M \quad \text{and} \quad \bar{a} \mathop{\downarrow}^{\mathrm{f}}_{\operatorname{acl}^{\mathrm{eq}}(U)} M .$$

Moreover, $\operatorname{tp}(\bar{a}/\operatorname{acl}^{\mathrm{eq}}(U))$ is stationary according to Proposition 3.2. Therefore, $\sigma(\bar{a}') \equiv_M \bar{a}$ and there exists an automorphism $\rho \in \operatorname{Aut} \mathbb{M}_M$ mapping $\sigma(\bar{a}')$ to $\bar{a}$. Since

$$\rho[\sigma[M]] = \rho[\sigma_0[M]] = \rho[M] = M ,$$

the composition $\rho \circ \sigma \in \operatorname{Aut} \mathbb{M}_U$ is the desired automorphism mapping $\bar{a}'$ to $\bar{a}$. □

The second lemma characterises those free extensions that are unique.

Definition 3.5. We write

$$\bar{a} \mathop{\downarrow}^{!}_{U} B \quad :\text{iff} \quad \operatorname{tp}(\bar{a}/UB) \text{ is the unique free extension of } \operatorname{tp}(\bar{a}/U) \text{ over } U \cup B .$$

Lemma 3.6. *Let T be a stable theory, $\kappa > |T|$ a cardinal, $\mathfrak{M}$ a strongly κ-homogeneous model of T, and $U \subseteq M$ a set of size $|U| < \kappa$. Then*

$$\bar{a} \mathop{\smile\!\!\!\!\!|}^{!}_{U} M \quad \text{iff} \quad \bar{a} \mathop{\smile\!\!\!\!\!|}^{\mathrm{f}}_{U} M \quad \textit{and} \quad \bar{a} \mathop{\sqrt[\mathrm{s}]{}}_{U} M\,.$$

Proof. ($\Leftarrow$) As $\mathrm{tp}(\bar{a}/M)$ is a free extension of $\mathrm{tp}(\bar{a}/U)$, we only need to prove uniqueness. If $\mathrm{tp}(\bar{a}/M)$ and $\mathrm{tp}(\bar{a}'/M)$ are two free extensions of $\mathrm{tp}(\bar{a}/U)$, we can use Lemma 3.4 to find an automorphism $\pi \in \mathrm{Aut}\,\mathbb{M}_U$ with $\pi(\bar{a}') = \bar{a}$ and $\pi[M] = M$. Hence, for every formula $\varphi(\bar{x};\bar{y})$ over U and every $\bar{b} \subseteq M$, $\bar{a} \mathop{\sqrt[\mathrm{s}]{}}_{U} M$ implies that

$$\mathbb{M} \vDash \varphi(\bar{a}';\bar{b}) \quad \text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a};\pi(\bar{b})) \quad \text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a};\bar{b})\,.$$

Consequently, $\bar{a} \equiv_M \bar{a}'$.

($\Rightarrow$) If $\bar{a} \mathop{\not\smile\!\!\!\!\!|}^{\mathrm{f}}_{U} M$, the type $\mathrm{tp}(\bar{a}/M)$ is not a free extension of $\mathrm{tp}(\bar{a}/U)$ and we are done. Hence, suppose that $\bar{a} \mathop{\smile\!\!\!\!\!|}^{\mathrm{f}}_{U} M$ and $\bar{a} \mathop{\not\sqrt[\mathrm{s}]{}}_{U} M$. We claim that $\mathrm{tp}(\bar{a}/U)$ has at least two free extensions over M. By assumption, we can find finite tuples $\bar{b}, \bar{b}' \subseteq M$ with $\bar{b} \equiv_U \bar{b}'$ and $\bar{b} \not\equiv_{U\bar{a}} \bar{b}'$. Let $\varphi(\bar{x};\bar{y})$ be a formula over U such that

$$\mathbb{M} \vDash \varphi(\bar{a};\bar{b}) \wedge \neg\varphi(\bar{a};\bar{b}')\,.$$

Since $\bar{b} \equiv_U \bar{b}'$ and $\mathfrak{M}$ is strongly κ-homogeneous, there exists an automorphism $\pi_0 \in \mathrm{Aut}\,\mathfrak{M}_U$ mapping $\bar{b}$ to $\bar{b}'$. Let π be an automorphism of $\mathbb{M}$ extending π_0. Then

$$\mathbb{M} \vDash \varphi(\bar{a};\bar{b}) \quad \text{implies} \quad \mathbb{M} \vDash \varphi(\pi(\bar{a});\bar{b}')\,.$$

Hence, $\pi(\bar{a}) \not\equiv_M \bar{a}$. Furthermore,

$$\bar{a} \mathop{\smile\!\!\!\!\!|}^{\mathrm{f}}_{U} M \quad \text{implies} \quad \pi(\bar{a}) \mathop{\smile\!\!\!\!\!|}^{\mathrm{f}}_{U} \pi[M]\,.$$

Since $\pi[M] = \pi_0[M] = M$, it follows that $\mathrm{tp}(\bar{a}/M)$ and $\mathrm{tp}(\pi(\bar{a})/M)$ are two different free extensions of $\mathrm{tp}(\bar{a}/U)$. □

Proposition 3.7. *Let T be stable. Then*

$$\mathrm{tp}(\bar{a}/U) \textit{ is stationary} \quad \text{iff} \quad \bar{a} \mathop{\sqrt[\mathrm{i}]{}}_{U} U\,.$$

Proof. ($\Rightarrow$) Suppose that $\mathfrak{p} := \mathrm{tp}(\bar{a}/U)$ is stationary. Since $\sqrt[i]{} = {}^*(\sqrt[s]{})$, it is sufficient to show that $\bar{a}\ {}^*(\sqrt[s]{})_U\ U$. Hence, consider a set $B \supseteq U$. We fix a strongly $(|T|^+ \oplus |U|^+)$-homogeneous model $\mathfrak{M}$ containing B. Let $\mathfrak{q}$ be the unique free extension of $\mathfrak{p}$ to M and let $\bar{a}'$ be a realisation of $\mathfrak{q}$. Then $\bar{a}' \equiv_U \bar{a}$ and it follows by Lemma 3.6 that $\bar{a}' \sqrt[s]{}_U M$. In particular, we have $\bar{a}' \sqrt[s]{}_U B$.

($\Leftarrow$) Suppose that $\bar{a} \sqrt[i]{}_U U$. To show that $\mathfrak{p} := \mathrm{tp}(\bar{a}/U)$ is stationary, consider a set $B \supseteq U$. We fix a strongly $(|T|^+ \oplus |U|^+)$-homogeneous model $\mathfrak{M}$ containing B. Let $\bar{a}' \equiv_U \bar{a}$ be a tuple with $\bar{a}' \sqrt[i]{}_U M$. By Proposition F3.1.12, it follows that $\bar{a}' \mathop{\downarrow}^{\mathrm{f}}_U M$ and $\bar{a}' \sqrt[s]{}_U M$. Therefore, Lemma 3.6 implies that $\bar{a}' \mathop{\downarrow}^{!}_U M$. In particular, $\mathrm{tp}(\bar{a}'/B)$ is the unique free extension of $\mathfrak{p}$ over B. $\square$

As an application of stationary types, we present the following topological characterisation of the set of free extensions of a type.

Theorem 3.8 (Open Mapping Theorem). *Let T be a stable theory, $U \subseteq A$ sets, and let $\mathfrak{F}^{\bar{s}}(A/U)$ denote the subspace of $\mathfrak{S}^{\bar{s}}(A)$ consisting of all types that do not fork over U.*

(a) *$F^{\bar{s}}(A/U)$ is a closed subset of $\mathfrak{S}^{\bar{s}}(A)$.*

(b) *The restriction map $\rho : \mathfrak{F}^{\bar{s}}(A/U) \to \mathfrak{S}^{\bar{s}}(U) : \mathfrak{p} \mapsto \mathfrak{p}|_U$ is continuous, closed, open, and surjective.*

Proof. (a) We use (DEF) to fix, for every type $\mathfrak{p} \in S^{\bar{s}}(A) \setminus F^{\bar{s}}(A/U)$, a formula $\varphi_{\mathfrak{p}}(\bar{x}) \in \mathfrak{p}$ such that

$$\mathbb{M} \vDash \varphi_{\mathfrak{p}}(\bar{c}) \quad \text{implies} \quad \bar{c} \mathop{\not\downarrow}^{\mathrm{f}}_U A\,.$$

Setting

$$\Phi := \{\, \neg\varphi_{\mathfrak{p}} \mid \mathfrak{p} \in S^{\bar{s}}(A) \setminus F^{\bar{s}}(A/U) \,\}$$

it follows that $\Phi \subseteq \mathfrak{p}$, for every $\mathfrak{p} \in F^{\bar{s}}(A/U)$, while $\neg\varphi_{\mathfrak{p}} \in \Phi \setminus \mathfrak{p}$, for every $\mathfrak{p} \notin F^{\bar{s}}(A/U)$. Hence,

$$F^{\bar{s}}(A/U) = \langle \Phi \rangle_{\mathfrak{S}^{\bar{s}}(A)}\,,$$

which is a closed set.

(b) We have seen in Corollary c3.2.22 that the restriction map

$$\rho^+ : \mathfrak{S}^{\bar{s}}(A) \to \mathfrak{S}^{\bar{s}}(U) : \mathfrak{p} \mapsto \mathfrak{p}|_U$$

is continuous, closed, and surjective. By (a) it follows that the restriction

$$\rho = \rho^+ \restriction F^{\bar{s}}(A/U) : \mathfrak{F}^{\bar{s}}(A/U) \to \mathfrak{S}^{\bar{s}}(U)$$

is also continuous and closed. Furthermore, it follows by (EXT) that every type over U has a free extension to a type over A. Hence, ρ is surjective and it remains to prove that it is open.

First, we consider the case where A is a strongly $|U|^+$-homogeneous model. Every open set $O \subseteq F^{\bar{s}}(A/U)$ is a union of basic open sets of the form

$$\langle\varphi\rangle_{\mathfrak{F}^{\bar{s}}(A/U)} := \{\, \mathfrak{p} \in F^{\bar{s}}(A/U) \mid \varphi \in \mathfrak{p} \,\}\,.$$

Therefore it is sufficient to prove that the image of a basic open set is open. Let $\varphi(\bar{x};\bar{y})$ be a formula over U, $\bar{c} \subseteq A$ parameters, and let

$$\overline{\langle\varphi(\bar{x};\bar{c})\rangle}_{\mathfrak{F}^{\bar{s}}(A/U)} := \bigcup \big\{\, \big\langle\varphi(\bar{x};\pi(\bar{c}))\big\rangle_{\mathfrak{F}^{\bar{s}}(A/U)} \;\big|\; \pi \in \operatorname{Aut}\mathfrak{M}_U \,\big\}$$

be the closure of $\langle\varphi(\bar{x};\bar{c})\rangle_{\mathfrak{F}^{\bar{s}}(A/U)}$ under conjugates. Being a union of open sets, this set is also open. Furthermore,

$$\rho^{-1}\big[\rho\big[\langle\varphi(\bar{x};\bar{c})\rangle_{\mathfrak{F}^{\bar{s}}(A/U)}\big]\big] = \overline{\langle\varphi(\bar{x};\bar{c})\rangle}_{\mathfrak{F}^{\bar{s}}(A/U)}\,,$$

since, for types $\mathfrak{p}, \mathfrak{q} \in F^{\bar{s}}(A/U)$ with $\mathfrak{p}|_U = \mathfrak{q}|_U$, we can use Lemma 3.4 to find an automorphism $\pi \in \operatorname{Aut}\mathfrak{M}_U$ with $\pi(\mathfrak{p}) = \mathfrak{q}$.

For a type $\mathfrak{p}_0 \in S^{\bar{s}}(U)$, it follows that

$$\begin{aligned}
&\mathfrak{p}_0 \notin \rho\big[\langle\varphi(\bar{x};\bar{c})\rangle_{\mathfrak{F}^{\bar{s}}(A/U)}\big]\\
\text{iff}\quad &\rho^{-1}(\mathfrak{p}_0) \cap \langle\varphi(\bar{x};\bar{c})\rangle_{\mathfrak{F}^{\bar{s}}(A/U)} = \varnothing\\
\text{iff}\quad &\rho^{-1}(\mathfrak{p}_0) \cap \overline{\langle\varphi(\bar{x};\bar{c})\rangle}_{\mathfrak{F}^{\bar{s}}(A/U)} = \varnothing
\end{aligned}$$

$$\begin{array}{ll} \text{iff} & \rho^{-1}(\mathfrak{p}_0) \setminus \overline{\langle \varphi(\bar{x};\bar{c}) \rangle_{\mathfrak{F}^{\bar{s}}(A/U)}} \neq \emptyset \\ \text{iff} & \mathfrak{p}_0 \in \rho\big[F^{\bar{s}}(A/U) \setminus \overline{\langle \varphi(\bar{x};\bar{c}) \rangle_{\mathfrak{F}^{\bar{s}}(A/U)}}\big]. \end{array}$$

Hence, the complement of $\rho[\langle \varphi(\bar{x};\bar{c}) \rangle_{\mathfrak{F}^{\bar{s}}(A/U)}]$ is the image of a closed set. As we have shown above that the map ρ is closed, it follows that the complement is closed and the image $\rho[\langle \varphi(\bar{x};\bar{c}) \rangle_{\mathfrak{F}^{\bar{s}}(A/U)}]$ is open.

It remains to prove the general case. We fix a strongly $|U|^+$-homogeneous model $\mathfrak{M}$ containing A and let

$$\rho' : \mathfrak{F}^{\bar{s}}(M/U) \to \mathfrak{F}^{\bar{s}}(A/U)$$

be the corresponding restriction map. Consider a basic open set

$$\langle \varphi(\bar{x};\bar{c}) \rangle_{\mathfrak{F}^{\bar{s}}(A/U)}$$

in $\mathfrak{F}^{\bar{s}}(A/U)$. Then $\langle \varphi(\bar{x};\bar{c}) \rangle_{\mathfrak{F}^{\bar{s}}(M/U)}$ is basic open in $\mathfrak{F}^{\bar{s}}(M/U)$ and

$$\langle \varphi(\bar{x};\bar{c}) \rangle_{\mathfrak{F}^{\bar{s}}(M/U)} = (\rho')^{-1}\big[\langle \varphi(\bar{x};\bar{c}) \rangle_{\mathfrak{F}^{\bar{s}}(A/U)}\big].$$

As ρ' is surjective, we have

$$\begin{aligned} \rho\big[\langle \varphi(\bar{x};\bar{c}) \rangle_{\mathfrak{F}^{\bar{s}}(A/U)}\big] &= \rho\big[\rho'\big[(\rho')^{-1}\big[\langle \varphi(\bar{x};\bar{c}) \rangle_{\mathfrak{F}^{\bar{s}}(A/U)}\big]\big]\big] \\ &= (\rho \circ \rho')\big[\langle \varphi(\bar{x};\bar{c}) \rangle_{\mathfrak{F}^{\bar{s}}(M/U)}\big] \end{aligned}$$

This set is open, since we have shown above that the composition $\rho \circ \rho'$ is an open map. □

4. *The multiplicity of a type*

Most types have several free extensions. In this section we study their number. We will prove in Theorem 4.6 below that a theory is stable if, and only if, the number of such extensions is bounded.

Definition 4.1. Let T be a theory and $\surd$ a forking relation.

(a) The $\surd$*-multiplicity* $\mathrm{mult}_{\surd}(\mathfrak{p})$ of a type $\mathfrak{p}$ over A is the minimal cardinal κ such that, for every set $B \supseteq A$, there are at most κ complete types over B that are $\surd$-free extensions of $\mathfrak{p}$. If such a cardinal κ does not exist, we set $\mathrm{mult}_{\surd}(\mathfrak{p}) := \infty$. For $\surd = \overset{\mathrm{f}}{\surd}$, we drop the subscript and simply write $\mathrm{mult}(\mathfrak{p})$.

(b) The *multiplicity* $\mathrm{mult}(\surd)$ of $\surd$ is the maximal $\surd$-multiplicity of some complete type (with finitely many variables). If there is no maximum, we set $\mathrm{mult}(\surd) := \infty$.

We start by proving that, in a stable theory, every type has bounded multiplicity.

Lemma 4.2. *Let T be a stable theory.*

(a) *For every type $\mathfrak{p} \in S^{\bar{s}}(A)$ with $|\bar{s}| < \omega$, there exists some model $\mathfrak{M}$ of T of size $|M| \leq |T|$ such that*

$$\mathrm{mult}_{\downarrow^{\mathrm{f}}}(\mathfrak{p}) \leq |S^{\bar{s}}(M)|\,.$$

(b) $\mathrm{mult}(\downarrow^{\mathrm{f}}) \leq \sup \big\{ |S^{<\omega}(U)| \bigm| |U| \leq |T| \big\} \leq 2^{|T|}$

Proof. (a) For every type $\mathfrak{p} \in S^{\bar{s}}(A)$, there exists a set $U \subseteq A$ of size

$$|U| < \mathrm{loc}_0(\downarrow^{\mathrm{f}}) \leq \mathrm{fc}(\downarrow^{\mathrm{f}}) \leq |T|^+$$

such that $\mathfrak{p}$ does not fork over U. Since $|U| \leq |T|$, we can find a model $\mathfrak{M}$ of size $|M| = |T|$ containing U. Since every type has at least one free extension over any given set, it is sufficient to bound the number of free extensions of $\mathfrak{p}$ over sets B containing M. Hence, let $\mathfrak{q} \in S^{\bar{s}}(B)$ be an extension of $\mathfrak{p}$ with $B \supseteq M$. We have seen in Corollary 3.3 that types over models are stationary. Hence, $\mathfrak{q}$ is the unique free extension of $\mathfrak{q}|_M$. Consequently, if $\mathfrak{q}, \mathfrak{q}' \in S^{\bar{s}}(B)$ are distinct free extensions of $\mathfrak{p}$, then $\mathfrak{q}|_M \neq \mathfrak{q}'|_M$. Therefore, $\mathfrak{p}$ has at most $|S^{\bar{s}}(M)|$ free extensions.

(b) The first inequality follows immediately from (a). The second one follows from the fact that there are at most $2^{|T|}$ types over a set of size $|T|$. □

A characterisation of stable theories

Recall that we write $\mathfrak{p} \leq_{\sqrt{}} \mathfrak{q}$ if $\mathfrak{q}$ is a $\sqrt{}$-free extension of $\mathfrak{p}$. This is a definition of $\leq_{\sqrt{}}$ in terms of a given preforking relation $\sqrt{}$. Conversely, given an extension relation $\leq$ we can recover a corresponding preforking relation $\sqrt{}$. In the remainder of this section we present two characterisations of stable theories one in terms of the extension relation $\leq$ and one in terms of the parameter $\mathrm{mult}(\downarrow^{\mathrm{f}})$.

Proposition 4.3. *If* $\downarrow$ *is a symmetric forking relation with* $\mathrm{mult}(\downarrow) < \infty$*, then* $\leq_{\downarrow}$ *satisfies the following conditions:*

(INV) *Invariance.* $\mathfrak{p} \leq \mathfrak{q}$ *implies* $\pi(\mathfrak{p}) \leq \pi(\mathfrak{q})$*, for every automorphism* $\pi \in \mathrm{Aut}(\mathbb{M})$.

(LC) *Local Character. There exists a cardinal* κ *such that, for every set* U *and every type* $\mathfrak{p} \in S^{<\omega}(U)$*, there exists a subset* $U_0 \subseteq U$ *of size* $|U_0| < \kappa$ *such that* $\mathfrak{p} \upharpoonright U_0 \leq \mathfrak{p}$.

(BND) *Boundedness. For every type* $\mathfrak{p} \in S^{<\omega}(U)$*, there exists a cardinal* μ *such that, for every set* $C \subseteq \mathbb{M}$*,* $\mathfrak{p}$ *has at most* μ *extensions* $\mathfrak{q} \in S^{<\omega}(U \cup C)$ *with* $\mathfrak{p} \leq \mathfrak{q}$.

(EXT) *Extension. For every* $\mathfrak{p} \in S^{<\omega}(U)$ *and every set* $C \subseteq \mathbb{M}$*, there exists some type* $\mathfrak{q} \in S^{<\omega}(U \cup C)$ *with* $\mathfrak{p} \leq \mathfrak{q}$.

(TR) *Transitivity.* $\mathfrak{p} \leq \mathfrak{q} \leq \mathfrak{r}$ *implies* $\mathfrak{p} \leq \mathfrak{r}$.

(MON) *Monotonicity.* $\mathfrak{p} \leq \mathfrak{r}$ *implies* $\mathfrak{p} \leq \mathfrak{q}$*, for all* $\mathfrak{p} \subseteq \mathfrak{q} \subseteq \mathfrak{r}$.

Proof. (BND) holds since $\mathrm{mult}(\downarrow) < \infty$. The other axioms follow from the fact that $\downarrow$ is a symmetric forking relation: (INV) follows by invariance; (LC) follows by right locality; (EXT) follows by the extension axiom; (TR) follows by left transitivity and symmetry; and (MON) follows by monotonicity. □

For the converse statement, we need a technical lemma.

Lemma 4.4. *Let $\mathfrak{M}$ be a κ^+-saturated, strongly κ^+-homogeneous model, $U \subseteq M$ a set of size $|U| < \kappa$, and $\bar a \in \mathbb{M}^{<\omega}$. If $\bar a \not\smile^{\mathrm{d}}_U M$, then* $\mathrm{tp}(\bar a/M)$ *has at least κ conjugates over U.*

Proof. If $\bar a \not\smile^{\mathrm{d}}_U M$, there exists a finite tuple $\bar b \subseteq M$ such that $\bar a \not\smile^{\mathrm{d}}_U \bar b$. By Lemma F3.1.3, we can find an indiscernible sequence $(\bar b_n)_{n<\omega}$ over U with $\bar b = \bar b_0$ such that, for each tuple $\bar a' \equiv_{U\bar b} \bar a$, there are indices $m, n < \omega$ with

$$\bar b_m \not\equiv_{U\bar a'} \bar b_n \,.$$

Fix a formula $\varphi(\bar x; \bar y)$ over U such that

$$\mathbb{M} \vDash \varphi(\bar a; \bar b_m) \wedge \neg\varphi(\bar a; \bar b_n)\,, \quad \text{for some } m, n < \omega\,.$$

Replacing φ be its negation, if necessary, we may assume that there are infinitely many indices n such that $\mathbb{M} \vDash \neg\varphi(\bar a; \bar b_n)$. By compactness, it follows that there exists a tuple $\bar a' \equiv_U \bar a$ and an indiscernible sequence $(\bar b'_\alpha)_{\alpha<\kappa}$ of length κ such that

$$\mathbb{M} \vDash \varphi(\bar a'; \bar b'_\alpha) \quad \text{iff} \quad \alpha = 0\,.$$

As $\mathfrak{M}$ is κ^+-saturated, we may choose the sequence $(\bar b'_\alpha)_{\alpha<\kappa}$ to be in M. By strong κ^+-homogeneity we can find, for every $\alpha < \kappa$, an automorphism $\sigma_\alpha \in \mathrm{Aut}\,\mathfrak{M}_U$ such that

$$\sigma_\alpha(\bar b'_\beta) = \bar b'_{\alpha+\beta}\,, \quad \text{for all } \beta < \kappa\,.$$

Let $\pi_\alpha \in \mathrm{Aut}\,\mathbb{M}$ be an extension of σ_α and set

$$\bar a_\alpha := \pi_\alpha(\bar a')\,, \quad \text{for } \alpha < \kappa\,.$$

Then

$$\begin{aligned} \mathbb{M} \vDash \varphi(\bar a_\alpha; \bar b'_{\alpha+\beta}) \quad &\text{iff} \quad \mathbb{M} \vDash \varphi\big(\pi_\alpha(\bar a'); \pi_\alpha(\bar b'_\beta)\big) \\ &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar a'; \bar b'_\beta) \\ &\text{iff} \quad \beta = 0\,. \end{aligned}$$

Consequently, $\mathrm{tp}(\bar{a}_\alpha/M) \neq \mathrm{tp}(\bar{a}_\beta/M)$, for $\alpha \neq \beta$. Since these types are extensions of $\mathrm{tp}(\bar{a}/U)$ and they are conjugate over U, the claim follows. □

Theorem 4.5. *Let T be a complete first-order theory.*

(a) *T is stable if, and only if, there exists a relation $\leq$ on complete types satisfying* (INV), (LC), *and* (BND).

(b) *If $\leq$ is an extension relation satisfying* (INV), (LC), (BND), (EXT), (TR), *and* (MON), *then $\leq\, = \,\leq_{\downarrow\mathrm{f}}$.*

Proof. (a) ($\Rightarrow$) If T is stable, the relation $\leq_{\downarrow\mathrm{f}}$ has the desired properties by Proposition 4.3. For ($\Leftarrow$), suppose that $\leq$ satisfies (INV), (LC), and (BND). Let κ be the cardinal from (LC) and fix a κ-saturated model $\mathfrak{M}$. For $U \subseteq M$ and $\mathfrak{p} \in S^{<\omega}(U)$, we denote by $\mu(\mathfrak{p}; U)$ the cardinal from (BND). Set

$$\mu := \sup \left\{ \mu(\mathfrak{p}; U) \;\middle|\; U \subseteq M \text{ with } |U| < \kappa \text{ and } \mathfrak{p} \in S^{<\omega}(U) \right\}.$$

Since, for every subset $U \subseteq \mathbb{M}$ of size $|U| < \kappa$, there is some automorphism $\pi \in \mathrm{Aut}\,\mathbb{M}$ with $\pi[U] \subseteq M$, it follows by (INV) that, for all sets $U, C \subseteq \mathbb{M}$ with $|U| < \kappa$ and for every type $\mathfrak{p} \in S^{<\omega}(U)$, there are at most μ types $\mathfrak{q} \in S^{<\omega}(U \cup C)$ with $\mathfrak{p} \leq \mathfrak{q}$.

Fix a set $U \subseteq \mathbb{M}$ and a finite tuple $\bar{s}$ of sorts. For every type $\mathfrak{p} \in S^{\bar{s}}(U)$, we can fix, by (LC), a subset $C(\mathfrak{p}) \subseteq U$ of size $|C(\mathfrak{p})| < \kappa$ such that $\mathfrak{p}|_{C(\mathfrak{p})} \leq \mathfrak{p}$. Let $C \subseteq U$ be a set of size $|C| < \kappa$. Then

- $|S^{\bar{s}}(C)| \leq 2^{|T| \oplus \kappa}$ and,
- for every type $\mathfrak{q} \in S^{\bar{s}}(C)$, there are at most μ types $\mathfrak{p} \in S^{\bar{s}}(U)$ with $C(\mathfrak{p}) = C$ and $\mathfrak{p}|_C = \mathfrak{q}$.

Consequently, we have

$$\left|S^{\bar{s}}(U)\right| \leq |U|^{<\kappa} \otimes 2^{|T| \oplus \kappa} \otimes \mu .$$

Setting $\lambda_0 := \mu \oplus 2^{|T|}$ and $\lambda := \lambda_0^\kappa$, it follows that

$$\left|S^{\bar{s}}(U)\right| \leq \lambda^\kappa \otimes \lambda_0^\kappa \otimes \mu = \lambda , \quad \text{for every set } U \text{ of size } |U| \leq \lambda .$$

Hence, T is λ-stable.

(b) Let $\mathfrak{q} \in S^{\bar{s}}(B)$ be an extension of $\mathfrak{p} \in S^{\bar{s}}(A)$. We have to show that $\mathfrak{p} \leq \mathfrak{q}$ if, and only if, $\mathfrak{q}$ is a $\downfree^{\mathrm{f}}$-free extension of $\mathfrak{p}$.

($\Rightarrow$) Suppose that $\mathfrak{p} \leq \mathfrak{q}$. By (BND), there is a cardinal μ such that, for every set C, $\mathfrak{p}$ has at most μ $\leq$-free extensions over $A \cup C$. Set $\kappa := \mu^+ \oplus |A|^+$ and let $\mathfrak{M}$ be a κ^+-saturated and strongly κ^+-homogeneous model containing B. We use (EXT) to find a type $\mathfrak{r} \geq \mathfrak{q}$ over M. By (TR), it follows that $\mathfrak{p} \leq \mathfrak{r}$. Hence, (INV) implies that $\mathfrak{r}$ has at most μ conjugates over A. By Lemma 4.4, it follows that $\mathfrak{r}$ does not fork over A. In particular, $\mathfrak{q}$ does not fork over A.

($\Leftarrow$) Suppose that $\mathfrak{q}$ is a free extension of $\mathfrak{p}$. Fix a strongly $(|T| \oplus |A|)^+$-homogeneous model $\mathfrak{M}$ containing B and let $\mathfrak{r}$ be a free extension of $\mathfrak{q}$ over M. By (EXT), there exists a type $\mathfrak{r}' \geq \mathfrak{p}$ over M. By the first part of the proof, $\mathfrak{p} \leq \mathfrak{r}'$ implies that $\mathfrak{r}'$ is a free extension of $\mathfrak{p}$. Let $\bar{a}$ and $\bar{a}'$ be realisations of, respectively, $\mathfrak{r}$ and $\mathfrak{r}'$. Then we can use Lemma 3.4 to find an automorphism $\pi \in \operatorname{Aut} \mathbb{M}_A$ such that $\pi(\bar{a}') = \bar{a}$ and $\pi[M] = M$. By (INV),

$$\mathfrak{p} \leq \mathfrak{q} \quad \text{implies} \quad \mathfrak{p} \leq \operatorname{tp}(\pi(\bar{a}')/\pi[M]) = \operatorname{tp}(\bar{a}/M)\,.$$

Hence, it follows by (MON) that $\mathfrak{p} \leq \operatorname{tp}(\bar{a}/B) = \mathfrak{q}$. □

Translating this theorem into the language of forking relations, we obtain the following characterisation of stable theories.

Theorem 4.6. *Let T be a complete first-order theory. The following statements are equivalent.*

(1) *T is stable.*

(2) *$\sqrt[\mathrm{f}]{}$ is symmetric and* $\operatorname{mult}(\sqrt[\mathrm{f}]{}) < \infty$.

(3) *There exists a symmetric forking relation $\downfree$ with* $\operatorname{mult}(\downfree) < \infty$.

Proof. (1) $\Rightarrow$ (2) follows by Lemma 4.2 and the implication (2) $\Rightarrow$ (3) is trivial. For (3) $\Rightarrow$ (1), suppose that $\downfree$ is a symmetric forking relation with $\operatorname{mult}(\downfree) < \infty$. By Proposition 4.3, the corresponding extension relation $\leq_{\downarrow}$ satisfies the three axioms from Theorem 4.5 (a). Hence, T is stable. □

Proposition 4.7. *If* $\downarrow$ *is a symmetric forking relation with* $\mathrm{mult}(\downarrow) < \infty$, *then* $\downarrow = \downarrow^{\mathrm{f}}$.

Proof. By Proposition 4.3, the extension relation $\leq_{\downarrow}$ satisfies the six axioms from Theorem 4.5 (b). Consequently, $\leq_{\downarrow} = \leq_{\downarrow^{\mathrm{f}}}$. For a finite tuple $\bar{a}$ and sets $U, B \subseteq \mathbb{M}$, it follows that

$$\begin{aligned} \bar{a} \downarrow_U B \quad &\text{iff} \quad \mathrm{tp}(\bar{a}/U) \leq_{\downarrow} \mathrm{tp}(\bar{a}/B) \\ &\text{iff} \quad \mathrm{tp}(\bar{a}/U) \leq_{\downarrow^{\mathrm{f}}} \mathrm{tp}(\bar{a}/B) \quad \text{iff} \quad \bar{a} \downarrow^{\mathrm{f}}_U B \,. \end{aligned}$$

Hence, finite character implies that $\downarrow = \downarrow^{\mathrm{f}}$. □

As a further application we derive a characterisation of forking in totally transcendental theories by showing that the relation of being a Morley-free extension (which was defined in Section F2.1) satisfies the conditions of the above theorem.

Corollary 4.8. *Let* T *be a totally transcendental theory.*

(a) $\bar{a} \downarrow^{\mathrm{f}}_U B \quad$ iff $\quad \mathrm{rk}_{\mathrm{M}}(\bar{a}/UB) = \mathrm{rk}_{\mathrm{M}}(\bar{a}/U)\,, \quad$ *for all finite* $\bar{a}$.

(b) $\mathrm{mult}_{\downarrow^{\mathrm{f}}}(\mathfrak{p}) < \aleph_0$.

Proof. (a) For types $\mathfrak{p} \in S^{\bar{s}}(U)$ and $\mathfrak{q} \in S^{\bar{s}}(V)$, we define

$$\mathfrak{p} \leq_{\mathrm{M}} \mathfrak{q} \quad :\text{iff} \quad \mathfrak{q} \text{ is a Morley-free extension of } \mathfrak{p}\,.$$

It is sufficient to show that $\leq_{\mathrm{M}}$ satisfies the conditions in Theorem 4.5 (b).

(INV) follows immediately from the definition. (BND) and (EXT) were already shown in Lemma F2.1.9 (a) and (b), respectively, while (LC) was proved in Lemma F2.1.6 (c).

For (TR), suppose that $\mathfrak{p} \leq_{\mathrm{M}} \mathfrak{q} \leq_{\mathrm{M}} \mathfrak{r}$. Then

$$\mathrm{rk}_{\mathrm{M}}(\mathfrak{p}) = \mathrm{rk}_{\mathrm{M}}(\mathfrak{q}) = \mathrm{rk}_{\mathrm{M}}(\mathfrak{r})\,,$$

which implies that $\mathfrak{p} \leq_M \mathfrak{r}$.

(MON) Suppose that $\mathfrak{p} \leq_M \mathfrak{r}$ and $\mathfrak{p} \subseteq \mathfrak{q} \subseteq \mathfrak{r}$. By Lemma F2.1.6 (b), we have

$$\mathrm{rk}_M(\mathfrak{p}) \geq \mathrm{rk}_M(\mathfrak{q}) \geq \mathrm{rk}_M(\mathfrak{r}) = \mathrm{rk}_M(\mathfrak{p}) .$$

Hence, $\mathfrak{p} \leq_M \mathfrak{q}$.

(b) We have seen in Lemma F2.1.9 that every type has only finitely many Morley-free extensions. Hence, the claim follows by (a). □

5. Morley sequences in stable theories

Let us collect several results on Morley sequences in stable theories. Many of the proofs rely on the notion of a stationary type. We start with a proof that we can drop the requirement of indiscernibility from the definition of a Morley sequence if the type in question is stationary.

Lemma 5.1. *Let T be a stable theory, $\mathfrak{p}$ a stationary type over U, and let $(\bar{a}_i)_{i \in I}$ be a sequence of realisations of $\mathfrak{p}$. If*

$$\bar{a}_i \mathop{\downarrow}^{\mathrm{f}}_U \bar{a}[<i] , \quad \text{for all } i \in I ,$$

then $(\bar{a}_i)_{i \in I}$ is a $\mathop{\downarrow}^{\mathrm{f}}$-Morley sequence for $\mathfrak{p}$ over U.

Proof. We have to show that $(\bar{a}_i)_{i \in I}$ is indiscernible over U. By induction on $n < \omega$, we prove that

$$\bar{a}[\bar{\imath}] \equiv_{U\bar{a}[<l]\bar{a}[>m]} \bar{a}[\bar{k}] , \quad \text{for all } \bar{\imath}, \bar{k} \in [I]^n \text{ with } l < \bar{\imath}, \bar{k} < m .$$

Hence, let $\bar{\imath}, \bar{k} \in [I]^n$ and $l < \bar{\imath}, \bar{k} < m$. By symmetry, we may assume that $i_{n-1} < k_{n-1}$. According to Lemma F2.4.9, we have

$$\bar{a}_{i_{n-1}} \mathop{\downarrow}^{\mathrm{f}}_U \bar{a}[<i_{n-1}]\bar{a}[>m] \quad \text{and} \quad \bar{a}_{k_{n-1}} \mathop{\downarrow}^{\mathrm{f}}_U \bar{a}[<i_{n-1}]\bar{a}[>m] ,$$

Since $\mathfrak{p}$ is stationary, it follows that

$$\bar{a}_{i_{n-1}} \equiv_{U\bar{a}[<i_{n-1}]\bar{a}[>m]} \bar{a}_{k_{n-1}} .$$

By inductive hypothesis, we further have

$$\bar{a}_{i_{n-2}} \dots \bar{a}_{i_0} \equiv_{U\bar{a}[<l]\bar{a}[\geq k_{n-1}]} \bar{a}_{k_{n-2}} \dots \bar{a}_{k_0} .$$

Consequently,

$$\begin{aligned} \bar{a}_{i_{n-1}} \bar{a}_{i_{n-2}} \dots \bar{a}_{i_0} &\equiv_{U\bar{a}[<l]\bar{a}[>m]} \bar{a}_{k_{n-1}} \bar{a}_{i_{n-2}} \dots \bar{a}_{i_0} \\ &\equiv_{U\bar{a}[<l]\bar{a}[>m]} \bar{a}_{k_{n-1}} \bar{a}_{k_{n-2}} \dots \bar{a}_{k_0} . \end{aligned}$$

□

Lemma 5.2. *Let T be a stable theory, $\mathfrak{p}$ a stationary type over U and $\mathfrak{q}$ the unique free extensions of $\mathfrak{p}$ over $U \cup C$.*

(a) *Every $\mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}$-Morley sequences $(\bar{a}_i)_{i\in I}$ for $\mathfrak{q}$ over $U \cup C$ is also a $\mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}$-Morley sequence for $\mathfrak{p}$ over U.*

(b) *Let $(\bar{a}_i)_{i\in I}$ be a $\mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}$-Morley sequence for $\mathfrak{p}$ over U. If $I_0 \subseteq I$ and $C \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}_{U\bar{a}[I_0]} \bar{a}[I]$, then $(\bar{a}_i)_{i\in I \setminus I_0}$ is a Morley sequence for $\mathfrak{q}$ over $U \cup C$.*

Proof. (a) As $(\bar{a}_i)_{i\in I}$ is indiscernible over $U \cup C$, it is trivially indiscernible over U. Furthermore,

$$\bar{a}_i \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}_U C \quad \text{and} \quad \bar{a}_i \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}_{UC} \bar{a}[<i] \quad \text{implies} \quad \bar{a}_i \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}_U \bar{a}[<i] .$$

(b) Set $A_0 := \bar{a}[I_0]$. We start by showing that

$$C\bar{a}[K] \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}_{UA_0} \bar{a}[I \setminus K] , \quad \text{for all finite } K \subseteq I \setminus I_0 .$$

The proof is by induction on $|K|$. If $K = \varnothing$, the claim holds by assumption. Hence, suppose that $K = K_0 \cup \{k\}$ and that we have already shown that

$$C\bar{a}[K_0] \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}_{UA_0} \bar{a}[I \setminus K_0] .$$

Then

$$C\bar{a}[K_0] \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}_{UA_0\bar{a}_k} \bar{a}[I \setminus K] .$$

Furthermore, we have seen in Lemma F2.4.9 that

$$\bar{a}_k \mathop{\downfree}^{\mathrm{f}}_U \bar{a}[I \setminus K]\,,$$

which implies that

$$\bar{a}_k \mathop{\downfree}^{\mathrm{f}}_{UA_0} \bar{a}[I \setminus K]\,.$$

Consequently, it follows by transitivity that

$$C\bar{a}[K_0]\bar{a}_k \mathop{\downfree}^{\mathrm{f}}_{UA_0} \bar{a}[I \setminus K]\,.$$

Having proved the claim, it follows by (FIN) that

$$C\bar{a}[<i] \mathop{\downfree}^{\mathrm{f}}_{UA_0} \bar{a}_i\,, \quad \text{for all } i \in I\,,$$

which implies that

$$\bar{a}_i \mathop{\downfree}^{\mathrm{f}}_{UCA_0} \bar{a}[<i]\,, \quad \text{for all } i \in I \setminus I_0\,.$$

Hence, it follows by Lemma 5.1 that $(\bar{a}_i)_{i \in I \setminus I_0}$ is a $\downfree^{\mathrm{f}}$-Morley sequences over $U \cup C \cup A_0$. ☐

For stable theories, we can turn every indiscernible sequence into a Morley sequence by increasing the domain of the type.

Proposition 5.3. *Let T be a stable theory, κ an infinite cardinal, and let $(\bar{a}_i)_{i \in I}$ be an infinite indiscernible sequence over U such that $|\bar{a}_i| < \kappa$, for all $i \in I$. There exist a set C of size $|C| < \kappa \oplus \aleph_1$ and a stationary type $\mathfrak{p} \in S^{<\kappa}(U \cup C)$ such that $(\bar{a}_i)_{i \in I}$ is a $\downfree^{\mathrm{f}}$-Morley sequence for $\mathfrak{p}$ over $U \cup C$.*

Proof. We have seen in Corollary E5.4.13 and Corollary E5.4.14 that, for every formula $\varphi(\bar{x}; \bar{c})$ the set

$$[\![\varphi(\bar{a}_i; \bar{c})]\!]_{i \in I} = \{\, i \in I \mid \mathbb{M} \vDash \varphi(\bar{a}_i; \bar{c}) \,\}$$

is either finite or cofinite and that, for every set $C \subseteq \mathbb{M}$, the type

$$\mathrm{Av}_1((\bar{a}_i)_i/UC) = \{\, \varphi(\bar{x}) \mid \varphi \text{ a formula over } U \cup C \text{ such that } [\![\varphi(\bar{a}_i)]\!]_{i\in I} \text{ is cofinite} \,\}$$

is complete. According to Proposition E5.4.12, there exists, for every formula $\varphi(\bar{x};\bar{y})$ over U, a finite constant $k(\varphi) < \omega$ such that, for all $\bar{c} \subseteq \mathbb{M}$,

$$\varphi(\bar{x};\bar{c}) \in \mathrm{Av}_1((\bar{a}_i)_i/UC) \quad \text{iff} \quad \big|I \smallsetminus [\![\varphi(\bar{a}_i;\bar{c})]\!]_{i\in I}\big| \leq k(\varphi)\,.$$

Choose an injective function $\mu : \omega \to I$ and set $I_0 := \operatorname{rng}\mu$. It follows that

$$\begin{aligned} &\varphi(\bar{x};\bar{c}) \in \mathrm{Av}_1((\bar{a}_i)_i/UC) \\ \text{iff} \quad &\big|[\![\varphi(\bar{a}_{\mu(n)};\bar{c})]\!]_{n<2k(\varphi)+1}\big| > k(\varphi) \\ \text{iff} \quad &\mathbb{M} \vDash \bigvee\big\{ \bigwedge_{n\in K} \varphi(\bar{a}_{\mu(n)};\bar{c}) \,\big|\, K \subseteq [2k(\varphi)+1],\ |K| = k(\varphi)+1 \big\}\,. \end{aligned}$$

Consequently, the type $\mathrm{Av}_1((\bar{a}_i)_i/UC)$ is definable over $\bar{a}[I_0]$, for every $C \subseteq \mathbb{M}$. For $C \subseteq \mathbb{M}$, fix a tuple $\bar{a}(C)$ realising $\mathrm{Av}_1((\bar{a}_i)_i/UC\bar{a}[I_0])$. For every C, we have

$$\bar{a}(C) \equiv_{U\bar{a}[I_0]} \bar{a}(\varnothing) \quad \text{and} \quad \bar{a}(C) \overset{\mathrm{df}}{\diagup\!\!\!\!\smile}_{\bar{a}[I_0]} UC\,.$$

Consequently,

$$\bar{a}(\varnothing)\ {}^*\big(\overset{\mathrm{df}}{\diagup\!\!\!\!\smile}\big)_{\bar{a}[I_0]}\ U$$

By Propositions 3.7 and 2.5 it follows that the type

$$\mathfrak{p} := \mathrm{tp}(\bar{a}(\varnothing)/U\bar{a}[I_0]) = \mathrm{Av}_1((\bar{a}_i)_i/U\bar{a}[I_0])$$

is stationary. Since the tuple $\bar{a}_i$ realises $\mathrm{Av}_1((\bar{a}_i)_i/U\bar{a}[I_0]\bar{a}[<i])$, for $i \in I \smallsetminus I_0$,

$$\bar{a}_i\ {}^*\big(\overset{\mathrm{df}}{\diagup\!\!\!\!\smile}\big)_{\bar{a}[I_0]}\ U\bar{a}[I_0]\bar{a}[<i] \quad \text{implies} \quad \bar{a}_i \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{f}}_{\bar{a}[I_0]} U\bar{a}[I_0]\bar{a}[<i]\,,$$

by Propositions 1.5 and F3.1.12. Therefore, it follows by Lemma 5.1 that $(\bar{a}_i)_{i\in I\smallsetminus I_0}$ is a $\mathop{\smile\!\!\!\!\!\!|}^{\mathrm{f}}$-Morley sequence for $\mathfrak{p}$ over $U \cup \bar{a}[I_0]$. ☐

As an application, it follows that every indiscernible sequence can be turned into an indiscernible sequence over a larger set if we remove some of its elements.

Lemma 5.4. *Let T be a stable theory and $(\bar{a}_i)_{i \in I}$ and indiscernible sequence over U with $|\bar{a}_i| < \aleph_0$. For every set $C \subseteq \mathbb{M}$ there exists a set $I_0 \subseteq I$ of size $|I_0| \leq |C| \otimes \mathrm{loc}_0(\downarrow^{\mathrm{f}})$ such that $(\bar{a}_i)_{i \in I \setminus I_0}$ is indiscernible over $U \cup C \cup \bar{a}[I_0]$.*

Proof. We have seen in Proposition 5.3 that there exist a set $U' \supseteq U$ and a stationary type $\mathfrak{p}$ over U' such that $(\bar{a}_i)_{i \in I}$ is a Morley sequence for $\mathfrak{p}$ over U'. For every finite $C_0 \subseteq C$, we can find a set $J(C_0) \subseteq I$ of size $|J(C_0)| < \mathrm{loc}_0(\downarrow^{\mathrm{f}})$ such that

$$C_0 \downarrow^{\mathrm{f}}_{U' \bar{a}[J(C_0)]} \bar{a}[I] .$$

Setting $I_0 := \bigcup \{ J(C_0) \mid C_0 \subseteq C \text{ finite} \}$ it follows that

$$|I_0| \leq |C| \otimes \mathrm{loc}_0(\downarrow^{\mathrm{f}}) \quad \text{and} \quad C \downarrow^{\mathrm{f}}_{U' \bar{a}[I_0]} \bar{a}[I] .$$

Consequently, we can use Lemma 5.2 to show that $(\bar{a}_i)_{i \in I \setminus I_0}$ is a Morley sequence over $U' C \bar{a}[I_0]$. In particular, it is indiscernible over $U C \bar{a}[I_0]$. □

In totally transcendental theories, it is particularly simple to find Morley sequences.

Definition 5.5. Let $\surd$ be an abstract independence relation. A family $(A_i)_{i \in I}$ of sets is $\surd$-*independent* over U if

$$A_k \surd_U \bigcup_{i \neq k} A_i , \quad \text{for all } k \in I .$$

Lemma 5.6. *Let T be a totally transcendental theory and $\mathfrak{p} \in S^{<\omega}(U)$ a type. Every set $I \subseteq \mathfrak{p}^{\mathbb{M}}$ that is $\downarrow^{\mathrm{f}}$-independent over U has a finite partition $I = I_0 \cup \dots \cup I_{n-1}$ such that each I_i is totally indiscernible over U.*

Proof. We have seen in Corollary 4.8 (b) that $\mathrm{mult}_{\downarrow^{\mathrm f}}(\mathfrak p) < \aleph_0$. Thus, $\mathfrak p$ has only finitely many free extensions $\mathfrak q_0, \dots, \mathfrak q_{n-1}$ over $\mathrm{acl}^{\mathrm{eq}}(U)$. By Lemma 5.1, each set $I_i := I \cap \mathfrak q_i^{\mathbb M}$ forms a Morley sequence over U. In particular, it is totally indiscernible. $\square$

6. The stability spectrum

The *stability spectrum* of a theory T is the class of all cardinals κ such that T is κ-stable. In this section, we will compute the stability spectrum from two parameters: $\mathrm{fc}(\downarrow^{\mathrm f})$ and $\mathrm{st}(T)$. Recall that $\mathrm{fc}(\downarrow^{\mathrm f})$ is the least cardinal κ such that there is no $\downarrow^{\mathrm f}$-forking chain of length κ for a finite set. The cardinal $\mathrm{st}(T)$ is defined as follows.

Definition 6.1. Let T be a complete theory. $\mathrm{st}(T)$ is the minimal infinite cardinal κ such that T is κ-stable. If there is no such cardinal, we set $\mathrm{st}(T) := \infty$.

The following technical lemma contains the main ingredients to determine the stability spectrum of a theory.

Lemma 6.2. *Let T be a stable theory and κ a cardinal.*

(a) *If $\kappa < \kappa^{<\mathrm{fc}(\downarrow^{\mathrm f})}$, then T is not κ-stable.*

(b) $\mathrm{fc}(\downarrow^{\mathrm f}) \leq |T|^+$.

(c) $|S^{<\omega}(U)| \leq \mathrm{st}(T) \leq 2^{|T|}$, *for every set U of size $|U| \leq \mathrm{st}(T)$.*

(d) $\mathrm{fc}(\downarrow^{\mathrm f}) \oplus \mathrm{mult}(\downarrow^{\mathrm f}) \leq \mathrm{st}(T)$.

(e) *If $\kappa \geq \mathrm{st}(T)$ and $\kappa = \kappa^{<\mathrm{fc}(\downarrow^{\mathrm f})}$, then T is κ-stable.*

Proof. (a) Let μ be the least cardinal with $\kappa^\mu > \kappa$. Then $\kappa < \kappa^{<\mathrm{fc}(\downarrow^{\mathrm f})}$ implies that $\mu < \mathrm{fc}(\downarrow^{\mathrm f})$. Hence, there exist a finite tuple $\bar a$ and a $\downarrow^{\mathrm f}$-forking chain $(\bar b_\alpha)_{\alpha<\mu}$ for $\bar a$ over $\varnothing$ of length μ. We construct a tree $(\bar c_\eta)_{\eta \in \kappa^{\leq\mu}}$ as follows. We start with $\bar c_{\langle\rangle} := \bar b_0$. For the inductive step, suppose that $\bar c_\eta$ is already defined for all $\eta \in \kappa^{<\mu}$ with $|\eta| < \alpha$ and set

$$C_\alpha := \bigcup \{\, \bar c_\eta \mid \eta \in \kappa^{<\alpha} \,\} .$$

If α is a limit ordinal, we choose, for every $\eta \in \kappa^\alpha$, a tuple $\bar{c}_\eta$ with

$$\bar{c}_\eta \bar{c}[<\eta] \equiv \bar{b}_\alpha \bar{b}[<\alpha] \quad \text{and} \quad \bar{c}_\eta \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}_{\bar{c}[<\eta]} C_\alpha \,.$$

For the successor step, suppose that $\alpha = \beta + 1$. For each $\eta \in \kappa^\beta$, we choose

- a tuple $\bar{d}$ such that $\bar{d}\bar{c}[\leq\eta] \equiv \bar{b}_\alpha \bar{b}[<\alpha]$,
- a $\mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}$-Morley sequence $(\bar{c}'_i)_{i<\kappa}$ for $\mathrm{tp}(\bar{d}/\bar{c}[\leq\eta])$ over $\bar{c}[\leq\eta]$, and
- a sequence $(\bar{c}''_i)_{i<\kappa}$ such that

$$\bar{c}''[<\kappa] \equiv_{\bar{c}[\leq\eta]} \bar{c}'[<\kappa] \quad \text{and} \quad \bar{c}''[<\kappa] \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}_{\bar{c}[\leq\eta]} C_\alpha \,.$$

Then we set $\bar{c}_{\eta\alpha} := \bar{c}''_\alpha$, for $\alpha < \kappa$.

Having constructed the tree $(\bar{c}_\eta)_{\eta\in\kappa^{<\mu}}$, we set $U := \bigcup_{\eta\in\kappa^{<\mu}} \bar{c}_\eta$. Then $|U| = \kappa^{<\mu} = \kappa$. For each $\zeta \in \kappa^\mu$, let $\bar{a}_\zeta$ be a tuple such that

$$\bar{a}_\zeta \bar{c}[<\zeta] \equiv \bar{a}\bar{b}[<\mu] \quad \text{and} \quad \bar{a}_\zeta \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}_{\bar{c}[<\zeta]} U \,.$$

We claim that

$$\bar{a}_\xi \not\equiv_U \bar{a}_\zeta \,, \qquad \text{for } \xi \neq \zeta \,.$$

This implies that $|S^{<\omega}(U)| \geq \kappa^\mu > \kappa = |U|$. Hence, T is not κ-stable.

It remains to prove the claim. Given $\xi \neq \zeta$, let η be the longest common prefix of ξ and ζ and let $\alpha \neq \beta$ be the indices such that $\eta\alpha \preceq \xi$ and $\eta\beta \preceq \zeta$. We start by showing that

$$\bar{c}[\leq\zeta_0] \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}_{\bar{c}[\leq\eta]} \bar{c}_{\eta\alpha} \,, \qquad \text{for all } \zeta_0 \preceq \zeta \,.$$

The proof is by induction on $|\zeta_0|$. Note that we have

$$\bar{c}_{\eta\beta} \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}_{\bar{c}[\leq\eta]} \bar{c}_{\eta\alpha}$$

by choice of $\bar{c}_{\eta\beta}$ and $\bar{c}_{\eta\alpha}$. By (NOR) this implies that

$$\bar{c}[\leq\eta\beta] \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}_{\bar{c}[\leq\eta]} \bar{c}_{\eta\alpha} \,.$$

Hence, the claim holds for $\zeta_\circ \leq \eta\beta$. For the inductive step, let $\zeta_\circ > \eta\beta$. Then

$$\bar{c}_{\zeta_\circ} \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}_{\bar{c}[<\zeta_\circ]} C_{|\zeta_\circ|} \quad \text{implies} \quad \bar{c}_{\zeta_\circ} \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}_{\bar{c}[<\zeta_\circ]} \bar{c}_{\eta\alpha} \,.$$

By inductive hypothesis, we furthermore have

$$\bar{c}[<\zeta_\circ] \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}_{\bar{c}[\leq\eta]} \bar{c}_{\eta\alpha} \,.$$

Hence, the claim follows by transitivity.

Having proved the claim, it follows by finite character that

$$\bar{c}[<\zeta] \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}_{\bar{c}[\leq\eta]} \bar{c}_{\eta\alpha} \,.$$

Since $\bar{a}_\zeta \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}_{\bar{c}[<\zeta]} \bar{c}_{\eta\alpha}$ this implies by transitivity that

$$\bar{a}_\zeta \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}_{\bar{c}[\leq\eta]} \bar{c}_{\eta\alpha} \,.$$

On the other hand,

$$\bar{a} \not\mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}_{\bar{b}[\leq\alpha]} \bar{b}_{\alpha+1} \quad \text{implies} \quad \bar{a}_\xi \not\mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}_{\bar{c}[\leq\eta]} \bar{c}_{\eta\alpha} \,.$$

Consequently, $\bar{a}_\xi \not\equiv_{\bar{c}[\leq\eta\alpha]} \bar{a}_\zeta$.

(b) follows from Theorem F3.2.18.

(c) For the upper bound, it is sufficient to note that, according to Theorem C3.5.17, $(2^{|T|})^{|T|} = 2^{|T|}$ implies that T is $2^{|T|}$-stable.

For the lower bound, let U be a set of size $|U| \leq \mathrm{st}(T)$. Fixing some set $A \supseteq U$ of size $|A| = \mathrm{st}(T)$, it follows by $\mathrm{st}(T)$-stability of T that

$$|S^{<\omega}(U)| \leq |S^{<\omega}(A)| = |A| = \mathrm{st}(T) \,.$$

(d) We start by showing that $\mathrm{fc}(\mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}) \leq \mathrm{st}(T)$. Note that

$$\kappa < \mathrm{fc}(\mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}) \quad \text{implies} \quad \kappa < \kappa^\kappa \leq \kappa^{<\mathrm{fc}(\mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}})}.$$

Therefore, it follows by (a) that T is not κ-stable for $\kappa < \mathrm{fc}(\mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}})$. Thus, $\mathrm{st}(T) \geq \mathrm{fc}(\mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}})$.

It therefore remains to prove that $\mathrm{mult}(\mathop{\downarrow}^{\mathrm{f}}) \leq \mathrm{st}(T)$. Let $\mathfrak{p} \in S^{<\omega}(U)$. By Proposition F2.3.24, there exists a set $U_\circ \subseteq U$ of size $|U_\circ| < \mathrm{loc}_\circ(\mathop{\downarrow}^{\mathrm{f}}) \leq \mathrm{fc}(\mathop{\downarrow}^{\mathrm{f}})$ such that $\mathfrak{p}|_{U_\circ} \leq \mathfrak{p}$. Since every free extension of $\mathfrak{p}$ is also a free extension of $\mathfrak{p}|_{U_\circ}$, it follows that $\mathrm{mult}(\mathfrak{p}|_{U_\circ}) \geq \mathrm{mult}(\mathfrak{p})$. As T is $\mathrm{st}(T)$-stable,

$$|U_\circ| < \mathrm{fc}(\mathop{\downarrow}^{\mathrm{f}}) \leq \mathrm{st}(T) \quad \text{implies} \quad |S^{\leq\omega}(U_\circ)| \leq \mathrm{st}(T)\,.$$

Consequently, there exists a model $\mathfrak{M}$ of size $\mathrm{st}(T)$ that contains $U_\circ$. We claim that $|S^{<\omega}(M)| \geq \mathrm{mult}(\mathfrak{p}|_{U_\circ})$. As T is $\mathrm{st}(T)$-stable, it then follows that

$$\mathrm{mult}(\mathfrak{p}) \leq \mathrm{mult}(\mathfrak{p}|_{U_\circ}) \leq |S^{<\omega}(M)| \leq \mathrm{st}(T)\,.$$

To prove the claim, consider a set $C \supseteq U_\circ$ and let $(\mathfrak{q}_\alpha)_{\alpha<\lambda}$ be a sequence of distinct free extensions of $\mathfrak{p}|_{U_\circ}$ over C. For each $\alpha < \lambda$, choose a free extension $\mathfrak{q}^+_\alpha \geq \mathfrak{q}_\alpha$ over $C \cup M$ and set $\mathfrak{r}_\alpha := \mathfrak{q}^+_\alpha|_M$. If we can show that $\mathfrak{r}_\alpha \neq \mathfrak{r}_\beta$, for $\alpha \neq \beta$, it will follow that $|S^{<\omega}(M)| \geq \lambda$, as desired.

Hence, suppose that $\mathfrak{r}_\alpha = \mathfrak{r}_\beta$. Since $\mathfrak{q}^+_\alpha$ and $\mathfrak{q}^+_\beta$ are free extensions of the stationary type $\mathfrak{r}_\alpha = \mathfrak{r}_\beta$, it follows that $\mathfrak{q}^+_\alpha = \mathfrak{q}^+_\beta$. In particular, $\mathfrak{q}_\alpha = \mathfrak{q}_\beta$, which implies that $\alpha = \beta$.

(e) Let $\kappa \geq \mathrm{st}(T)$ be a cardinal such that $\kappa^{<\mathrm{fc}(\mathop{\downarrow}^{\mathrm{f}})} = \kappa$ and let U be a set of size $|U| \leq \kappa$. By Proposition F2.3.24, we can find, for every type $\mathfrak{p} \in S^{<\omega}(U)$, a set $U_\circ \subseteq U$ of size $|U_\circ| < \mathrm{loc}_\circ(\mathop{\downarrow}^{\mathrm{f}}) \leq \mathrm{fc}(\mathop{\downarrow}^{\mathrm{f}})$ such that $\mathfrak{p}|_{U_\circ} \leq \mathfrak{p}$. As T is $\mathrm{st}(T)$-stable,

$$|U_\circ| < \mathrm{fc}(\mathop{\downarrow}^{\mathrm{f}}) \leq \mathrm{st}(T) \quad \text{implies} \quad |S^{\leq\omega}(U_\circ)| \leq \mathrm{st}(T)\,.$$

Consequently, it follows as in Theorem 4.5 that

$$\begin{aligned}\left|S^{<\omega}(U)\right| &\leq |U|^{<\mathrm{fc}(\mathop{\downarrow}^{\mathrm{f}})} \otimes \mathrm{st}(T) \otimes \mathrm{mult}(\mathop{\downarrow}^{\mathrm{f}}) \\ &\leq \kappa^{<\mathrm{fc}(\mathop{\downarrow}^{\mathrm{f}})} \otimes \mathrm{st}(T) \otimes \mathrm{mult}(\mathop{\downarrow}^{\mathrm{f}}) = \kappa\,,\end{aligned}$$

where the last equality follows by (c) and our choice of κ. $\square$

Combining statements (a) and (d) of Lemma 6.2 we obtain the following description of the stability spectrum.

Theorem 6.3. *A stable theory T is κ-stable if, and only if,*

$$\kappa = \mathrm{st}(T) \oplus \kappa^{<\mathrm{fc}(\mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}})}.$$

Proof. ($\Leftarrow$) follows by Lemma 6.2 (d). For ($\Rightarrow$), suppose that T is κ-stable. By definition, this implies that $\kappa \geq \mathrm{st}(T)$. Furthermore, it follows by Lemma 6.2 (a) that $\kappa \geq \kappa^{<\mathrm{fc}(\mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}})}$. Since the converse inequality $\kappa \leq \kappa^{<\mathrm{fc}(\mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}})}$ is trivial, the claim follows. ☐

Let us consider a subclass of stable theories where the stability spectrum is particularly simple.

Definition 6.4. A complete first-order theory T is called *supersimple* if $\mathrm{loc}(\overset{\mathrm{f}}{\surd}) \leq \aleph_0$. If T is supersimple and stable, we call it *superstable.*

Note that it follows by Theorem F2.4.17 that every supersimple theory is simple.

Theorem 6.5. *Let T be a complete first-order theory. The following conditions are equivalent.*

(1) *T is supersimple.*

(2) $\mathrm{fc}(\overset{\mathrm{f}}{\surd}) \leq \aleph_0$

(3) $\mathrm{loc}_0(\overset{\mathrm{f}}{\surd}) \leq \aleph_0$

Proof. (2) $\Leftrightarrow$ (3) follows by Proposition F2.3.24 and (1) $\Leftrightarrow$ (3) follows by Lemma F2.3.20. ☐

Theorem 6.6. *Let T be a complete first-order theory. The following conditions are equivalent.*

(1) *T is superstable.*

(2) *T is κ-stable if, and only if, $\kappa \geq \mathrm{st}(T)$.*

(3) *There is a cardinal λ such that T is κ-stable for all $\kappa \geq \lambda$.*

Proof. (2) ⇒ (3) is trivial.

(1) ⇒ (2) According to Theorem 6.3, T is κ-stable if, and only if,

$$\kappa \geq \mathrm{st}(T) \quad \text{and} \quad \kappa = \kappa^{<\aleph_0} .$$

As the second condition is vacuously true, the claim follows.

(3) ⇒ (1) Fix a cardinal $\kappa \geq \lambda$ with $\mathrm{cf}\,\kappa = \aleph_0$. As $\kappa \geq \lambda \geq \mathrm{st}(T)$, it follows by Theorem 6.3 that $\kappa = \kappa^{<\mathrm{fc}(\downarrow^{\mathrm{f}})}$. Since $\kappa^{\aleph_0} = \kappa^{\mathrm{cf}\,\kappa} > \kappa$, this implies that $\mathrm{fc}(\downarrow^{\mathrm{f}}) \leq \aleph_0$. Hence, the claim follows by Theorem 6.5. □

Corollary 6.7. *Every $\aleph_0$-stable theory is superstable.*

Proof. If T is $\aleph_0$-stable, then $\mathrm{fc}(\downarrow^{\mathrm{f}}) \leq \mathrm{st}(T) = \aleph_0$ and it follows by Theorem 6.5 that T is supersimple. □

We conclude this section by noting that, for countable theories, the characterisation in Theorem 6.3 leaves only four possibilities.

Theorem 6.8. *Every countable complete theory T satisfies exactly one of the following conditions:*

(1) *T is totally transcendental. Then*

$$T \text{ is } \kappa\text{-stable} \quad \text{iff} \quad \kappa \geq \aleph_0 .$$

(2) *T is superstable, but not totally transcendental. Then*

$$T \text{ is } \kappa\text{-stable} \quad \text{iff} \quad \kappa \geq 2^{\aleph_0} .$$

(3) *T is stable, but not superstable. Then*

$$T \text{ is } \kappa\text{-stable} \quad \text{iff} \quad \kappa = \kappa^{\aleph_0} .$$

(4) *T is unstable.*

Proof. Let T be a stable theory. If T is $\aleph_0$-stable, we have seen in Theorem C3.5.18 that T is totally transcendental and κ-stable for all infinite κ.

Hence, suppose that T is not $\aleph_0$-stable. Then there exists a countable set $U \subseteq \mathbb{M}$ with $|S^{<\omega}(U)| > \aleph_0$. According to Corollary B5.7.5, this implies that $|S^{<\omega}(U)| = 2^{\aleph_0}$. Consequently, $\mathrm{st}(T) \geq 2^{\aleph_0}$. By Lemma 6.2 (c), it follows that $\mathrm{st}(T) = 2^{\aleph_0}$.

Furthermore, we have $\mathrm{fc}(\downarrow^{\mathrm{f}}) \leq \aleph_1$, according to Lemma 6.2 (b). If $\mathrm{fc}(\downarrow^{\mathrm{f}}) = \aleph_0$, then T is superstable and it follows by Theorem 6.6 that

$$T \text{ is } \kappa\text{-stable} \quad \text{iff} \quad \kappa \geq \mathrm{st}(T) = 2^{\aleph_0}\,.$$

If $\mathrm{fc}(\downarrow^{\mathrm{f}}) = \aleph_1$, Theorem 6.3 implies that T is κ-stable if, and only if, $\kappa \geq \mathrm{st}(T) = 2^{\aleph_0}$ and $\kappa = \kappa^{\aleph_0}$. Note that $\kappa = \kappa^{\aleph_0}$ implies $2^{\aleph_0} \leq \kappa^{\aleph_0} = \kappa$. Hence, the first condition is superfluous and

$$T \text{ is } \kappa\text{-stable} \quad \text{iff} \quad \kappa = \kappa^{\aleph_0}\,. \qquad \square$$

G2. *Models of stable theories*

1. *Isolation relations*

In this chapter, we study the structure of models of a stable theory. Our main tool will be a generalisation of the notion of a construction which we introduced in Section E3.4. This generalisation is based on the notion of a so-called isolation relation.

Definition 1.1. (a) A ternary relation $\surd$ on small subsets of $\mathbb{M}$ is an *isolation relation* if it is an abstract independence relation satisfying the axioms (INV), (BMON), and

(RSH) *Right Shift.*

$$AC \surd_U B \quad \text{and} \quad C \surd_U AB \quad \text{implies} \quad A \surd_U BC\,.$$

If $\bar{a} \surd_U B$, we say that $\operatorname{tp}(\bar{a}/B)$ is $\surd$*-isolated* over U. For $U = B$, we sometimes drop the subscript and abbreviate

$$A \surd_U U \quad \text{by} \quad A \surd U\,.$$

(b) The *left base-monotonicity cardinal* $\operatorname{lbm}(\surd)$ of an isolation relation $\surd$ is the least cardinal κ such that there are sets A, B, C, U with

$$|C| \leq \kappa\,, \quad AC \surd_U B\,, \quad \text{and} \quad A \not\surd_{UC} B\,.$$

If such a cardinal does not exist, we set $\operatorname{lbm}(\surd) = \infty$.

Thus, the difference between an isolation relation and a preforking relation is that we have dropped (DEF) while we have provided a converse to Lemma F2.2.4 by (RSH).

Lemma 1.2. *Every symmetric preforking relation is an isolation relation with* $\mathrm{lbm}(\downarrow) = \infty$.

Proof. By symmetry and base monotonicity it follows that $\mathrm{lbm}(\downarrow) = \infty$. Hence, we only have to prove (RSH). Let $AC \downarrow_U B$ and $C \downarrow_U A$. Then $B \downarrow_U AC$ and Lemma F2.2.4 implies that $BC \downarrow_U A$. Hence, $A \downarrow_U BC$. □

Our main example of an isolation relation will be the relation $\overset{\mathrm{at}}{\surd}$. To illustrate the concept, we also introduce three variants.

Definition 1.3. Let κ be an infinite cardinal.

$$\begin{array}{rll} A \overset{\mathrm{at}\,\kappa}{\surd}_U B & :\text{iff} & \text{for every finite } \bar a \subseteq A\,, \text{ there exists a set} \\ & & \Phi \subseteq \mathrm{tp}(\bar a/U) \text{ of size } |\Phi| < \kappa \text{ such that} \\ & & \quad \mathbb{M} \vDash \Phi(\bar a') \;\Rightarrow\; \bar a' \equiv_{UB} \bar a\,. \\ A \downarrow^{\mathrm{wo}}_U B & :\text{iff} & A' \equiv_U A \qquad\;\; \Rightarrow\; A' \equiv_{UB} A\,. \\ A \downarrow^{\mathrm{a}}_U B & :\text{iff} & A' \equiv_{\mathrm{acl}^{\mathrm{eq}}(U)} A \;\Rightarrow\; A' \equiv_{UB} A\,. \end{array}$$

Lemma 1.4. (a) $\overset{\mathrm{at}}{\surd} \subseteq \overset{\mathrm{at}\,\kappa}{\surd} \subseteq \downarrow^{\mathrm{wo}} \subseteq \overset{\mathrm{d}}{\surd}$

(b) $\overset{\mathrm{at}}{\surd}$ *is an isolation relation with* $\mathrm{lbm}(\overset{\mathrm{at}}{\surd}) \geq \aleph_0$.

(c) $\overset{\mathrm{at}\,\kappa}{\surd}$ *satisfies all axioms of an isolation relation except for* (RSH). *For regular cardinals* κ, *we have* $\mathrm{lbm}(\overset{\mathrm{at}\,\kappa}{\surd}) \geq \kappa$.

(d) $\downarrow^{\mathrm{wo}}$ *is a symmetric isolation relation with* $\mathrm{lbm}(\downarrow^{\mathrm{wo}}) = \infty$.

Proof. (a) The first two inclusions follow immediately from the definitions. For the last one, suppose that $\bar a \downarrow^{\mathrm{wo}}_U \bar b$. To prove that $\bar a \overset{\mathrm{d}}{\surd}_U \bar b$, let $(\bar b_n)_{n<\omega}$ be an indiscernible sequence over U with $\bar b_0 = \bar b$. According to Lemma F3.1.3, it is sufficient to find a tuple $\bar a' \equiv_{U\bar b} \bar a$ such that

$$\bar b_m \equiv_{U\bar a'} \bar b_n\,, \quad \text{for all } m, n < \omega\,.$$

We will show in (d) that $\downarrow^{\mathrm{wo}}$ is symmetric. Hence, we also have $\bar b \downarrow^{\mathrm{wo}}_U \bar a$ and

$$\bar b_n \equiv_U \bar b \quad \text{implies} \quad \bar b_n \equiv_{U\bar a} \bar b\,, \qquad \text{for all } n < \omega\,.$$

Consequently, the tuple $\bar a$ itself has the desired properties.

(b) We have already seen in Lemma F2.3.3 that $\sqrt[\mathrm{at}]{}$ is an abstract independence relation satisfying (INV) and (BMON). The fact that $\mathrm{lbm}(\sqrt[\mathrm{at}]{}) \geq \aleph_0$ follows by (c) for $\kappa = \aleph_0$. Hence, the only axiom that remains to be verified is (RSH).

Suppose that $AC \sqrt[\mathrm{at}]{}_U B$ and $C \sqrt[\mathrm{at}]{}_U AB$. To check that $A \sqrt[\mathrm{at}]{}_U BC$, consider a finite tuple $\bar a \subseteq A$. Since $AC \sqrt[\mathrm{at}]{}_U B$, there is a formula $\varphi(\bar x)$ over U isolating $\mathrm{tp}(\bar a/UB)$. We claim that φ also isolates $\mathrm{tp}(\bar a/UBC)$. Let $\bar a'$ be a tuple satisfying φ. Then $\bar a' \equiv_{UB} \bar a$ and we have to show that $\bar a' \equiv_{UBC} \bar a$. Given a finite tuple $\bar c \subseteq C$, we fix some tuple $\bar c'$ such that

$$\bar a'\bar c \equiv_{UB} \bar a\bar c'\,.$$

Since $\bar c \sqrt[\mathrm{at}]{}_U AB$, $\bar c' \equiv_U \bar c$ implies that $\bar c' \equiv_{UB\bar a} \bar c$. Hence,

$$\bar a'\bar c \equiv_{UB} \bar a\bar c' \equiv_{UB} \bar a\bar c\,.$$

We have shown that

$$\bar a' \equiv_{UB\bar c} \bar a\,, \qquad \text{for all finite } \bar c \subseteq C\,.$$

Consequently, $\bar a' \equiv_{UBC} \bar a$.

(c) (INV) and (FIN) follow immediately from the definition.

(BMON) Suppose that $\bar a \sqrt[\mathrm{at}\,\kappa]{}_U BC$. By (FIN) we may assume that $\bar a$ is finite. Hence, there exists a set $\Phi(\bar x) \subseteq \mathrm{tp}(\bar a/U)$ of size $|\Phi| < \kappa$ such that

$$\mathbb{M} \vDash \Phi(\bar a') \quad \text{implies} \quad \bar a' \equiv_{UBC} \bar a\,.$$

Since $\Phi \subseteq \mathrm{tp}(\bar a/UC)$, the same set shows that $\bar a \sqrt[\mathrm{at}\,\kappa]{}_{UC} B$.

(MON) Suppose that $\bar a\bar c \;{}^{\mathrm{at}\,\kappa}\!\!\sqrt{}_U\; BD$. We claim that $\bar a \;{}^{\mathrm{at}\,\kappa}\!\!\sqrt{}_U\; B$. Again we may assume that $\bar a$ and $\bar c$ are finite. According to the definition, there exists a set $\Phi(\bar x, \bar y) \subseteq \mathrm{tp}(\bar a\bar c/U)$ of size $|\Phi| < \kappa$ such that

$$\mathbb{M} \vDash \Phi(\bar a', \bar c') \quad \text{implies} \quad \bar a'\bar c' \equiv_{UBD} \bar a\bar c\,.$$

We claim that the set

$$\Psi(\bar x) := \big\{\, \exists \bar y \bigwedge \Phi_0 \;\big|\; \Phi_0 \subseteq \Phi \text{ finite} \,\big\}$$

is the desired witness for $\bar a \;{}^{\mathrm{at}\,\kappa}\!\!\sqrt{}_U\; B$. Hence, suppose that $\mathbb{M} \vDash \Psi(\bar a')$. By definition of Ψ it follows that, for every finite subset $\Phi_0 \subseteq \Phi$, we can find some tuple $\bar c'$ with $\mathbb{M} \vDash \Phi_0(\bar a', \bar c')$. By compactness, this implies that there is some tuple $\bar c'$ with $\mathbb{M} \vDash \Phi(\bar a', \bar c')$. Consequently, $\bar a'\bar c' \equiv_{UBD} \bar a\bar c$. In particular, we have $\bar a' \equiv_{UB} \bar a$.

(NOR) Suppose that $A \;{}^{\mathrm{at}\,\kappa}\!\!\sqrt{}_U\; B$. To show that $AU \;{}^{\mathrm{at}\,\kappa}\!\!\sqrt{}_U\; BU$, let $\bar a \subseteq A$ and $\bar c = \langle c_0, \dots, c_{n-1}\rangle \subseteq U$ be finite. There exists a set $\Phi(\bar x) \subseteq \mathrm{tp}(\bar a/U)$ of size $|\Phi| < \kappa$ such that

$$\mathbb{M} \vDash \Phi(\bar a') \quad \text{implies} \quad \bar a' \equiv_{UBC} \bar a\,.$$

Setting

$$\Psi(\bar x, \bar y) := \Phi(\bar x) \cup \{y_0 = c_0, \dots, y_{n-1} = c_{n-1}\}$$

it follows that

$$\mathbb{M} \vDash \Psi(\bar a', \bar c') \quad \text{implies} \quad \bar a' \equiv_{UBC} \bar a \quad \text{and} \quad \bar c' = \bar c\,.$$

In particular, we have $\bar a'\bar c' \equiv_{UB} \bar a\bar c$.

(LRF) Let $A, B \subseteq \mathbb{M}$. To show that $A \;{}^{\mathrm{at}\,\kappa}\!\!\sqrt{}_A\; B$, consider a finite tuple $\bar a = \langle a_0, \dots, a_{n-1}\rangle \subseteq A$. We set

$$\Phi(\bar x) := \{x_0 = a_0, \dots, x_{n-1} = a_{n-1}\}\,.$$

Then $\mathbb{M} \vDash \Phi(\bar a')$ implies that $\bar a' = \bar a$. In particular, we have $\bar a' \equiv_{AB} \bar a$.

(LTR) Suppose that $A_0 A_1 A_2 \overset{\text{at}\,\kappa}{\surd}_{A_0 A_1} B$ and $A_0 A_1 \overset{\text{at}\,\kappa}{\surd}_{A_0} B$. By (NOR) it is sufficient to prove that $A_1 A_2 \overset{\text{at}\,\kappa}{\surd}_{A_0} B$. Hence, let $\bar a_1 \subseteq A_1$ and $\bar a_2 \subseteq A_2$ be finite. By assumption, there are sets $\Phi(\bar x_1, \bar x_2) \subseteq \operatorname{tp}(\bar a_1 \bar a_2 / A_0 A_1)$ and $\Psi(\bar x_1) \subseteq \operatorname{tp}(\bar a_1 / A_0)$ of size $|\Phi|, |\Psi| < \kappa$ such that

$$\begin{aligned} &\mathbb{M} \vDash \Phi(\bar a_1', \bar a_2') \quad \text{implies} \quad \bar a_1' \bar a_2' \equiv_{B A_0 A_1} \bar a_1 \bar a_2 \,, \\ \text{and} \quad &\mathbb{M} \vDash \Psi(\bar a_1') \qquad \text{implies} \quad \bar a_1' \equiv_{B A_0} \bar a_1 \,. \end{aligned}$$

Let $U \subseteq A_1$ be the set of parameters from A_1 that are used in Φ and let Φ' be the set of formulae obtained from Φ by replacing each parameter $c \in U$ by a variable y_c. For every finite $\bar c \subseteq U$, there exists a set $\Gamma_{\bar c}(\bar y) \subseteq \operatorname{tp}(\bar c / A_0)$ of size $|\Gamma_{\bar c}| < \kappa$ such that

$$\mathbb{M} \vDash \Psi(\bar c') \quad \text{implies} \quad \bar c' \equiv_{B A_0} \bar c \,.$$

Suppose that $\bar x_1 = \langle x_1^0, \dots, x_1^{n-1} \rangle$ and $\bar a_1 = \langle a_1^0, \dots, a_1^{n-1} \rangle$. We set

$$\begin{aligned} \Xi(\bar x_1, \bar x_2, (y_c)_{c \in U}) := \Phi' \cup \Psi \cup \bigcup_{\bar c \subseteq U} \Gamma_{\bar c} \\ \cup \{ x_1^0 = y_{a_1^0}, \dots, x_1^{n-1} = y_{a_1^{n-1}} \} \,. \end{aligned}$$

Then

$$\mathbb{M} \vDash \Xi(\bar a_1', \bar a_2', \bar c') \quad \text{implies} \quad \bar a_1' \equiv_{B A_0} \bar a_1 \quad \text{and} \quad \bar c' \equiv_{B A_0} \bar c \,,$$

where $\bar c$ is an enumeration of U. Hence, there exists an automorphism $\pi \in \operatorname{Aut} \mathbb{M}_{B A_0}$ with $\pi[\bar c'] = \bar c$. By the equations added to Ξ, it follows that $\pi[\bar a_1'] = \bar a_1$. Consequently,

$$\mathbb{M} \vDash \Xi(\bar a_1', \bar a_2', \bar c') \quad \text{implies} \quad \mathbb{M} \vDash \Xi(\pi[\bar a_1'], \pi[\bar a_2'], \pi[\bar c']) \,.$$

Hence, $\mathbb{M} \vDash \Phi(\bar a_1, \pi[\bar a_2'], \bar c)$, which means that

$$\pi[\bar a_2'] \equiv_{B A_0 A_1} \bar a_2 \,.$$

Consequently,

$$\bar a_1' \bar a_2' \equiv_{B A_0} \bar a_1 \pi[\bar a_2'] \equiv_{B A_0} \bar a_1 \bar a_2 \,.$$

To compute $\mathrm{lbm}(\sqrt[\mathrm{at}\kappa]{/})$, suppose that κ is regular and $AC \sqrt[\mathrm{at}\kappa]{/}_U B$ for $|C| < \kappa$. We have to show that $A \sqrt[\mathrm{at}\kappa]{/}_{UC} B$. Hence, let $\bar a \subseteq A$ be finite. For every finite tuple $\bar c \subseteq C$, there exists a set $\Phi_{\bar c}(\bar x, \bar x') \subseteq \mathrm{tp}(\bar a\bar c/U)$ of size $|\Phi_{\bar c}| < \kappa$ such that

$$\mathbb{M} \vDash \Phi_{\bar c}(\bar a', \bar c') \quad \text{implies} \quad \bar a'\bar c' \equiv_{UB} \bar a\bar c\,.$$

We set

$$\Psi(\bar x) := \bigcup \{\, \Phi_{\bar c}(\bar x, \bar c) \mid \bar c \subseteq C \text{ finite} \,\}\,.$$

Then $|\Psi| < \kappa$ since κ is regular. Furthermore, $\mathbb{M} \vDash \Psi(\bar a')$ implies

$$\bar a'\bar c \equiv_{UB} \bar a\bar c\,, \qquad \text{for all finite } \bar c \subseteq C\,.$$

Hence, $\bar a'C \equiv_{UB} \bar aC$, which implies that $\bar a' \equiv_{UBC} \bar a$.

(d) (INV) follows immediately form the definition.

(MON) Suppose that $AC \downarrow^{\mathrm{wo}}_U BD$. To show that $A \downarrow^{\mathrm{wo}}_U B$, consider a set $A' \equiv_U A$. Then there exists a set C' such that $A'C' \equiv_U AC$. By assumption, this implies that $A'C' \equiv_{UBD} AC$. In particular, we have $A' \equiv_{UB} A$.

(BMON) Suppose that $A \downarrow^{\mathrm{wo}}_U BC$. To show that $A \downarrow^{\mathrm{wo}}_{UC} B$, consider a set $A' \equiv_{UC} A$. Then $A' \equiv_U A$, which implies that $A' \equiv_{UBC} A'$.

(NOR) Suppose that $A \downarrow^{\mathrm{wo}}_U B$. To show that $AU \downarrow^{\mathrm{wo}}_U BU$, consider sets $A'U' \equiv_U AU$. Then $U' = U$ and $A' \equiv_U A$. Hence, $A' \equiv_{UB} A$, which implies that $A'U' \equiv_{UB} AU$.

(LRF) Let A and B be sets. To show that $A \downarrow^{\mathrm{wo}}_A B$, let $A' \equiv_A A$. Then $A' = A$, which implies that $A' \equiv_{AB} A$.

(SYM) Suppose that $A \downarrow^{\mathrm{wo}}_U B$. To show that $B \downarrow^{\mathrm{wo}}_U A$, consider a set $B' \equiv_U B$. We fix a set A' such that $B'A \equiv_U BA'$. Then $A' \equiv_U A$, which implies that $A' \equiv_{UB} A$. Hence, $B'A \equiv_U BA' \equiv_U BA$, that is, $B' \equiv_{UA} B$.

(LTR) Since we have already proved symmetry, it is sufficient to show that $\downarrow^{\mathrm{wo}}$ is right transitive. Hence, suppose that $A \downarrow^{\mathrm{wo}}_{B_0} B_1$ and $A \downarrow^{\mathrm{wo}}_{B_1} B_2$, for $B_0 \subseteq B_1 \subseteq B_2$. To show that $A \downarrow^{\mathrm{wo}}_{B_0} B_2$, consider a set $A' \equiv_{B_0} A$. Then $A' \equiv_{B_1} A$, which implies that $A' \equiv_{B_2} A$.

(FIN) Suppose that $A_0 \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{wo}}_U B$, for all finite $A_0 \subseteq A$. To show that $A \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{wo}}_U B$, let $A' \equiv_U A$. Fix an automorphism $\pi \in \operatorname{Aut} \mathbb{M}_U$ with $\pi[A'] = A$ and consider a finite subset $A'_0 \subseteq A'$. By assumption, $A'_0 \equiv_U \pi[A'_0]$ implies that $A'_0 \equiv_{UB} \pi[A'_0]$. Consequently, we have

$$A'_0 \equiv_{UB} \pi[A'_0], \quad \text{for all finite } A'_0 \subseteq A'.$$

This implies that $A' \equiv_{UB} \pi[A'] = A$.

(RSH) Suppose that $AC \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{wo}}_U B$ and $C \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{wo}}_U AB$. By symmetry, it follows that $B \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{wo}}_U AC$. Hence, $B \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{wo}}_{UC} A$, which implies that $BC \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{wo}}_{UC} A$. Together with $C \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{wo}}_U A$ it follows by left transitivity that $BC \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{wo}}_U A$. Hence, $A \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{wo}}_U BC$.

Finally, the fact that $\operatorname{lbm}(\mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{wo}}) = \infty$ follows by symmetry and base monotonicity. □

Exercise 1.1. Show that $\mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{a}}$ satisfies all axioms of a symmetric isolation relation except for (LTR) and (RSH). Furthermore, $\operatorname{lbm}(\mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{a}}) = \infty$.

Remark. One can show that, if the theory in question is stable, the relation $\mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{a}}$ is also transitive and, thus, an isolation relation.

Recall that we write $\bar{a} \mathop{\smile\!\!\!\!\!\!\mid}^{!}_U B$ if $\operatorname{tp}(\bar{a}/UB)$ is the unique free extension of $\operatorname{tp}(\bar{a}/U)$ over $U \cup B$. This relation will be used in Section 5 below. It inherits some, but not all properties from $\mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}$. In particular, it is a symmetric isolation relation, but not necessarily a forking relation.

Lemma 1.5. *Let T be a stable theory. The relation $\mathop{\smile\!\!\!\!\!\!\mid}^{!}$ is a symmetric isolation relation with $\mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{wo}} \subseteq \mathop{\smile\!\!\!\!\!\!\mid}^{!} \subseteq \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}$.*

Proof. The second inclusion follows immediately from the definition. For the first one, suppose that $\bar{a} \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{wo}}_U B$. By Lemma 1.4, it follows that $\bar{a} \sqrt[\mathrm{d}]{}_U B$. As T is stable, this is equivalent to $\bar{a} \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}_U B$. For uniqueness, suppose that $\bar{a}' \equiv_U \bar{a}$ is another tuple with $\bar{a}' \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}_U B$. Then $\bar{a} \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{wo}}_U B$ implies that $\bar{a}' \equiv_{UB} \bar{a}$. Consequently, $\bar{a} \mathop{\smile\!\!\!\!\!\!\mid}^{!}_U B$.

It remains to prove that $\mathop{\smile\!\!\!\!\!\!\mid}^{!}$ is a symmetric isolation relation.

(INV) follows immediately from the definition.

(MON) Suppose that $\bar a_0 \bar a_1 \mathop{\smile\!\!\!\!\!\!|}^{!}_{U} B$ and let $B_0 \subseteq B$. If $\operatorname{tp}(\bar a_0/UB)$ is the unique free extension of $\operatorname{tp}(\bar a_0/U)$ over $U \cup B$, then $\operatorname{tp}(\bar a_0/UB_0)$ is its unique free extension over $U \cup B_0$. Hence, it is sufficient show that $\bar a_0 \bar a_1 \mathop{\smile\!\!\!\!\!\!|}^{!}_{U} B$ implies $\bar a_0 \mathop{\smile\!\!\!\!\!\!|}^{!}_{U} B$. By monotonicity of $\mathop{\smile\!\!\!\!\!\!|}^{\mathrm f}$, we only need to prove uniqueness.

Consider a tuple $\bar a'_0 \equiv_U \bar a_0$ with $\bar a'_0 \mathop{\smile\!\!\!\!\!\!|}^{\mathrm f}_{U} B$. We have to show that $\bar a'_0 \equiv_{UB} \bar a_0$. Choose $\bar a'_1$ such that $\bar a'_0 \bar a'_1 \equiv_U \bar a_0 \bar a_1$, and let $\bar a''_1$ be a tuple with

$$\bar a''_1 \equiv_{U\bar a'_0} \bar a'_1 \quad\text{and}\quad \bar a''_1 \mathop{\smile\!\!\!\!\!\!|}^{\mathrm f}_{U\bar a'_0} B\,.$$

Since $\bar a'_0 \mathop{\smile\!\!\!\!\!\!|}^{\mathrm f}_{U} B$, transitivity implies that

$$\bar a'_0 \bar a''_1 \mathop{\smile\!\!\!\!\!\!|}^{\mathrm f}_{U} B\,.$$

As $\bar a'_0 \bar a''_1 \equiv_U \bar a'_0 \bar a'_1 \equiv_U \bar a_0 \bar a_1$, it follows that

$$\bar a_0 \bar a_1 \mathop{\smile\!\!\!\!\!\!|}^{!}_{U} B \quad\text{implies}\quad \bar a'_0 \bar a''_1 \equiv_{UB} \bar a_0 \bar a_1\,.$$

In particular, $\bar a'_0 \equiv_{UB} \bar a_0$.

(NOR) Suppose that $\bar a \mathop{\smile\!\!\!\!\!\!|}^{!}_{U} B$. Then $\bar a \mathop{\smile\!\!\!\!\!\!|}^{\mathrm f}_{U} B$ implies $U\bar a \mathop{\smile\!\!\!\!\!\!|}^{\mathrm f}_{U} BU$. Hence, we only need to prove uniqueness. Let $\bar c$ be an enumeration of U and suppose that there are tuples $\bar a'\bar c' \equiv_U \bar a\bar c$ such that $\bar a'\bar c' \mathop{\smile\!\!\!\!\!\!|}^{\mathrm f}_{U} BU$. Then $\bar c' = \bar c$, and $\bar a' \equiv_U \bar a$ implies that $\bar a' \equiv_{UB} \bar a$. Hence, $\bar a'\bar c' = \bar a'\bar c \equiv_{UB} \bar a\bar c$.

(BMON) Suppose that $\bar a \mathop{\smile\!\!\!\!\!\!|}^{!}_{U} BC$. We claim that $\bar a \mathop{\smile\!\!\!\!\!\!|}^{!}_{UC} B$. Since $\mathop{\smile\!\!\!\!\!\!|}^{\mathrm f}$ is base monotone, we only need to prove uniqueness. Hence, consider a tuple $\bar a' \equiv_{UC} \bar a$ with $\bar a' \mathop{\smile\!\!\!\!\!\!|}^{\mathrm f}_{UC} B$. Then

$$\bar a' \equiv_U \bar a \quad\text{and}\quad \bar a \mathop{\smile\!\!\!\!\!\!|}^{!}_{U} BC \quad\text{implies}\quad \bar a' \equiv_{UBC} \bar a\,.$$

(FIN) If $A \mathop{\smile\!\!\!\!\!\!|}^{!}_{U} B$, then (MON) implies that $A_0 \mathop{\smile\!\!\!\!\!\!|}^{!}_{U} B$, for every finite $A_0 \subseteq A$. Conversely, suppose that $A \not\mathop{\smile\!\!\!\!\!\!|}^{!}_{U} B$. If $A \not\mathop{\smile\!\!\!\!\!\!|}^{\mathrm f}_{U} B$, there is a finite subset $A_0 \subseteq A$ with $A_0 \not\mathop{\smile\!\!\!\!\!\!|}^{\mathrm f}_{U} B$ and we are done. Hence, suppose that $A \mathop{\smile\!\!\!\!\!\!|}^{\mathrm f}_{U} B$. Then there is a set $A' \equiv_U A$ such that $A' \mathop{\smile\!\!\!\!\!\!|}^{\mathrm f}_{U} B$ and $A' \not\equiv_{UB} A$. Let $\pi \in \operatorname{Aut} \mathbb{M}_U$ be an automorphism with $\pi[A] = A'$. Since $A' \not\equiv_{UB} A$, we can find a finite tuple $\bar a \subseteq A$ with $\pi(\bar a) \not\equiv_{UB} \bar a$. As $\pi(\bar a) \equiv_U \bar a$ and $\pi(\bar a) \mathop{\smile\!\!\!\!\!\!|}^{\mathrm f}_{U} B$, it follows that $\bar a \not\mathop{\smile\!\!\!\!\!\!|}^{!}_{U} B$.

(SYM) Let $\bar{a} \mathop{\smile\!\!\!\!\!\!\mid}^{!}_{U} \bar{b}$. By symmetry of $\mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}$, we have $\bar{b} \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}_{U} \bar{a}$. For uniqueness, consider a tuple $\bar{b}' \equiv_U \bar{b}$ with $\bar{b}' \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}_{U} \bar{a}$. There exists an automorphism $\pi \in \operatorname{Aut} \mathbb{M}_U$ with $\pi(\bar{b}') = \bar{b}$. Since $\mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}$ is invariant under automorphisms,

$$\bar{b}' \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}_{U} \bar{a} \quad \text{implies} \quad \bar{b} \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}_{U} \pi(\bar{a}) \,.$$

By symmetry, $\pi(\bar{a}) \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}_{U} \bar{b}$. Since $\pi(\bar{a}) \equiv_U \bar{a}$ and $\bar{a} \mathop{\smile\!\!\!\!\!\!\mid}^{!}_{U} \bar{b}$, it follows that

$$\pi(\bar{a}) \equiv_{U \cup \bar{b}} \bar{a} \,.$$

For every formula $\varphi(\bar{x}, \bar{y})$ over U, we therefore have

$$\begin{aligned} \mathbb{M} \vDash \varphi(\bar{a}, \bar{b}) \quad &\text{iff} \quad \mathbb{M} \vDash \varphi(\pi(\bar{a}), \bar{b}) \\ &\text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}, \pi^{-1}(\bar{b})) \quad \text{iff} \quad \mathbb{M} \vDash \varphi(\bar{a}, \bar{b}') \,. \end{aligned}$$

Consequently, $\bar{b}' \equiv_{U \cup \bar{a}} \bar{b}$.

(LTR) As we already have proved symmetry, it is sufficient to show that, for sets $B_0 \subseteq B_1 \subseteq B_2$,

$$\bar{a} \mathop{\smile\!\!\!\!\!\!\mid}^{!}_{B_0} B_1 \quad \text{and} \quad \bar{a} \mathop{\smile\!\!\!\!\!\!\mid}^{!}_{B_1} B_2 \quad \text{implies} \quad \bar{a} \mathop{\smile\!\!\!\!\!\!\mid}^{!}_{B_0} B_2 \,.$$

By transitivity of $\mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}$, we only need to prove uniqueness. Hence, consider a tuple $\bar{a}' \equiv_{B_0} \bar{a}$ with $\bar{a}' \mathop{\smile\!\!\!\!\!\!\mid}^{\mathrm{f}}_{B_0} B_2$. Then

$$\bar{a} \mathop{\smile\!\!\!\!\!\!\mid}^{!}_{B_0} B_1 \quad \text{implies} \quad \bar{a}' \equiv_{B_1} \bar{a} \,.$$

Hence,

$$\bar{a} \mathop{\smile\!\!\!\!\!\!\mid}^{!}_{B_1} B_2 \quad \text{implies} \quad \bar{a}' \equiv_{B_2} \bar{a} \,.$$

(RSH) Suppose that $AC \mathop{\smile\!\!\!\!\!\!\mid}^{!}_{U} B$ and $C \mathop{\smile\!\!\!\!\!\!\mid}^{!}_{U} AB$. Then

$$B \mathop{\smile\!\!\!\!\!\!\mid}^{!}_{U} AC \quad \text{and} \quad C \mathop{\smile\!\!\!\!\!\!\mid}^{!}_{U} A \,,$$

and it follows by Lemma F2.2.4 that $BC \mathop{\smile\!\!\!\!\!\!\mid}^{!}_{U} A$. Thus, $A \mathop{\smile\!\!\!\!\!\!\mid}^{!}_{U} BC$. □

2. *Constructions*

We can use isolation relations to stratify a structure such that every part is isolated over the previous ones. This leads to a generalised notion of a construction.

Throughout the chapter, we will use the following notation. Given a sequence $(A_\alpha)_{\alpha<\gamma}$ of sets, an ordinal $\alpha \leq \gamma$, and a set $I \subseteq \gamma$, we will write

$$A[I] := \bigcup_{i\in I} A_i\,, \quad A[<\alpha] := \bigcup_{i<\alpha} A_i\,, \quad \text{and} \quad A[\leq\alpha] := \bigcup_{i\leq\alpha} A_i\,.$$

Definition 2.1. Let $\surd$ be a ternary relation on small subsets of $\mathbb{M}$ and $A, U \subseteq \mathbb{M}$.

(a) A $\surd$*-stratification* of A over U is a sequence $\zeta = (B_\alpha)_{\alpha<\gamma}$ of disjoint sets $B_\alpha \subseteq A$ such that $A = B[<\gamma]$ and

$$B_\alpha \surd UB[<\alpha]\,, \quad \text{for all } \alpha < \gamma\,.$$

(b) A $\surd$-stratification ζ is a $\surd$*-construction* if each set B_α is a singleton. In this case, we identify ζ with the corresponding sequence $(b_\alpha)_{\alpha<\gamma}$ of elements $b_\alpha \in B_\alpha$. We say that a set A is $\surd$*-constructible* over U if there exists a $\surd$-construction of A over U.

(c) Let $\zeta = (B_\alpha)_{\alpha<\gamma}$ be a $\surd$-stratification. The *locality cardinal* $\mathrm{loc}(\zeta)$ of ζ is the least cardinal κ such that, for every $\alpha < \gamma$, there exists a set $C_\alpha \subseteq U \cup B[<\alpha]$ of size $|C_\alpha| < \kappa$ such that

$$B_\alpha \surd_{C_\alpha} UB[<\alpha]\,.$$

Remark. In this terminology, the kind of constructions introduced in Section E3.4 are $\overset{\mathrm{at}}{\surd}$-constructions.

We will use $\surd$-stratifications to study the structure of $\surd$-constructible models. To do so, we will frequently be interested in whether a given subset of a $\surd$-constructible set is itself $\surd$-constructible. A simple sufficient condition is given by the notion of a *closed* set.

Definition 2.2. Let $\zeta = (B_\alpha)_{\alpha<\gamma}$ be a $\surd$-stratification of A. A subset $I \subseteq \gamma$ of indices is *ζ-closed* over U if

$$B_\alpha \surd_{UB[I\cap\alpha]} UB[<\alpha], \quad \text{for all } \alpha \in I.$$

Similarly, we call a set $C \subseteq A$ *ζ-closed* if it is of the form $C = B[I]$, for some ζ-closed set $I \subseteq \gamma$.

Lemma 2.3. *Let $\surd$ be an isolation relation and $\zeta = (B_\alpha)_{\alpha<\gamma}$ a $\surd$-stratification of A over U. If $I \subseteq \gamma$ is ζ-closed over U, then $(B_\alpha)_{\alpha\in I}$ is a $\surd$-stratification of $B[I]$ over U.*

Proof. Consider an index $\alpha \in I$. As I is ζ-closed, we have

$$B_\alpha \surd_{UB[I\cap\alpha]} UB[<\alpha].$$

By (MON), this implies that

$$B_\alpha \surd_{UB[I\cap\alpha]} UB[I\cap\alpha]. \qquad \square$$

In particular, ζ-closed subsets of a $\surd$-constructible set are themselves $\surd$-constructible. Before proving further properties of ζ-closed sets, let us present a lemma with several ways to construct such sets.

Lemma 2.4. *Let $\surd$ be a relation satisfying* (BMON), *$\zeta = (B_\alpha)_{\alpha<\gamma}$ a $\surd$-stratification of A over U, and let $\kappa \geq \mathrm{loc}(\zeta)$ be a regular cardinal.*

(a) *A union of ζ-closed sets is ζ-closed.*

(b) *If $I \subseteq \gamma$ is ζ-closed over U, then so is $I \cap \beta$ for every $\beta \leq \gamma$.*

(c) *Every index $\alpha < \gamma$ is contained in a ζ-closed set $I \subseteq \Downarrow\alpha$ over U of size $|I| < \kappa$.*

(d) *For every set $C \subseteq A$ of size $|C| < \kappa$, there is some ζ-closed set $I \subseteq \gamma$ over U of size $|I| < \kappa$ with $C \subseteq B[I]$.*

Proof. (a) Let $\mathcal{I}$ be a set of ζ-closed sets. To show that $K := \bigcup \mathcal{I}$ is ζ-closed, consider an index $\alpha \in K$. Then $\alpha \in I$, for some ζ-closed set $I \in \mathcal{I}$, and

$$B_\alpha \surd_{UB[I \cap \alpha]} UB[<\alpha] \quad \text{implies} \quad B_\alpha \surd_{UB[K \cap \alpha]} UB[<\alpha]$$

by (BMON).

(b) Let I be ζ-closed and fix $\beta \leq \gamma$. For $\alpha \in I \cap \beta$, we have $I \cap \alpha = (I \cap \beta) \cap \alpha$. Hence,

$$B_\alpha \surd_{UB[I \cap \alpha]} UB[<\alpha] \quad \text{implies} \quad B_\alpha \surd_{UB[(I \cap \beta) \cap \alpha]} UB[<\alpha] .$$

(c) We prove the claim by induction on α. There exists a set $C_\alpha \subseteq U \cup B[<\alpha]$ of size $|C_\alpha| < \operatorname{loc}(\zeta) \leq \kappa$ such that

$$B_\alpha \surd_{C_\alpha} UB[<\alpha] .$$

Set $J := \{ \beta < \alpha \mid C_\alpha \cap B_\beta \neq \varnothing \}$. By inductive hypothesis, every index $\beta \in J$ is contained in some ζ-closed set $I_\beta \subseteq \Downarrow\beta$ of size $|I_\beta| < \kappa$. By (a), the union $I := \bigcup_{\beta \in J} I_\beta$ is also ζ-closed. Since

$$C_\alpha \subseteq U \cup B[J] \subseteq U \cup B[I] = U \cup B[I \cap \alpha] ,$$

it follows by (BMON) that

$$B_\alpha \surd_{C_\alpha} UB[<\alpha] \quad \text{implies} \quad B_\alpha \surd_{UB[I \cap \alpha]} UB[<\alpha] .$$

Since $I \subseteq \alpha$, this implies that $I \cup \{\alpha\}$ is also ζ-closed. Furthermore, $|I \cup \{\alpha\}| < \kappa$, as κ is regular.

(d) Given $C \subseteq A$, set

$$J := \{ \alpha < \gamma \mid C \cap B_\alpha \neq \varnothing \} .$$

By (c), every $\alpha \in J$ is contained in some ζ-closed set $K_\alpha \subseteq \gamma$ of size $|K_\alpha| < \kappa$. By (a), the union $I := \bigcup_{\alpha < \gamma} K_\alpha$ is ζ-closed. As κ is regular and $|J| \leq |C| < \kappa$, we have $|I| < \kappa$. Since $C \subseteq B[J] \subseteq B[I]$, the claim follows. $\square$

The next proposition collects several properties of ζ-closed sets. In particular, we show that a $\surd$-stratification over U is also a $\surd$-stratification over $U \cup C$, for every ζ-closed set C.

Proposition 2.5. *Let $\surd$ be an isolation relation and $\zeta = (B_\alpha)_{\alpha<\gamma}$ a $\surd$-stratification of A over U.*

(a) *If $I \subseteq \gamma$ is ζ-closed over U, then*

$$B[I] \surd_{UB[I\cap\alpha]} B[<\alpha], \quad \text{for all } \alpha \leq \gamma.$$

(b) *ζ is a $\surd$-stratification of A over $U \cup B[I]$, for every ζ-closed set $I \subseteq \gamma$.*

(c) *If $K \subseteq \gamma$ is a ζ-closed set over U of size $|B[K]| < \mathrm{lbm}(\surd)$, then every set $I \subseteq \gamma$ that is ζ-closed over U is also ζ-closed over $U \cup B[K]$.*

Proof. (a) Fix $\alpha \leq \gamma$. We prove the statement by induction on the minimal ordinal β such that $I \subseteq \beta$. If $I \subseteq \alpha$, then (LRF) and (NOR) imply that

$$B[I] \surd_{UB[I]} B[<\alpha].$$

As $I = I \cap \alpha$, the claim follows.

For the successor step, suppose that $I = \{\beta\} \cup I_0$ where $I_0 \subseteq \beta$ and $\beta \geq \alpha$. Since I is ζ-closed, we have

$$B_\beta \surd_{UB[I\cap\beta]} B[<\beta]$$

which, by (MON), implies that

$$B_\beta \surd_{UB[I_0]B[I\cap\alpha]} B[<\alpha].$$

Furthermore, the set $I_0 = I \cap \beta$ is ζ-closed according to Lemma 2.4 (b). Consequently, the inductive hypothesis yields

$$B[I_0] \surd_{UB[I_0\cap\alpha]} B[<\alpha].$$

Since $I_0 \cap \alpha = I \cap \alpha$, it follows by (NOR) and (LTR) that

$$B_\beta B[I_0] \mathop{\sqrt{}}\nolimits_{UB[I\cap\alpha]} B[<\alpha] .$$

Finally, suppose that I has no maximal element. We have seen in Lemma 2.4 (b) that $I \cap \beta$ is ζ-closed, for all $\beta < \gamma$. By inductive hypothesis, it therefore follows that

$$B[I \cap \beta] \mathop{\sqrt{}}\nolimits_{UB[I\cap\alpha]} B[<\alpha] , \quad \text{for all } \beta \in I .$$

Consequently, (FIN) implies that

$$B[I] \mathop{\sqrt{}}\nolimits_{UB[I\cap\alpha]} B[<\alpha] .$$

(b) We have to show that

$$B_\alpha \sqrt{} \ UB[I]B[<\alpha] , \quad \text{for all } \alpha < \gamma .$$

Hence, let $\alpha < \gamma$. If $\alpha \in I$, the claim follows by (NOR). Thus, suppose that $\alpha \notin I$. By (a), we have

$$B[I] \mathop{\sqrt{}}\nolimits_{UB[I\cap(\alpha+1)]} B[\leq\alpha] .$$

Since $\alpha \notin I$, this implies by (BMON) that

$$B[I] \mathop{\sqrt{}}\nolimits_{UB[<\alpha]} B[<\alpha]B_\alpha .$$

Hence, (BMON) and (NOR) yield

$$UB[I]B_\alpha \mathop{\sqrt{}}\nolimits_{UB[<\alpha]B_\alpha} B[<\alpha] .$$

As $B_\alpha \mathop{\sqrt{}}\nolimits_{UB[<\alpha]} B[<\alpha]$, it follows by (NOR) and (LTR) that

$$B[I]B_\alpha \mathop{\sqrt{}}\nolimits_{UB[<\alpha]} B[<\alpha] .$$

Hence, $B[I] \mathrel{\surd_{UB[<\alpha]}} B[<\alpha]B_\alpha$ implies by (rsh) that

$$B_\alpha \mathrel{\surd_{UB[<\alpha]}} B[I]B[<\alpha] .$$

Consequently, it follows by (bmon) and (nor) that

$$B_\alpha \mathrel{\surd_{UB[I]B[<\alpha]}} UB[I]B[<\alpha] .$$

(c) Let $I, K \subseteq \gamma$ be ζ-closed over U. Then $I \cup K$ is also ζ-closed. Hence, it follows by (a) that

$$B[I \cup K] \mathrel{\surd_{UB[(I \cup K) \cap \alpha]}} B[<\alpha] .$$

For $\alpha \in I$, this implies that

$$B_\alpha B[K] \mathrel{\surd_{UB[K \cap \alpha]B[I \cap \alpha]}} B[<\alpha] .$$

Since $|B[K]| < \mathrm{lbm}(\surd)$, it follows that

$$B_\alpha \mathrel{\surd_{UB[K]B[I \cap \alpha]}} B[<\alpha] ,$$

as desired. □

Corollary 2.6. *Let* $\surd$ *be an isolation relation and* $\zeta = (B_\alpha)_{\alpha<\gamma}$ *a* $\surd$*-stratification of* A *over* U*. Then* $A \mathrel{\surd} UB[<\alpha]$*, for all* $\alpha \leq \gamma$*.*

Proof. Since the set $I := \gamma$ is ζ-closed, this follows immediately from Proposition 2.5 (a). □

We conclude this section by presenting conditions for the existence of a stratification or a construction.

Lemma 2.7. *Let* $\surd$ *be an isolation relation,* $\zeta = (B_\alpha)_{\alpha<\gamma}$ *a* $\surd$*-stratification of* A *over* U*, and let* $(C_\alpha)_{\alpha<\delta}$ *be a sequence of subsets of* A *such that each set* C_α *is* ζ*-closed over* $U \cup C[<\alpha]$*.*

(a) *The union* $C[<\delta]$ *is* ζ*-closed over* U*.*

(b) $(C_\alpha \setminus C[<\alpha])_{\alpha<\delta}$ *is a* $\surd$*-stratification of* $C[<\delta]$ *over* U.

Proof. (a) For each $\beta < \delta$, fix a set $I_\beta \subseteq \gamma$ such that $C_\beta = B[I_\beta]$. We prove that $I[<\delta]$ is ζ-closed over U by induction on δ. Consider an index $\alpha \in I[<\delta]$. Then $\alpha \in I_\beta$ for some ordinal $\beta < \delta$. By inductive hypothesis, $I[<\beta]$ is ζ-closed over U. Hence, it follows by Proposition 2.5 (a) that

$$C[<\beta] \;\surd_{UB[I[<\beta]\cap\alpha]}\; B[<\alpha]\,.$$

By (BMON), this implies that

$$C[<\beta] \;\surd_{UB[I[<\beta]\cap\alpha]B[I_\beta\cap\alpha]}\; B[<\alpha]\,.$$

Furthermore, as C_β is ζ-closed over $U \cup C[<\beta]$, we have

$$B_\alpha \;\surd_{UC[<\beta]B[I_\beta\cap\alpha]}\; B[<\alpha]\,.$$

Therefore, it follows by (LTR) that

$$B_\alpha C[<\beta] \;\surd_{UB[I[<\beta]\cap\alpha]B[I_\beta\cap\alpha]}\; B[<\alpha]\,.$$

By (BMON), this implies that

$$B_\alpha \;\surd_{UB[I[<\delta]\cap\alpha]}\; B[<\alpha]\,.$$

(b) Let $\alpha < \delta$. By (a) and Proposition 2.5 (b), ζ is a $\surd$-stratification of A over $U \cup C[<\alpha]$. As C_α is ζ-closed over $U \cup C[<\alpha]$, it follows therefore by Proposition 2.5 (a) that $C_\alpha \;\surd\; UC[<\alpha]$. □

Lemma 2.8. *Let* $\surd$ *be an isolation relation and* $(B_\alpha)_{\alpha<\gamma}$ *a sequence of sets.*

(a) *If every* B_α *is* $\surd$*-constructible over* $U \cup B[<\alpha]$*, then* $B[<\gamma]$ *is* $\surd$*-constructible over* U.

(b) *Let* ζ *be a* $\surd$*-construction of some set* $A \supseteq B[<\gamma]$ *over* U*. If each* B_α *is* ζ*-closed over* U*, then* $B[<\gamma]$ *is* $\surd$*-constructible over* U.

Proof. (a) For each $\alpha < \gamma$, fix a $\surd$-construction $(b_i^\alpha)_{i<\eta(\alpha)}$ of B_α over $U \cup B[<\alpha]$. Set $\delta(\alpha) := \sum_{i<\alpha} \eta(i)$ and let $(a_\beta)_{\beta<\delta(\gamma)}$ be the concatenation of all sequences $(b_i^\alpha)_i$, for $\alpha < \gamma$, that is,

$$a_{\delta(\alpha)+i} := b_i^\alpha \,, \quad \text{for } \alpha < \gamma \text{ and } i < \eta(\alpha) \,.$$

To prove that $(a_\beta)_{\beta<\delta(\gamma)}$ is a $\surd$-construction of $B[<\gamma]$ over U, consider an index $\beta < \delta(\gamma)$. Then $\beta = \delta(\alpha) + i$, for some $\alpha < \gamma$ and $i < \eta(\alpha)$, and

$$b_i^\alpha \surd UB[<\alpha]b^\alpha[<i] \quad \text{implies} \quad a_\beta \surd Ua[<\beta] \,.$$

(b) According to Lemma 2.7 (a), B_α is $\surd$-constructible over U. Furthermore, Lemma 2.4 (a) implies that each set of the form $B[<\alpha]$ is ζ-closed over U. Hence, it follows by Proposition 2.5 (b) that B_α is $\surd$-constructible over $U \cup B[<\alpha]$. Consequently, the claim follows by (a). □

Lemma 2.9. *Let $\surd$ be an isolation relation. A set A of size $|A| \leq \mathrm{lbm}(\surd)$ is $\surd$-constructible over a set U if, and only if, $A \surd U$.*

Proof. ($\Rightarrow$) follows by Corollary 2.6. For ($\Leftarrow$), let $A \surd U$ and let $\zeta = (a_\alpha)_{\alpha<\kappa}$ be an enumeration of A of length $\kappa := |A|$. We claim that ζ is a $\surd$-construction of A over U. For each $\alpha < \kappa$,

$$A \surd U \quad \text{implies} \quad a_\alpha a[<\alpha] \surd U \,.$$

Since $\big|a[<\alpha]\big| < \kappa \leq \mathrm{lbm}(\surd)$, it follows that $a_\alpha \surd_{Ua[<\alpha]} U$. □

Corollary 2.10. *Let $\surd$ be an isolation relation and let $(B_\alpha)_{\alpha<\gamma}$ be a $\surd$-stratification of A over U where*

$$|B_\alpha| \leq \mathrm{lbm}(\surd) \,, \quad \textit{for all } \alpha < \gamma \,.$$

Then A is $\surd$-constructible over U.

Proof. Since $B_\alpha \surd UB[<\alpha]$, it follows by Lemma 2.9 that each B_α is $\surd$-constructible over $U \cup B[<\alpha]$. Consequently, the claim follows by Lemma 2.8 (a). □

Clearly, if a set A has a $\surd$-construction of length γ, then $|A| \leq \gamma < |A|^+$. The next lemma can be used to obtain constructions of length exactly $|A|$.

Lemma 2.11. *Let ζ be a $\surd$-construction of A over U. If $\mathrm{loc}(\zeta)^{\mathrm{reg}} \leq |A|$, then A has a $\surd$-construction over U of length $|A|$.*

Proof. Let $(a_\alpha)_{\alpha<\kappa}$ be an enumeration of A of length $\kappa := |A|$. By induction on $\alpha < \kappa$, we can use Lemma 2.4 to choose subsets $B_\alpha \subseteq A$ of size $|B_\alpha| < \mathrm{loc}(\zeta)^{\mathrm{reg}}$ such that $a_\alpha \in B_\alpha$ and B_α is ζ-closed over $U \cup B[<\alpha]$. By Lemma 2.3, each set $B_\alpha \setminus B[<\alpha]$ has a $\surd$-construction $\xi_\alpha = (b_i^\alpha)_{i<\gamma_\alpha}$ of length

$$\gamma_\alpha < |B_\alpha|^+ \leq \mathrm{loc}(\zeta)^{\mathrm{reg}} \leq \kappa\,.$$

We have seen in the proof of Lemma 2.8 (a) that the concetanation of these $\surd$-constructions is a $\surd$-construction of $B[<\kappa] = A$ over U of length $\sum_{\alpha<\kappa} \gamma_\alpha = \kappa$. □

3. *Prime models*

Using $\surd$-constructions we can generalise the results of Section E3.4 to arbitrary isolation relations. One important property of the relation $\overset{\mathrm{at}}{\surd}$ that is not captured by the notion of an isolation relation is the fact that every model realises all isolated types. When generalising certain results about $\overset{\mathrm{at}}{\surd}$-constructions we have to require this property separately. This leads to the notion of $\surd$-saturation.

Definition 3.1. Let $\surd$ be an isolation relation, κ a cardinal, and $A, U \subseteq \mathbb{M}$ sets.

(a) A is $\surd$-*κ-saturated* if, for all sets $C \subseteq A$ of size $|C| < \kappa$ and every finite set $B \subseteq \mathbb{M}$ with $B \surd C$, there is some set $B' \subseteq A$ with $B' \equiv_C B$.

(b) A is $\surd$-*κ-prime* over U if it is $\surd$-κ-saturated and, for every $\surd$-κ-saturated set $B \supseteq U$, there exists an automorphism $\pi \in \mathrm{Aut}\,\mathbb{M}_U$ with $\pi[A] \subseteq B$.

Our aim is to prove that $\surd$-prime models are unique, up to isomorphism. We start by setting up the required back-and-forth machinery.

Lemma 3.2. *Let A be $\surd$-κ-saturated and $U \subseteq A$ a set of size $|U| < \kappa$. For every $\surd$-constructible set B over U of size $|B| \leq \kappa$, there exists some set $B' \subseteq A$ with $B' \equiv_U B$.*

Proof. Let $\zeta = (b_\alpha)_{\alpha<\gamma}$ be a $\surd$-construction of B over U of length $\gamma \leq \kappa$. We inductively construct a sequence $(b'_\alpha)_{\alpha<\gamma}$ in A such that

$$b'[<\alpha] \equiv_U b[<\alpha]\,, \quad \text{for all } \alpha \leq \gamma\,.$$

Suppose that we have already defined b'_α, for all $\alpha < \beta$. Fix an element c such that

$$b_\beta b[<\beta] \equiv_U cb'[<\beta]\,.$$

Then

$$b_\beta \surd Ub[<\beta] \quad \text{implies} \quad c \surd Ub'[<\beta]\,.$$

Since A is $\surd$-κ-saturated and $|U| \oplus |\beta| < \kappa$, we can therefore find an element $b'_\beta \in A$ with $b'_\beta \equiv_{Ub'[<\beta]} c$. It follows that

$$b'_\beta b'[<\beta] \equiv_U cb'[<\beta] \equiv_U b_\beta b[<\beta]\,. \qquad \square$$

Lemma 3.3. *Let $\surd$ be an isolation relation, $\kappa \geq \mathrm{lbm}(\surd)$ an uncountable cardinal, $M, N \subseteq \mathbb{M}$ $\surd$-κ-saturated, $U \subseteq \mathbb{M}$ a set of size $|U| < \kappa$, and let ξ and ζ be $\surd$-constructions of, respectively, M and N over U such that $\mathrm{loc}(\xi)^{\mathrm{reg}}, \mathrm{loc}(\zeta)^{\mathrm{reg}} \leq \mathrm{lbm}(\surd)$. Then*

$$H : M \cong^\kappa_{\mathrm{iso}} N\,,$$

where H is the set of all elementary maps $f : A \to B$ such that

- *$A \subseteq M$ is a ξ-closed set of size $|A| < \kappa$,*
- *$B \subseteq N$ is a ζ-closed set of size $|B| < \kappa$, and*

- $f \upharpoonright U = \mathrm{id}_U$.

Proof. Set $\lambda := \mathrm{loc}(\xi)^{\mathrm{reg}} \oplus \mathrm{loc}(\zeta)^{\mathrm{reg}}$. By symmetry it is sufficient to check the forth property. Consider a map $f : A \to B$ in H and an element $c \in M$. We will construct an increasing chain of elementary maps $g_n : C_n \to D_n$, for $n < \omega$, such that

- $f \subseteq g_0$ and $c \in \mathrm{dom}(g_0)$,
- $|C_n \setminus C_{n-1}|, |D_n \setminus D_{n-1}| < \lambda$ (where $C_{-1} := A$ and $D_{-1} := B$),
- C_n is $\surd$-constructible over $U \cup C_{n-1}$,
- D_n is $\surd$-constructible over $U \cup D_{n-1}$,
- C_n is ξ-closed over U, for even $n < \omega$, and
- D_n is ζ-closed over U, for odd $n < \omega$.

Then we can set $g := \bigcup_{n<\omega} g_n$. By Lemma 2.4 (a),

$$\mathrm{dom}(g) = \bigcup_{n<\omega} C_n = \bigcup_{n<\omega} C_{2n}$$

is ξ-closed and $\mathrm{rng}(g) = \bigcup_{n<\omega} D_{2n+1}$ is ζ-closed. Furthermore,

$$|\mathrm{dom}(g)| < |A|^+ \oplus \lambda \oplus \aleph_1 \leq \kappa\,.$$

Hence, $g \in H$.

It remains to construct $(g_n)_n$. By Lemma 2.4 (d), we can find a ξ-closed set $C'_0 \subseteq M$ of size $|C'_0| < \lambda$ with $c \in C'_0$. Choose a set $D'_0 \subseteq \mathbb{M}$ with

$$AC'_0 \equiv_U BD'_0\,.$$

Note that it follows by Proposition 2.5 (b) that C'_0 is $\surd$-constructible over $U \cup A$. Consequently, D'_0 is $\surd$-constructible over $U \cup B$. Since N is $\surd$-κ-saturated, we can therefore use Lemma 3.2 to find a set $D''_0 \subseteq N$ with

$$D''_0 \equiv_{UB} D'_0\,.$$

Consequently, $AC'_0 \equiv_U BD''_0$. Let $g_0 : A \cup C'_0 \to B \cup D''_0$ be the corresponding extension of f.

For the successor step, suppose that we have already defined $g_n : C_n \to D_n$. First, consider the case where n is even. As λ is regular, we can use Lemma 2.4 (d) to find a ζ-closed set $D'_{n+1} \subseteq N$ of size $|D'_{n+1}| < \lambda$ with $D_n \setminus D_{n-1} \subseteq D'_{n+1}$. Choose a set $C'_{n+1} \subseteq \mathbb{M}$ with

$$C_n C'_{n+1} \equiv_U D_n D'_{n+1} .$$

By Proposition 2.5 (b), D'_{n+1} is $\surd$-constructible over $U \cup D_{n-1}$. According to Lemma 2.9 this implies that

$$D'_{n+1} \surd_{UD_{n-1}} UD_{n-1} .$$

Since $|D_n \setminus D_{n-1}| < \lambda \leq \mathrm{lbm}(\surd)$, it follows that

$$D'_{n+1} \surd_{UD_n} UD_n .$$

Applying Lemma 2.9 again, we see that the set D'_{n+1} is $\surd$-constructible over $U \cup D_n$. By invariance, it follows that C'_{n+1} is $\surd$-constructible over $U \cup C_n$. Hence, we can use Lemma 3.2 to find a set $C''_{n+1} \subseteq M$ with

$$C''_{n+1} \equiv_{UC_n} C'_{n+1} .$$

Consequently, $C_n C''_{n+1} \equiv_U D_n D'_{n+1}$. Let $g_{n+1} : C_n \cup C''_{n+1} \to D_n \cup D'_{n+1}$ be the corresponding extension of g_n.

If n is odd, we proceed similarly by choosing a ξ-closed set $C_{n+1} \subseteq M$ containing $C_n \setminus C_{n-1}$. □

With the back-and-forth machinery in place we can construct isomorphisms between $\surd$-κ-saturated sets.

Proposition 3.4. *Let $\surd$ be an isolation relation and $A, B \subseteq \mathbb{M}$ $\surd$-κ-saturated sets that are $\surd$-constructible over some set $U \subseteq A$ of size $|U| < \kappa$.*

(a) *If $|A| \leq \kappa$, there exists an automorphism $\pi : \mathrm{Aut}\,\mathbb{M}_U$ such that $\pi[A] \subseteq B$.*

(b) *If* $|A|, |B| \leq \kappa$, *there exists an automorphism* $\pi : \mathrm{Aut}\,\mathbb{M}_U$ *such that* $\pi[A] = B$.

Proof. By Lemma 3.3, we have $H : A \equiv^{\kappa}_{\mathrm{iso}} B$. Consequently, we can use Lemma C4.4.10 (a) or (b) to find an elementary embedding $h : A \to B$ such that $h \upharpoonright U = \mathrm{id}_U$ and, in case $|B| \leq \kappa$, such that h is surjective. As $\mathbb{M}$ is strongly κ^+-homogeneous, we can extend h to the desired automorphism π. □

Corollary 3.5. *Let* U *be a set of size* $|U| < \kappa$. *Up to isomorphism, there is at most one* $\surd$*-*κ*-saturated set* A *of size* $|A| \leq \kappa$ *that is* $\surd$*-constructible over* U.

Proof. If A and B are $\surd$-κ-saturated sets of size at most κ that are $\surd$-constructible over U, we can use Proposition 3.4 (b) to find an automorphism mapping A to B. □

Corollary 3.6. *Let* A *be a* $\surd$*-*κ*-saturated set of size* $|A| \leq \kappa$ *that is* $\surd$*-constructible over a set* $U \subseteq A$ *of size* $|U| < |A|$. *Then* A *is* $\surd$*-*κ*-prime over* U.

Proof. If $B \supseteq U$ is $\surd$-κ-saturated and $\surd$-constructible over U, we can use Proposition 3.4 (a) to find the desired automorphism mapping A to a subset of B. □

Having proved that $\surd$-κ-saturated sets are unique, it remains to show that such sets exists. For the relation $\overset{\mathrm{at}}{\surd}$ we will prove below that every model is $\overset{\mathrm{at}}{\surd}$-$\kappa$-saturated. For isolation relations $\surd$ that satisfy the extension axiom, we have the following lemma.

Lemma 3.7. *Let* $\surd$ *be an isolation relation satisfying* (EXT) *such that* $\mathrm{lbm}(\surd) \geq \aleph_0$, *and let* κ *be an infinite cardinal. For every set* U, *there exists some* $\surd$*-*κ*-saturated set* A *that is* $\surd$*-constructible over* U.

Proof. We construct an increasing sequence $(A_\alpha)_{\alpha<\kappa^+}$ of sets $A_\alpha \subseteq \mathbb{M}$ such that each A_α is $\surd$-constructible over $U \cup A[<\alpha]$ and, for every

$\alpha < \kappa^+$, every set $C \subseteq A[<\alpha]$ of size $|C| < \kappa$, and every finite set $B \subseteq \mathbb{M}$, there is some set $B' \subseteq A[\leq\alpha]$ with $B' \equiv_C B$.

We start with $A_0 := U$. For the inductive step, suppose that we have already defined A_α, for all $\alpha < \beta$. To find A_β, we fix an enumeration $\langle C_\alpha, \mathfrak{p}_\alpha \rangle_{\alpha<\gamma}$ of all pairs $\langle C, \mathfrak{p} \rangle$ where $C \subseteq A[<\beta]$ has size $|C| < \kappa$ and $\mathfrak{p} \in S^{<\omega}(C)$ is a type such that $\bar{a} \mathrel{\sqrt{}} C$, for every realisation $\bar{a}$ of $\mathfrak{p}$. By induction on $\alpha < \gamma$, we choose finite tuples $\bar{b}_\alpha \in M^{<\omega}$ as follows. Let $\bar{b}'_\alpha$ be a realisation of $\mathfrak{p}_\alpha$. Then $\bar{b}'_\alpha \mathrel{\sqrt{}} C$ implies, by (EXT), that there is some tuple $\bar{b}_\alpha \equiv_C \bar{b}'_\alpha$ with

$$\bar{b}_\alpha \mathrel{\sqrt{}_C} A[<\beta]\bar{b}[<\alpha] \,.$$

We set $A_\beta := A[<\beta] \cup \bar{b}[<\gamma]$. Note that $(\bar{b}_\alpha \setminus \bar{b}[<\alpha])_{\alpha<\gamma}$ is a $\sqrt{}$-stratification of A_β over $A[<\beta]$. Since $\mathrm{lbm}(\sqrt{}) \geq \aleph_0$, it follows by Corollary 2.10 that A_β is $\sqrt{}$-constructible over $A[<\beta]$.

Having constructed $(A_\alpha)_{\alpha<\kappa^+}$, we claim that the union $A := A[<\kappa^+]$ has the desired properties. Since every set A_α is $\sqrt{}$-constructible over $U \cup A[<\alpha]$, it follows by Lemma 2.8 (a) that A is $\sqrt{}$-constructible over U. To show that it is also $\sqrt{}$-κ-saturated, let $C \subseteq A$ be a set of size $|C| < \kappa$ and let $\bar{b} \subseteq \mathbb{M}^{<\omega}$ be a tuple with $\bar{b} \mathrel{\sqrt{}} C$. As κ^+ is regular, there is some index $\alpha < \kappa^+$ such that $C \subseteq A[<\alpha]$. Since the pair $\langle C, \mathrm{tp}(\bar{b}/C) \rangle$ appears in the sequence used in the construction of A_α, it follows that there is some $\bar{b}' \subseteq A_\alpha$ with $\bar{b}' \equiv_C \bar{b}$. ☐

4. $\sqrt[\mathrm{at}]{}$-constructible models

In this section we take a closer look at $\sqrt[\mathrm{at}]{}$-constructible sets. We have already seen in Section E3.4 that a model which is $\sqrt[\mathrm{at}]{}$-constructible over some set U is atomic over U, prime over U, and unique up to isomorphisms over U. These facts also follow from the general results we have have derived in the present chapter once we have shown that models are always $\sqrt[\mathrm{at}]{}$-saturated.

Proposition 4.1. *Every model $M \subseteq \mathbb{M}$ is $\sqrt[\mathrm{at}]{}$-$|M|^+$-saturated.*

Proof. Suppose that $\bar{a} \sqrt[\mathrm{at}]{} U$ where $\bar{a} \in \mathbb{M}^{<\omega}$ and $U \subseteq M$. Let $\varphi(\bar{x})$ be a formula over U isolating $\operatorname{tp}(\bar{a}/U)$. Then

$$\mathbb{M} \vDash \varphi(\bar{a}) \quad \text{implies} \quad \mathbb{M} \vDash \exists \bar{x} \varphi(\bar{x}) \quad \text{implies} \quad \mathfrak{M} \vDash \exists \bar{x} \varphi(\bar{x}) .$$

Hence, there is some tuple $\bar{a}' \in M^{<\omega}$ with $\mathfrak{M} \vDash \varphi(\bar{a}')$. By choice of φ this implies that $\operatorname{tp}(\bar{a}'/U) = \operatorname{tp}(\bar{a}/U)$. ☐

Example. Consider the theory T of the structure $\mathfrak{C} := \langle 2^\omega, (P_n)_{n<\omega} \rangle$ where

$$P_n := \{ \alpha \in 2^\omega \mid \alpha(n) = 1 \} .$$

For this theory, we have

$$\begin{aligned} & A \sqrt[\mathrm{d}]{}_U B && \text{iff} \quad A \cap B \subseteq U , \\ \text{and} \quad & A \sqrt[\mathrm{at}]{}_U B && \text{iff} \quad A \subseteq U . \end{aligned}$$

Furthermore,

$$|S^{<\omega}(U)| \leq |U| \oplus |S^{<\omega}(\varnothing)| = |U| \oplus 2^{\aleph_0} ,$$

since

$$\operatorname{tp}(\bar{a}/\bar{a} \cap U) \vDash \operatorname{tp}(\bar{a}/U) , \quad \text{for all } \bar{a}, U \subseteq \mathbb{M} .$$

Consequently, the theory T is superstable with $\operatorname{st}(T) = 2^{\aleph_0}$, and a set A is $\sqrt[\mathrm{at}]{}$-constructible over U if, and only if, $A \subseteq U$. In particular, no model of T is $\sqrt[\mathrm{at}]{}$-constructible over $\varnothing$.

For stable theories one can show that subsets of $\sqrt[\mathrm{at}]{}$-constructible sets are again $\sqrt[\mathrm{at}]{}$-constructible. We start with three technical lemmas.

Lemma 4.2. *Let $\downarrow$ be a symmetric preforking relation. If $A \downarrow_U B$ then, for every set $D \subseteq \mathbb{M}$, there exists a set $C \subseteq A$ of size $|C| < \mathrm{loc}(\downarrow) \oplus |D|^+$ such that*

$$A \downarrow_{UC} BD .$$

Proof. By right locality, we can choose a set $C \subseteq A$ of size

$$|C| < \mathrm{loc}(\downarrow) \oplus |D|^+ \quad \text{such that} \quad D \downarrow_{UBC} UBA .$$

It follows that $D \downarrow_{UBC} A$. Furthermore, $A \downarrow_U B$ implies $B \downarrow_{UC} A$. By transitivity it follows that $BD \downarrow_{UC} A$. □

Lemma 4.3. *Let $\downarrow$ be a symmetric preforking relation, $\sqrt{}$ an isolation relation, and let $\kappa \geq \mathrm{loc}(\downarrow)$ a be regular cardinal. Let ζ be a $\sqrt{}$-construction of some set A over U with $\mathrm{loc}(\zeta) \leq \kappa$ and let $C \subseteq A$ be a subset. For every ζ-closed set $B \subseteq A$ with*

$$B \downarrow_{U(C \cap B)} C$$

and every set $D \subseteq A$ of size $|D| < \kappa$, there exists a ζ-closed set $B_+ \supseteq B \cup D$ such that

$$|B_+ \setminus B| < \kappa \quad \textit{and} \quad B_+ \downarrow_{U(C \cap B_+)} C .$$

Proof. Let $(d_\alpha)_{\alpha<\gamma}$ be an enumeration of D. Starting with $B_0 := B$, we construct an increasing chain $(B_\alpha)_{\alpha<\gamma}$ of ζ-closed sets such that

$$d_\alpha \in B_{\alpha+1} , \quad |B_\alpha \setminus B| < \kappa , \quad \text{and} \quad B_\alpha \downarrow_{U(C \cap B_\alpha)} C .$$

Then we can set $B_+ := \bigcup_{\alpha<\gamma} B_\alpha$.

For the successor step, suppose that we have already defined B_α. By Lemma 2.4 (c), there exists a ζ-closed set $Z \subseteq A$ of size $|Z| < \kappa$ containing d_α. We can choose a set $W \subseteq U \cup B_\alpha \cup C$ of size

$$|W| < \mathrm{loc}(\downarrow) \oplus |Z|^+ \leq \kappa \quad \text{such that} \quad Z \downarrow_W UB_\alpha C .$$

It follows that $Z \downfree_{UB_\alpha(W \cap C)} C$. Furthermore,

$$B_\alpha \downfree_{U(C \cap B_\alpha)} C \quad \text{implies} \quad B_\alpha \downfree_{U(C \cap B_\alpha)(C \cap Z)(C \cap W)} C\,.$$

With transitivity it follows that

$$B_\alpha Z(W \cap C) \downfree_{U(C \cap B_\alpha)(C \cap Z)(C \cap W)} C\,.$$

Setting $B_{\alpha+1} := B_\alpha \cup Z \cup (W \cap C)$, we have

$$|B_{\alpha+1} \setminus B| \leq |B_\alpha \setminus B| \oplus |Z| \oplus |W \cap C| < \kappa\,,$$

as desired.

For the limit step, let δ be a limit ordinal and suppose that B_α is already defined for all $\alpha < \delta$. Then we set $B_\delta := \bigcup_{\alpha<\delta} B_\alpha$. □

Lemma 4.4. *Let T be a stable theory.*

$$A \overset{\text{at}}{\bigvee} UB \quad \textit{and} \quad A \downfree^{\text{f}}_U B \quad \textit{implies} \quad A \overset{\text{at}}{\bigvee} U\,.$$

Proof. Let $\bar{a} \subseteq A$ be finite. We have to show that $\operatorname{tp}(\bar{a}/U)$ is isolated. Since $\bar{a} \overset{\text{at}}{\bigvee} UB$, the type $\mathfrak{p} := \operatorname{tp}(\bar{a}/UB)$ is isolated, that is, the set $\{\mathfrak{p}\}$ is open in $\mathfrak{S}^{\bar{s}}(UB)$. By assumption, $\mathfrak{p}$ does not fork over U. Hence, $\mathfrak{p} \in F^{\bar{s}}(UB/U)$ in the notation of the Open Mapping Theorem. According to that theorem, the restriction map $\mathfrak{F}^{\bar{s}}(UB/U) \to \mathfrak{S}^{\bar{s}}(U)$ is open. Consequently, the image $\{\mathfrak{p}|_U\}$ is open in $\mathfrak{S}^{\bar{s}}(U)$ and $\mathfrak{p}|_U$ is isolated. □

Using these lemmas, we can show that subsets of constructible sets are again constructible.

Theorem 4.5. *Let T be a stable theory with $\operatorname{fc}(\downfree^{\text{f}}) \leq \aleph_1$ and let A be a $\overset{\text{at}}{\bigvee}$-constructible set over U. Every subset $C \subseteq A$ is also $\overset{\text{at}}{\bigvee}$-constructible over U.*

Proof. Let ζ be a $\overset{\text{at}}{\bigvee}$-construction of A over U. Since

$$\operatorname{loc}(\zeta) \leq \operatorname{loc}_0(\overset{\text{at}}{\bigvee}) \leq \aleph_0$$

$$\text{and} \quad \mathrm{loc}(\downarrow^{\mathrm{f}}) \le \mathrm{loc}_0(\downarrow^{!})^{\mathrm{reg}} = \mathrm{fc}(\downarrow^{\mathrm{f}})^{\mathrm{reg}} \le \aleph_1\,,$$

we can use Lemma 4.3 to inductively construct an increasing sequence $(B_\alpha)_{\alpha<\gamma}$ of sets $B_\alpha \subseteq A$ such that

- $B_0 = \varnothing$ and $B[<\gamma] = A$,
- each B_α is ζ-closed over $U \cup B[<\alpha]$,
- $|B_\alpha \setminus B[<\alpha]| \le \aleph_0$,
- $B[<\alpha] \downarrow^{\mathrm{f}}_{U(C\cap B[<\alpha])} C$.

Set $C_\alpha := C\cap(B_\alpha \setminus B[<\alpha])$, for $\alpha < \gamma$. Then $|C_\alpha| \le \aleph_0 = \mathrm{lbm}(\sqrt[\mathrm{at}]{})$ and, by Corollary 2.10, it is sufficient to prove that $(C_\alpha)_{\alpha<\gamma}$ is a $\sqrt[\mathrm{at}]{}$-stratification of C over U.

Hence, let $\alpha < \gamma$. Since B_α is ζ-closed over $U \cup B[<\alpha]$, it follows by Lemma 2.8 (a) that $(B_\alpha)_{\alpha<\gamma}$ is a $\sqrt[\mathrm{at}]{}$-stratification of A over U. Hence,

$$B_\alpha \sqrt[\mathrm{at}]{}\; UB[<\alpha]\,, \quad \text{which implies that} \quad C_\alpha \sqrt[\mathrm{at}]{}\; UB[<\alpha]\,.$$

Since $C_\alpha \downarrow^{\mathrm{f}}_{UC[<\alpha]} B[<\alpha]$, Lemma 4.4 implies that $C_\alpha \sqrt[\mathrm{at}]{}\; UC[<\alpha]$, as desired. □

Corollary 4.6. *Let T be a stable theory with* $\mathrm{fc}(\downarrow^{\mathrm{f}}) \le \aleph_1$ *and let $\mathfrak{M}$ be a $\sqrt[\mathrm{at}]{}$-constructible model over U. Then $\mathfrak{M}$ is the unique prime model of T over U.*

Proof. By Corollary 3.6 $\mathfrak{M}$ is prime over U. For uniqueness, let $\mathfrak{N}$ be another prime model over U. Then there exists an elementary embedding $h : \mathfrak{N} \to \mathfrak{M}$. By Theorem 4.5, $\mathfrak{N}$ is $\sqrt[\mathrm{at}]{}$-constructible over U. Hence, it follows by Corollary 3.5 that $\mathfrak{N} \cong \mathfrak{M}$. □

Finally, let us take a look at $\sqrt[\mathrm{at}]{}$-constructible sets in totally transcendental theories. We will prove below that a model of such a theory is prime if, and only if, it is $\sqrt[\mathrm{at}]{}$-constructible. We also give a characterisation in terms of the length of indiscernible sequences. One direction is contained in the following proposition.

Proposition 4.7. *Let T be totally transcendental and let A be $\sqrt[\mathrm{at}]{}$-constructible over U. Every indiscernible sequence over U that is contained in A is countable.*

Proof. Let $\zeta = (a_\alpha)_{\alpha<\gamma}$ be a $\sqrt[\mathrm{at}]{}$-construction of A over U and suppose that A contains an uncountable indiscernible sequence $(c_\alpha)_{\alpha<\omega_1}$ over U. Note that $\operatorname{loc}(\zeta) \leq \operatorname{loc}_0(\sqrt[\mathrm{at}]{}) = \aleph_0$. Hence, we can use Lemma 2.4 (c) to fix, for every $\alpha < \omega_1$, a finite ζ-closed set $B_\alpha \subseteq A$ over U with $c_\alpha \in B_\alpha$. By Lemma E5.3.11 there exists, for every $\alpha < \omega_1$, an ordinal $\delta_\alpha < \omega_1$ such that $(c_\beta)_{\delta_\alpha \leq \beta < \omega_1}$ is indiscernible over $U \cup B[<\alpha]$. For $\alpha < \omega_1$, set

$$\gamma^\alpha_0 := \alpha\,, \quad \gamma^\alpha_{n+1} := \delta_{\gamma^\alpha_n}\,, \quad \text{and} \quad \gamma^\alpha_* := \sup_{n<\omega} \gamma^\alpha_n\,.$$

Note that $\gamma^\alpha_* < \omega_1$, for all $\alpha < \omega_1$. Since $(c_\beta)_{\beta \geq \gamma^\alpha_*}$ is indiscernible over $U \cup B[<\gamma^\alpha_n]$, for all $n < \omega$, it follows that it is indiscernible over

$$\bigcup_{n<\omega} (U \cup B[<\gamma^\alpha_n]) = U \cup B[<\gamma^\alpha_*]\,.$$

Set

$$D_0 := \{\, \gamma^\alpha_* \mid \alpha < \omega_1 \,\} \quad \text{and} \quad D := \{\, \sup I \mid I \subseteq D_0\,,\ \sup I < \omega_1 \,\}\,.$$

The set D is closed by definition and it is unbounded since $\gamma^\alpha_* \geq \alpha$, for all α. By Lemma A4.6.8 (a) it follows in particular that D is stationary. Furthermore, for every $\delta \in D$, the suffix $(c_\alpha)_{\delta \leq \alpha < \omega_1}$ is indiscernible over $U \cup B[<\delta]$. By Lemma 2.4 and Proposition 2.5, $B[<\delta]$ is ζ-closed and ζ is a $\sqrt[\mathrm{at}]{}$-construction over $U \cup B[<\delta]$. By Corollary 2.6, it follows that

$$A \sqrt[\mathrm{at}]{}\, UB[<\delta]\,, \quad \text{which implies that} \quad c_\delta \sqrt[\mathrm{at}]{}\, UB[<\delta]\,.$$

Since $\operatorname{loc}(\sqrt[\mathrm{at}]{}) = \aleph_0$, we can fix, for every $\delta \in D$, a finite set $W_\delta \subseteq U \cup B[<\delta]$ such that

$$c_\delta \sqrt[\mathrm{at}]{}_{W_\delta}\, UB[<\delta]\,.$$

By the Theorem of Fodor, there exist an index $\delta_0 \in D$ and a stationary set $E \subseteq D$ such that $W_\varepsilon \subseteq B[<\delta_0]$, for all $\varepsilon \in E$. Fix two ordinals $\varepsilon < \eta$ in E. By indiscernibility, we have $c_\varepsilon \equiv_{U \cup B[<\delta_0]} c_\eta$. But $W_\eta \subseteq U \cup B[<\delta_0]$ implies that

$$\mathrm{tp}(c_\eta/UB[<\delta_0]) \vDash \mathrm{tp}(c_\eta/UB[<\delta_0]c_\varepsilon)\,.$$

Hence, $c_\eta \neq c_\varepsilon$ implies that $c_\varepsilon \neq c_\varepsilon$. A contradiction. □

For the converse, we need several lemmas.

Definition 4.8. Let $\mathfrak{M}$ be a model. A set $A \subseteq M$ is *invariant* over $U \subseteq M$ if, for all finite tuples $\bar{a}, \bar{b} \in M^{<\omega}$,

$$\bar{a} \equiv_U \bar{b} \quad \text{implies} \quad \bar{a} \subseteq A \Leftrightarrow \bar{b} \subseteq A\,.$$

Lemma 4.9. *Let T be a totally transcendental theory, $\mathfrak{M}$ a model of T, and let $U \subseteq M$ be a set such that $M \sqrt[\mathrm{at}]{} U$. Then $M \sqrt[\mathrm{at}]{} UC$ for every set $C \subseteq M$ that is invariant over U.*

Proof. Let $\bar{a} \in M^{<\omega}$. Then the type $\mathrm{tp}(\bar{a}/U)$ is isolated by some formula $\varphi(\bar{x})$ over U. We have seen in Lemma E3.4.12 that the isolated types in $\mathfrak{S}^{<\omega}(UC)$ are dense. Consequently, we can find some isolated type

$$\mathfrak{p} \in \langle \varphi \rangle_{\mathfrak{S}^{<\omega}(UC)}\,.$$

Let $\psi(\bar{x}, \bar{d})$ be a formula over U isolating $\mathfrak{p}$ with $\bar{d} \subseteq C$ and fix a tuple $\bar{b}$ realising $\mathfrak{p}$. Then $\mathfrak{M} \vDash \varphi(\bar{b})$ implies that $\bar{a} \equiv_U \bar{b}$. Consequently, we can find some tuple $\bar{c}$ with $\bar{a}\bar{c} \equiv_U \bar{b}\bar{d}$. Then $\bar{d} \subseteq C$ implies $\bar{c} \subseteq C$ by invariance of C over U. Furthermore,

$$\mathfrak{M} \vDash \psi(\bar{b}; \bar{d}) \quad \text{implies} \quad \mathfrak{M} \vDash \psi(\bar{a}; \bar{c})\,.$$

We claim that $\psi(\bar{x}; \bar{c})$ isolates $\mathrm{tp}(\bar{a}/UC)$. Let $\vartheta(\bar{x}; \bar{c}') \in \mathrm{tp}(\bar{a}/UC)$. Fix some tuple $\bar{d}'$ such that

$$\bar{a}\bar{c}\bar{c}' \equiv_U \bar{b}\bar{d}\bar{d}'\,.$$

By invariance of C over U, it follows that $\bar{d}' \subseteq U \cup C$. Since $\psi(\bar{x};\bar{d})$ isolates $\mathrm{tp}(\bar{b}/UC)$ and $\mathfrak{M} \vDash \vartheta(\bar{a};\bar{c}')$ implies $\mathfrak{M} \vDash \vartheta(\bar{b};\bar{d}')$, we have

$$T(U \cup \bar{d}\bar{d}') \vDash \psi(\bar{x};\bar{d}) \to \vartheta(\bar{x};\bar{d}')\,.$$

Consequently,

$$T(U \cup \bar{c}\bar{c}') \vDash \psi(\bar{x};\bar{c}) \to \vartheta(\bar{x};\bar{c}')\,,$$

as desired. □

Lemma 4.10. *Let T be a totally transcendental theory, $\mathfrak{M}, \mathfrak{N}$ models of T, $\mathfrak{p} \in S^1(U)$ a type over $U \subseteq M$, and $U \subseteq C \subseteq M$ a set that is invariant over U. If $M \overset{\mathrm{at}}{\bigvee} U$ and M does not contain an uncountable indiscernible sequence over U, then every elementary map $f : C \to N$ can be extended to an elementary map $C \cup \mathfrak{p}^{\mathfrak{M}} \to N$.*

Proof. We prove the claim by induction on $\alpha := \mathrm{rk}_{\mathrm{M}}(\mathfrak{p})$. If $\alpha = 0$, then $\mathfrak{p}^{\mathfrak{M}} \subseteq \mathrm{acl}(U)$ and the claim holds trivially.

For the inductive step, suppose that we have proved the statement already for all types of Morley rank less than α. Let $I \subseteq \mathfrak{p}^{\mathfrak{M}}$ be a maximal set such that

$$a \mathop{\downarrow}^{\mathrm{f}}_{U} I \setminus \{a\}\,, \quad \text{for all } a \in I\,.$$

According to Lemma G1.5.6, we can partition I into finitely many totally indiscernible sequences. By assumption, each of them is countable. Hence, so is I. Let $(a_n)_{n<\omega}$ be an enumeration of I and set

$$C_n := C \cup \{\, c \in M \mid \mathrm{rk}_{\mathrm{M}}(c/Ua[<n]) < \alpha \,\}\,, \quad \text{for } n < \omega\,.$$

Then $(C_n)_{n<\omega}$ forms an increasing chain starting with $C_0 = C$. Furthermore, every element $b \in \mathfrak{p}^{\mathfrak{M}} \setminus \bigcup_{n<\omega} C_n$ satisfies

$$\mathrm{rk}_{\mathrm{M}}(b/UI) = \mathrm{rk}_{\mathrm{M}}(b/U)\,.$$

By Corollary G1.4.8 (a), this implies that $b \mathop{\downarrow}^{\mathrm{f}}_{U} I$. By maximality of I, it therefore follows that

$$\mathfrak{p}^{\mathfrak{M}} \subseteq \bigcup_{n<\omega} C_n \,.$$

Note that $M \sqrt[\mathrm{at}]{} U$ implies $M \sqrt[\mathrm{at}]{} Ua[<n]$ since $\mathrm{lbm}(\sqrt[\mathrm{at}]{}) \geq \aleph_0$. As C_n is invariant over $U \cup a[<n]$, it follows by Lemma 4.9 that $M \sqrt[\mathrm{at}]{} C_n$. We prove by induction on n, that f can be extended to C_n. For $n = 0$, we have $C_0 = C$ and there is nothing to do. For the inductive step, suppose that we have already extended f to C_n. We first extend f to $C_n \cup \{a_n\}$. Since $M \sqrt[\mathrm{at}]{} C_n$, there exists some formula $\varphi(x;\bar{c})$ with parameters $\bar{c} \subseteq C_n$ isolating $\mathrm{tp}(a_n/C_n)$. Then

$$\mathfrak{M} \vDash \exists x \varphi(x;\bar{c}) \quad \text{implies} \quad \mathfrak{N} \vDash \exists x \varphi(x; f(\bar{c}))\,.$$

Hence, there exists some element $b \in N$ such that we can extend f by setting $f(a_n) := b$. Having done so we obtain the desired extension of f to C_{n+1} by applying the inductive hypothesis on α to C_n (for U) and $C_n \cup \{a_n\}$ (for C). □

We obtain the following characterisation of prime models of totally transcendental theories.

Theorem 4.11. *Let T be a totally transcendental theory and U a set of parameters.*

(a) *T has a unique prime model over U.*

(b) *For a model $\mathfrak{M}$ of T, the following statements are equivalent:*

  (1) *$\mathfrak{M}$ is prime over U*

  (2) *$\mathfrak{M}$ is $\sqrt[\mathrm{at}]{}$-constructible over U.*

  (3) *$\mathfrak{M}$ is atomic over U and M does not contain an uncountable indiscernible sequence over U.*

Proof. (a) The existence and uniqueness of a prime model were already proved in Theorem E3.4.14.

(b) (2) ⇒ (1) was shown in Proposition E3.4.3.

(1) ⇒ (3) Suppose that $\mathfrak{M}$ is prime over U. By Proposition E3.4.13, there exists a model $\mathfrak{A}$ that is $\overset{\mathrm{at}}{\surd}$-constructible over U. As $\mathfrak{M}$ is prime over U, we can find an elementary embedding $h : \mathfrak{M} \to \mathfrak{A}$ fixing U. By Corollary 2.6, we have $A \overset{\mathrm{at}}{\surd} U$. Since $h[M] \subseteq A$, it follows by invariance that $M \overset{\mathrm{at}}{\surd} U$. Thus, $\mathfrak{M}$ is atomic over U.

Furthermore, every indiscernible sequence in M is mapped by h to an indiscernible sequence in A. We have seen in Proposition 4.7 that A only contains countable indiscernible sequences over U. Hence, so does M.

(3) ⇒ (2) Let $\mathfrak{N}$ be an arbitrary model of $T(U)$. To find the desired elementary embedding $\mathfrak{M}_U \to \mathfrak{N}$, we choose a maximal elementary map $f : C \to N$ with a domain $C \subseteq M$ that is invariant over U. We claim that $C = M$. For a contradiction, suppose otherwise. Then there is some element $a \in M \setminus C$. By Lemma 4.9, we have $M \overset{\mathrm{at}}{\surd} C$. Hence, setting $\mathfrak{p} := \mathrm{tp}(a/C)$ we can use Lemma 4.10 to extend f to an elementary map $C \cup \mathfrak{p}^{\mathfrak{M}} \to N$. This contradicts the maximality of f. □

5. *Strongly independent stratifications*

In contrast to $\overset{\mathrm{at}}{\surd}$, the isolation relation $\downarrow^{!}$ admits arbitrarily large constructible sets. Consequently, the length of $\downarrow^{!}$-constructions is unbounded. But we will show that there are $\downarrow^{!}$-stratifications of bounded length such that every enumeration refining them is a $\downarrow^{!}$-construction.

Unique free extensions

As a preliminary step, we start with computing $\mathrm{fc}(\downarrow^{!})$. We employ the following characterisation of $\downarrow^{!}$ which is a variant of Lemma G1.3.6.

Definition 5.1. For sets $A, B, U \subseteq \mathbb{M}$, we define

$$A \perp^{\mathrm{do}}_U B \quad :\text{iff} \quad \text{every relation that is definable over both } A \cup U \text{ and } B \cup U \text{ is already definable over } U\,.$$

If $A \perp^{\mathrm{do}}_U B$, we call A and B *definably orthogonal* over U.

Proposition 5.2. *Let T be a stable theory. Then*

$$A \mathop{\smile\!\!\!\!\!\!\downarrow}^{\mathrm{f}}_U C \quad \textit{and} \quad C \perp^{\mathrm{do}}_U \operatorname{acl}^{\mathrm{eq}}(U) \quad \textit{implies} \quad A \mathop{\smile\!\!\!\!\!\!\downarrow}^{!}_U C\,.$$

Proof. For a contradiction, suppose that $C \perp^{\mathrm{do}}_U \operatorname{acl}^{\mathrm{eq}}(U)$ and there are tuples $\bar a$ and $\bar b$ such that

$$\bar a \mathop{\smile\!\!\!\!\!\!\downarrow}^{\mathrm{f}}_U C\,, \quad \bar b \mathop{\smile\!\!\!\!\!\!\downarrow}^{\mathrm{f}}_U C\,, \quad \text{and} \quad \bar a \equiv_U \bar b\,,$$

but $\mathfrak{p} := \operatorname{tp}(\bar a/UC)$ and $\mathfrak{q} := \operatorname{tp}(\bar b/UC)$ are different. By Lemma **??**, we can find some $\chi \in \mathrm{FE}(U)$ such that

$$\mathfrak{p}(\bar x) \cup \mathfrak{q}(\bar y) \vDash \neg\chi(\bar x, \bar y)\,.$$

Let E be the set of those equivalence classes $[\bar c]_\chi$ containing some tuple $\bar c' \in [\bar c]_\chi$ realising $\mathfrak{p}$. Then E is finite and we can choose an enumeration $[\bar c_0]_\chi, \dots, [\bar c_{m-1}]_\chi$ of E where each representative $\bar c_i$ realises $\mathfrak{p}$. We set

$$\varphi(x) := \bigvee_{i<m} \chi(\bar x, \bar c_i)\,.$$

Then $\mathbb{M} \vDash \varphi(\bar a)$, while

$$\mathfrak{p}(\bar x) \cup \mathfrak{q}(\bar y) \vDash \neg\chi(\bar x, \bar y) \quad \text{implies} \quad \mathbb{M} \vDash \neg\chi(\bar c_i, \bar b)\,,$$

for all $i < m$. Hence, $\mathbb{M} \vDash \varphi(\bar a) \wedge \neg\varphi(\bar b)$ and it is sufficient to prove that φ is equivalent to a formula over U.

Since $\varphi^{\mathbb{M}} \equiv \bigvee_{i<m} \iota_\chi \bar x = [\bar c_i]_\chi$ is definable over $\operatorname{acl}^{\mathrm{eq}}(U)$ (recall that the structure $\mathbb{M}^{\mathrm{eq}}$ is equipped with projection functions $\iota_\chi : \mathbb{M}^{\bar s} \to \mathbb{M}_\chi$), it is sufficient to show that it is also definable over C. By Theorem E2.1.11 we only have to check that $\pi[\varphi^{\mathbb{M}}] = \varphi^{\mathbb{M}}$, for all $\pi \in \operatorname{Aut}\mathbb{M}_C$. Hence, consider an automorphism $\pi \in \operatorname{Aut}\mathbb{M}_C$. Then π can be extended to an automorphism of $\mathbb{M}^{\mathrm{eq}}_C$ and, therefore, it induces a permutation on the

equivalence classes of χ. Since $\pi[\mathfrak{p}^{\mathbb{M}}] = \mathfrak{p}^{\mathbb{M}}$ it follows that π induces a permutation of E. Hence, we have

$$\begin{aligned}\pi[\varphi^{\mathbb{M}}] = \bigcup_{i<m} \pi[\chi(\bar{x}, \bar{c}_i)]^{\mathbb{M}} &= \bigcup_{i<m} \chi(\bar{x}, \pi(\bar{c}_i))^{\mathbb{M}} \\ &= \bigcup_{i<m} \chi(\bar{x}, \bar{c}_i)^{\mathbb{M}} = \varphi^{\mathbb{M}},\end{aligned}$$

as desired. □

Proposition 5.3. *Let T be a stable theory.*

(a) $\mathrm{fc}(\mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}) \leq \mathrm{fc}(\mathop{\smile\!\!\!\!\!\mid}^{!}) \leq |T|^+$.

(b) $\mathrm{fc}(\mathop{\smile\!\!\!\!\!\mid}^{!}) \leq \mathrm{fc}(\mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}})^{\mathrm{reg}} \oplus \mathrm{mult}(\mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}})^+$.

(c) *If T is $\aleph_0$-stable, then* $\mathrm{fc}(\mathop{\smile\!\!\!\!\!\mid}^{!}) \leq \aleph_0$.

Proof. (a) For the lower bound, note that $\mathop{\smile\!\!\!\!\!\mid}^{!} \subseteq \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}$ implies that every $\mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}$-forking chain is also a $\mathop{\smile\!\!\!\!\!\mid}^{!}$-forking chain. For the upper bound, we prove that $\mathrm{loc}_0(\mathop{\smile\!\!\!\!\!\mid}^{!}) \leq |T|^+$. Since $|T|^+$ is regular, it then follows by Proposition F2.3.24 that $\mathrm{fc}(\mathop{\smile\!\!\!\!\!\mid}^{!}) \leq \mathrm{loc}_0(\mathop{\smile\!\!\!\!\!\mid}^{!})^{\mathrm{reg}} \leq |T|^+$.

Hence, consider a finite set A and an arbitrary set B. We construct an increasing sequence $(U_n)_{n<\omega}$ of subsets $U_n \subseteq B$ of size $|U_n| \leq |T|$ such that the union $U := \bigcup_{n<\omega} U_n$ is a set of size $|U| \leq |T|$ with $A \mathop{\smile\!\!\!\!\!\mid}^{!}_{U} B$.

We start with some set $U_0 \subseteq B$ of size $|U_0| < \mathrm{fc}(\mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}) \leq |T|^+$ such that $A \mathop{\smile\!\!\!\!\!\mid}^{\mathrm{f}}_{U_0} B$. For the inductive step, suppose that U_n is already defined. Note that there are at most

$$|T^{\mathrm{eq}}| \oplus |\mathrm{acl}^{\mathrm{eq}}(U_n)| \leq |T^{\mathrm{eq}}| \oplus |U_n| = |T|$$

formulae over $\mathrm{acl}^{\mathrm{eq}}(U_n)$ and, consequently, at most that many relations that are definable over both B and $\mathrm{acl}^{\mathrm{eq}}(U_n)$. Consequently, we can choose a set $C_n \subseteq B$ of size $|C_n| \leq |T|$ such that every relation definable over both B and $\mathrm{acl}^{\mathrm{eq}}(U_n)$ is definable over C_n. Setting $U_{n+1} := U_n \cup C_n$, it follows that

$$B \perp^{\mathrm{do}}_{U_{n+1}} \mathrm{acl}^{\mathrm{eq}}(U_n) \quad \text{and} \quad |U_{n+1}| = |U_n| \oplus |C_n| \leq |T|.$$

To see that $U := \bigcup_{n<\omega} U_n$ has the desired properties, first note that every relation definable over $\mathrm{acl}^{\mathrm{eq}}(U) = \bigcup_{n<\omega} \mathrm{acl}^{\mathrm{eq}}(U_n)$ is definable over $\mathrm{acl}^{\mathrm{eq}}(U_n)$, for some n, and hence over $U_{n+1} \subseteq U$. Consequently,

$$B \perp^{\mathrm{do}}_U \mathrm{acl}^{\mathrm{eq}}(U)$$

and it follows by Proposition 5.2 that $A \downarrow^{\mathrm{f}}_U B$ implies $A \downarrow^{!}_U B$.

(b), (c) We prove both bounds simultaneously. Let κ be the least regular cardinal such that

$$\mathrm{mult}_{\downarrow^{\mathrm{f}}}(\mathfrak{p}) < \kappa\,, \quad \text{for all types } \mathfrak{p}\,.$$

Then $\mathrm{mult}(\downarrow^{\mathrm{f}}) \leq \kappa \leq \mathrm{mult}(\downarrow^{\mathrm{f}})^+$. Furthermore, for an $\aleph_0$-stable theory, we have $\mathrm{fc}(\downarrow^{\mathrm{f}})^{\mathrm{reg}} \leq \mathrm{st}(T)^{\mathrm{reg}} = \aleph_0$ and we have seen in Corollary G1.4.8 that $\kappa \leq \aleph_0$. Therefore, both (b) and (c) follow if we can prove that

$$\mathrm{fc}(\downarrow^{!}) \leq \mathrm{fc}(\downarrow^{\mathrm{f}})^{\mathrm{reg}} \oplus \kappa\,.$$

For a contradiction, suppose that we can find some $\downarrow^{!}$-forking chain $(B_\alpha)_{\alpha<\gamma}$ for some finite tuple $\bar a$ over $\varnothing$ whose length is $\gamma := \mathrm{fc}(\downarrow^{\mathrm{f}})^{\mathrm{reg}} \oplus \kappa$. Let $I \subseteq \gamma$ be the set of all indices $\alpha < \gamma$ such that $\bar a \not\downarrow^{\mathrm{f}}_{B[<\alpha]} B_\alpha$. Then $(B_\alpha)_{\alpha\in I}$ is a $\downarrow^{\mathrm{f}}$-forking chain for $\bar a$ over $\varnothing$ since

$$A \not\downarrow^{\mathrm{f}}_{B[<\alpha]} B_\alpha \quad \text{implies} \quad A \not\downarrow^{\mathrm{f}}_{B[I]\cap B[<\alpha]} B_\alpha\,, \quad \text{for all } \alpha \in I\,.$$

Hence, $|I| < \mathrm{fc}(\downarrow^{\mathrm{f}}) \leq \gamma$. As γ is regular, it follows that $I \subseteq \beta_0$, for some index $\beta_0 < \gamma$. Note that $\beta_0 + \gamma = \gamma$. Hence, replacing $(B_\alpha)_{\alpha<\gamma}$ by the subsequence $(B_\alpha)_{\beta_0\leq\alpha<\gamma}$ and setting $U := B[<\beta_0]$, we may assume that

$$\bar a \not\downarrow^{!}_{UB[<\alpha]} B_\alpha \quad \text{and} \quad \bar a \downarrow^{\mathrm{f}}_{UB[<\alpha]} B_\alpha\,, \quad \text{for all } \alpha < \gamma\,.$$

We construct a sequence $(\bar c_\alpha)_{\alpha<\gamma}$ of tuples as follows. As

$$\bar a \not\downarrow^{!}_{UB[<\alpha]} B_\alpha \quad \text{and} \quad \bar a \downarrow^{\mathrm{f}}_{UB[<\alpha]} B_\alpha\,,$$

there exists a tuple $\bar{c}'_\alpha$ such that

$$\bar{c}'_\alpha \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{f}}_{UB[<\alpha]} B_\alpha\,, \quad \bar{c}'_\alpha \equiv_{UB[<\alpha]} \bar{a}\,, \quad \text{and} \quad \bar{c}'_\alpha \not\equiv_{UB[<\alpha]B_\alpha} \bar{a}\,.$$

We choose some tuple $\bar{c}_\alpha \equiv_{UB[<\alpha]B_\alpha} \bar{c}'_\alpha$ such that

$$\bar{c}_\alpha \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{f}}_{UB[<\alpha]} B[<\gamma]\,.$$

By a straightforward induction on α, one can show that $\bar{a} \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{f}}_U B[<\alpha]$. Consequently, we have $\bar{c}_\alpha \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{f}}_U B[<\alpha]$ and it follows by transitivity that $\bar{c}_\alpha \mathop{\smile\!\!\!\!\!\!|}^{\mathrm{f}}_U B[<\gamma]$. Since

$$\bar{c}_\alpha \not\equiv_{UB[<\alpha+1]} \bar{a} \equiv_{UB[<\alpha+1]} c_{\alpha'}\,, \quad \text{for all } \alpha < \alpha' < \gamma\,,$$

the types $\mathrm{tp}(\bar{c}_\alpha/UB[<\gamma])$ are distinct $\mathop{\smile\!\!\!\!\!\!|}^{\mathrm{f}}$-free extensions of $\mathrm{tp}(\bar{a}/U)$. By choice of κ, it follows that $\gamma < \kappa$. A contradiction. □

Strongly independent stratifications

Instead of working with single $\mathop{\smile\!\!\!\!\!\!|}^{!}$-constructions, we will work with families of them that are encoded by a stratification for the following relation.

Definition 5.4. Let $\mathfrak{M}$ be the model of a stable theory. We define

$$A \overset{\mathrm{si}}{\surd}_U B \quad :\text{iff} \quad a \mathop{\smile\!\!\!\!\!\!|}^{!}_U B \cup (A \setminus \{a\})\,, \quad \text{for all } a \in A\,.$$

A set A is *strongly independent* over U if $A \overset{\mathrm{si}}{\surd} U$.

Note that $A \overset{\mathrm{si}}{\surd}_U B$ implies that every enumeration of A is a $\mathop{\smile\!\!\!\!\!\!|}^{!}$-construction over U. It follows that every $\overset{\mathrm{si}}{\surd}$-stratification can be refined to a whole family of $\mathop{\smile\!\!\!\!\!\!|}^{!}$-constructions.

Lemma 5.5. *If there exists a $\overset{\mathrm{si}}{\surd}$-stratification ζ of A over U, then A is $\mathop{\smile\!\!\!\!\!\!|}^{!}$-constructible over U.*

Proof. Suppose that $\zeta = (B_\beta)_{\beta<\gamma}$. By Lemma 2.8 (a), it is sufficient to prove that each set B_β is $\mathop{\downarrow}^!$-constructible over $U \cup B[<\beta]$. Hence, let $(b_\alpha)_{\alpha<\delta}$ be an arbitrary enumeration of B_β. Then

$$B_\beta \overset{\text{si}}{\surd} UB[<\beta] \quad \text{implies} \quad b_\alpha \mathop{\downarrow}^!_{UB[<\beta]} b[<\alpha]\,, \quad \text{for all } \alpha < \delta. \square$$

Lemma 5.6. *In a stable theory, the relation* $\overset{\text{si}}{\surd}$ *satisfies all axioms of an isolation relation except for* (LTR).

Proof. (INV) follows immediately form the definition.

(MON) Suppose that $AC \overset{\text{si}}{\surd}_U BD$. Then

$$a \mathop{\downarrow}^!_U BD \cup ((A \cup C) \smallsetminus \{a\})\,, \quad \text{for all } a \in A \cup C\,,$$

which implies that

$$a \mathop{\downarrow}^!_U B \cup (A \smallsetminus \{a\})\,, \quad \text{for all } a \in A\,.$$

Hence, $A \overset{\text{si}}{\surd}_U B$.

(BMON) Suppose that $A \overset{\text{si}}{\surd}_U BC$. Then

$$a \mathop{\downarrow}^!_U BC \cup (A \smallsetminus \{a\})\,, \quad \text{for all } a \in A\,,$$

which implies that

$$a \mathop{\downarrow}^!_{UC} B \cup (A \smallsetminus \{a\})\,, \quad \text{for all } a \in A\,.$$

Hence, $A \overset{\text{si}}{\surd}_{UC} B$.

(NOR) Suppose that $A \overset{\text{si}}{\surd}_U B$. Then

$$a \mathop{\downarrow}^!_U B \cup (A \smallsetminus \{a\})\,, \quad \text{for all } a \in A\,,$$

which implies that

$$a \mathop{\downarrow}^!_U UB \cup (A \smallsetminus \{a\})\,, \quad \text{for all } a \in A\,.$$

Since, by (NOR),

$$c \mathrel{\downarrow}^{!}_{U} UB \cup (A \setminus \{c\}) , \quad \text{for all } c \in U ,$$

it follows that $AU \mathrel{\overset{\mathrm{si}}{\surd}}_{U} BU$.

(LRF) Let A and B be sets. Since

$$a \mathrel{\downarrow}^{!}_{A} B \cup (A \setminus \{a\}) , \quad \text{for all } a \in A ,$$

it follows that $A \mathrel{\overset{\mathrm{si}}{\surd}}_{A} B$.

(FIN) Suppose that $A_0 \mathrel{\overset{\mathrm{si}}{\surd}}_{U} B$, for all finite $A_0 \subseteq A$. To show that $A \mathrel{\overset{\mathrm{si}}{\surd}}_{U} B$, consider an element $a \in A$. For every finite $A_0 \subseteq A \setminus \{a\}$,

$$A_0 a \mathrel{\overset{\mathrm{si}}{\surd}}_{U} B \quad \text{implies} \quad a \mathrel{\downarrow}^{!}_{U} BA_0 .$$

Since $\mathrel{\downarrow}^{!}$ is a symmetric relation with finite character, it follows that

$$a \mathrel{\downarrow}^{!}_{U} B(A \setminus \{a\}) .$$

(RSH) Suppose that $AC \mathrel{\overset{\mathrm{si}}{\surd}}_{U} B$ and $C \mathrel{\overset{\mathrm{si}}{\surd}}_{U} AB$. In order to show that $A \mathrel{\overset{\mathrm{si}}{\surd}}_{U} BC$, we consider an element $a \in A$. Then

$$AC \mathrel{\overset{\mathrm{si}}{\surd}}_{U} B \quad \text{implies} \quad a \mathrel{\downarrow}^{!}_{U} B \cup ((A \cup C) \setminus \{a\}) .$$

If $a \notin C$, then $B \cup ((A \cup C) \setminus \{a\}) = B \cup C \cup (A \setminus \{a\})$ and we are done. Hence, suppose that $a \in C$. Then

$$C \mathrel{\overset{\mathrm{si}}{\surd}}_{U} AB \quad \text{implies} \quad a \mathrel{\downarrow}^{!}_{U} AB \cup (C \setminus \{a\}) .$$

As $AB \cup (C \setminus \{a\}) = BC \cup (A \setminus \{a\})$, the claim follows. □

The next lemma is our main tool to construct $\overset{\mathrm{si}}{\surd}$-stratifications.

Lemma 5.7. *Let T be a stable theory and $c \in \mathbb{M}$. Then*

$$A \mathrel{\overset{\mathrm{si}}{\surd}}_{U} B \quad \textit{and} \quad c \mathrel{\downarrow}^{!}_{U} AB \quad \textit{implies} \quad Ac \mathrel{\overset{\mathrm{si}}{\surd}}_{U} B .$$

Proof. Let $a \in A$ and set $A_0 := A \setminus \{a\}$. We have to show that

$$a \mathop{\smile\!\!\!\!\!\!\downarrow}^!_U BA_0c\,.$$

By symmetry, $c \mathop{\downarrow}^!_{UBA_0} a$ implies $a \mathop{\downarrow}^!_{UBA_0} c$. Since $a \mathop{\downarrow}^!_U BA_0$, it follows by transitivity that $a \mathop{\downarrow}^!_U BA_0c$, as desired. □

We are finally able to prove that $\overset{\mathrm{si}}{\sqrt{}}$-stratifications always exist and that their length can be bounded.

Theorem 5.8. *Let T be a stable theory. Every set $A \subseteq \mathbb{M}$ has a $\overset{\mathrm{si}}{\sqrt{}}$-stratification $\zeta = (B_\alpha)_{\alpha<\gamma}$ over $\varnothing$ of length $\gamma \leq \mathrm{fc}(\mathop{\downarrow}^!)$.*

Proof. Set $\kappa := \mathrm{fc}(\mathop{\downarrow}^!)$. By induction on α, we choose a sequence $(B_\alpha)_{\alpha<\kappa}$ of disjoint subsets $B_\alpha \subseteq A$ as follows. Suppose that B_α has already been defined for all $\alpha < \beta$. Since the union of an increasing chain of strongly independent sets is again strongly independent, we can use the Lemma of Zorn to find a maximal subset $B_\beta \subseteq A \setminus B[<\beta]$ such that

$$B_\beta \overset{\mathrm{si}}{\sqrt{}} B[<\beta]\,.$$

The sequence $(B_\alpha)_{\alpha<\kappa}$ defined in this way is a $\overset{\mathrm{si}}{\sqrt{}}$-stratification of $B[<\kappa]$ over $\varnothing$.

It remains to prove that $B[<\kappa] = A$. For a contradiction, suppose that there is some element $a \in A \setminus B[<\kappa]$. By definition of $\kappa = \mathrm{fc}(\mathop{\downarrow}^!)$ there exists some index $\alpha < \kappa$ such that $a \mathop{\downarrow}^!_{B[<\alpha]} B_\alpha$. Since $B_\alpha \overset{\mathrm{si}}{\sqrt{}}_{B[<\alpha]} B[<\alpha]$, it follows by Lemma 5.7 that

$$B_\alpha a \overset{\mathrm{si}}{\sqrt{}}_{B[<\alpha]} B[<\alpha]\,.$$

This contradicts the maximality of B_α. □

By the special nature of the relation $\overset{\mathrm{si}}{\sqrt{}}$, $\overset{\mathrm{si}}{\sqrt{}}$-stratifications can always be refined.

Definition 5.9. Let $(B_\alpha)_{\alpha<\gamma}$ and $(C_\alpha)_{\alpha<\delta}$ be partitions of a set A. We call $(B_\alpha)_{\alpha<\gamma}$ a *refinement* of $(C_\alpha)_{\alpha<\delta}$ if there exists an increasing function $f : \gamma \to \delta$ such that

$$B_\alpha \subseteq C_{f(\alpha)}, \quad \text{for all } \alpha < \gamma .$$

Lemma 5.10. *Let $\zeta = (C_\alpha)_{\alpha<\delta}$ be a $\overset{\text{si}}{\bigvee}$-stratification of A over U. Every refinement $(B_\alpha)_{\alpha<\gamma}$ of ζ is also a $\overset{\text{si}}{\bigvee}$-stratification of A over U.*

Proof. Let $f : \gamma \to \delta$ be the function such that $B_\alpha \subseteq C_{f(\alpha)}$. To show that $B_\alpha \overset{\text{si}}{\bigvee} B[<\alpha]$, we consider an element $a \in B_\alpha$. Since

$$C[<f(\alpha)] \subseteq B[<\alpha] \quad \text{and} \quad B[\leq\alpha] \subseteq C[\leq f(\alpha)],$$

it follows by monotonicity of $\mathop{\downarrow}^{!}$ that

$$a \mathop{\downarrow}^{!}_{UC[<f(\alpha)]} C[\leq f(\alpha)] \smallsetminus \{a\}$$

implies $a \mathop{\downarrow}^{!}_{UB[<\alpha]} B[\leq\alpha] \smallsetminus \{a\}$. □

When considering $\overset{\text{si}}{\bigvee}$-stratifications as families of $\mathop{\downarrow}^{!}$-constructions, we need to modify the notion of a closed set as follows.

Definition 5.11. Let $\zeta = (B_\alpha)_{\alpha<\gamma}$ be a $\overset{\text{si}}{\bigvee}$-stratification of A over U.

(a) For every $\alpha < \gamma$ and every element $a \in B_\alpha$, let $W(a) \subseteq U \cup B[<\alpha]$. We call the family $(W(a))_{a\in A}$ a *system of bases* for ζ if

$$a \mathop{\downarrow}^{!}_{W(a)} U \cup (B[\leq\alpha] \smallsetminus \{a\}), \quad \text{for all } \alpha < \gamma \text{ and all } a \in B_\alpha .$$

(b) Let $(W(a))_{a\in A}$ be a system of bases for ζ. A set $C \subseteq A$ is W-*closed* if $W(a) \subseteq U \cup C$, for all $a \in C$.

Lemma 5.12. *Let T be a stable theory and let $\zeta = (B_\alpha)_{\alpha<\gamma}$ be a $\overset{\text{si}}{\bigvee}$-stratification of a A over U. There exists a system of bases $(W(a))_{a\in A}$ for ζ such that $|W(a)| < \text{loc}_0(\mathop{\downarrow}^{!})$, for all $a \in A$.*

Proof. For every $a \in B_\alpha$, we choose a set $W(a) \subseteq U \cup B[<\alpha]$ of size $|W(a)| < \mathrm{loc}_o(\downarrow^!)$ such that

$$a \downarrow^!_{W(a)} UB[<\alpha]\,.$$

Since $a \downarrow^!_{UB[<\alpha]} U \cup (B[\leq\alpha] \setminus \{a\})$, it follows by transitivity that

$$a \downarrow^!_{W(a)} U \cup (B[\leq\alpha] \setminus \{a\})\,. \qquad \square$$

Lemma 5.13. *Let $(B_\alpha)_{\alpha<\gamma}$ be a $\sqrt[\mathrm{si}]{}$-stratification of A over U with system of bases $(W(a))_{a\in A}$. If $C \subseteq A$ is W-closed, then $(B_\alpha \cap C)_{\alpha<\gamma}$ is a $\sqrt[\mathrm{si}]{}$-stratification of C over U.*

Proof. Set $C_\alpha := B_\alpha \cap C$, for $\alpha < \gamma$. Consider an element $a \in C_\alpha$. We have to show that $a \downarrow^!_{UC[<\alpha]} C[\leq\alpha] \setminus \{a\}$. Note that $W(a) \subseteq U \cup C[<\alpha]$ and

$$a \downarrow^!_{W(a)} U \cup (B[\leq\alpha] \setminus \{a\})\,,$$

by choice of $W(a)$. By monotonicity, it follows that

$$a \downarrow^!_{UC[<\alpha]} U \cup (C[\leq\alpha] \setminus \{a\})\,,$$

as desired. $\square$

6. Representations

We have introduced indiscernible systems at the end of Section E5.3. Intuitively, if $a : I \to M$ in an indiscernible system over $\mathfrak{I}$, the structure $\mathfrak{M}$ is at least as complicated as $\mathfrak{I}$. If $a : I \to M$ is surjective, the converse is also true to some extend. In this section, we will characterise theories T by classes $\mathcal{C}$ of structures such that, every model $\mathfrak{M}$ of T has a bijective indiscernible system $a : I \to M$ with $\mathfrak{I} \in \mathcal{C}$.

Definition 6.1. Let $\mathcal{C}$ be a class of structures.

(a) A *representation* of a structure $\mathfrak{M}$ in $\mathcal{C}$ is a bijective indiscernible system $r : \mathfrak{I} \to \mathfrak{M}$, for some $\mathfrak{I} \in \mathcal{C}$. If the system $r : \mathfrak{I} \to \mathfrak{M}$ is only QF-indiscernible, i.e., if

$$\operatorname{atp}(\bar{\imath}) = \operatorname{atp}(\bar{k}) \quad \text{implies} \quad \operatorname{atp}(r[\bar{\imath}]) = \operatorname{atp}(r[\bar{k}]) ,$$

we call r a *quantifier-free representation.*

(b) We say that a theory T has *$\mathcal{C}$-representations* if every model $\mathfrak{M}$ of T has a representation in $\mathcal{C}$. Similarly, we say that a class $\mathcal{K}$ has *quantifier-free $\mathcal{C}$-representations* if every structure $\mathfrak{M} \in \mathcal{K}$ has a quantifier-free representation in $\mathcal{C}$.

First, let us note that representations are closed under composition.

Lemma 6.2. *If T has $\mathcal{K}$-representations and $\mathcal{K}$ has quantifier-free $\mathcal{C}$-representations, then T has $\mathcal{C}$-representations.*

Lemma 6.3. *If $\mathcal{C} \subseteq \mathcal{K}$, then $\mathcal{C}$ has quantifier-free $\mathcal{K}$-representations.*

In this section we will characterise stable theories in terms of representations in the following classes.

Definition 6.4. Let κ and λ be cardinals.

(a) $\Upsilon_{\kappa\lambda}$ is the signature consisting of unary predicates P_α, for $\alpha < \kappa$, and unary function symbols f_α, for $\alpha < \lambda$.

(b) We denote the class of all $\Upsilon_{\kappa\lambda}$-structures by $\operatorname{Un}(\kappa, \lambda)$ and the subclass consisting of all structures that are *locally finite* by $\operatorname{Lf}(\kappa, \lambda)$. Finally, $\operatorname{Wf}(\kappa, \lambda) \subseteq \operatorname{Un}(\kappa, \lambda)$ is the subclass of all structures such that the inverse R^{-1} of the relation $R := \bigcup_{\alpha<\lambda} f_\alpha$ is *well-founded.*

First, let us give some simple relationships between these classes.

Lemma 6.5. *Let κ and λ be infinite cardinals.*

(a) $\operatorname{Wf}(\kappa, \lambda)$ *has quantifier-free* $\operatorname{Un}(\kappa, \lambda)$*-representations.*

(b) *If $n < \aleph_0$,* $\operatorname{Wf}(\kappa, n)$ *has quantifier-free* $\operatorname{Lf}(\kappa, n)$*-representations.*

(c) *If* $\kappa_0 \leq \kappa$ *and* $\lambda_0 \leq \lambda$*,* $\mathrm{Un}(\kappa_0, \lambda_0)$ *has quantifier-free* $\mathrm{Un}(\kappa, \lambda)$*-representations.*

(d) $\mathrm{Un}(\kappa, \lambda)$ *has quantifier-free* $\mathrm{Un}(0, \lambda \oplus \kappa)$*-representations.*

(e) $\mathrm{Wf}(\kappa, \lambda)$ *has quantifier-free* $\mathrm{Wf}(0, \lambda \oplus \kappa)$*-representations.*

Proof. (a) follows by Lemma 6.3.

(b) If n is finite, every structure in $\mathrm{Wf}(\kappa, n)$ is locally finite. Therefore $\mathrm{Wf}(\kappa, n) \subseteq \mathrm{Lf}(\kappa, n)$ and the claim follows again by Lemma 6.3.

(c) For $\mathfrak{M} \in \mathrm{Un}(\kappa_0, \lambda_0)$, we construct a quantifier-free representation $r : I \to M$ where $\mathfrak{I} \in \mathrm{Un}(\kappa, \lambda)$ is the expansion of $\mathfrak{M}$ by the following functions and relations:

$$\begin{aligned} R_\alpha &:= \varnothing\,, && \text{for } \kappa_0 \leq \alpha < \kappa\,, \\ f_\alpha &:= \mathrm{id}\,, && \text{for } \lambda_0 \leq \alpha < \lambda\,. \end{aligned}$$

It follows that the identity function $\mathrm{id} : M \to M$ is a quantifier-free indiscernible system over $\mathfrak{I}$.

(d) Consider a structure $\mathfrak{M} \in \mathrm{Un}(\kappa, \lambda)$. If $|M| \leq 1$, let $\mathfrak{I} \in \mathrm{Un}(0, \lambda \oplus \kappa)$ be the unique structure of size $|I| = |M|$ and let $r : I \to M$ be the corresponding bijection. Then $r : I \to M$ is a quantifier-free indiscernible system.

If $|M| > 1$, we proceed as follows. Choosing distinct elements $0, 1 \in M$ we construct the structure $\mathfrak{I}$ with universe $I := M$ and functions

$$\begin{aligned} f_\alpha^{\mathfrak{I}}(a) &:= f_\alpha^{\mathfrak{M}}(a)\,, && \text{for } \alpha < \lambda\,, \\ g_\beta^{\mathfrak{I}}(a) &:= \begin{cases} 0 & \text{if } a \notin P_\beta\,, \\ 1 & \text{if } a \in P_\beta\,, \end{cases} && \text{for } \beta < \kappa\,, \\ h^{\mathfrak{I}}(a) &:= 0\,. \end{aligned}$$

Note that $\mathfrak{I}$ has $\kappa \oplus \lambda \oplus 1$ functions. Hence, after renaming the functions we obtain a structure in $\mathrm{Un}(0, \lambda \oplus \kappa)$. The identity $\mathrm{id} : I \to M$ is a quantifier-free indiscernible system.

(e) Let $\mathfrak{M} \in \mathrm{Wf}(\kappa, \lambda)$. If $|M| \leq 1$, we proceed as in (d). If $|M| \geq 2$, we modify the construction in (d) as follows. Since $\mathfrak{M} \in \mathrm{Wf}(\kappa, \lambda)$, we can choose $0, 1 \in M$ such that $f_\alpha^{\mathfrak{M}}(0) = 0$ and $f_\alpha^{\mathfrak{M}}(1) \in \{0, 1\}$, for all α. Let $\mathfrak{I}$ be the structure with universe M and functions

$$\begin{aligned} f_\alpha^{\mathfrak{I}}(a) &:= f_\alpha^{\mathfrak{M}}(a)\,, && \text{for } \alpha < \lambda\,, \\ g_\beta^{\mathfrak{I}}(a) &:= \begin{cases} 0 & \text{if } a \notin P_\beta \text{ or } a = 0\,, \\ 1 & \text{if } a \in P_\beta \text{ and } a \neq 0\,, \end{cases} && \text{for } \beta < \kappa\,, \\ h^{\mathfrak{I}}(a) &:= 0\,. \end{aligned}$$

After a suitable renaming of the functions, we obtain a structure $\mathfrak{I} \in \mathrm{Wf}(0, \lambda \oplus \kappa)$. Again $\mathrm{id} : I \to M$ is the desired quantifier-free indiscernible system. □

Representable theories are stable

We start by showing that theories represented in one of the above classes are stable. The key argument is contained in the following two Ramsey results.

Lemma 6.6. *Let* $\mathfrak{M} \in \mathrm{Un}(\kappa, \lambda)$, *for infinite cardinals* κ *and* λ, *and let* $n < \omega$. *For every sequence* $(\bar{a}^\alpha)_{\alpha<\mu}$ *where* $\bar{a}^\alpha \in M^n$ *and* $\mu := (2^{\kappa \oplus \lambda})^+$, *there exists an infinite subset* $I \subseteq \mu$ *such that the subsequence* $(\bar{a}^\alpha)_{\alpha \in I}$ *is totally* QF-*indiscernible.*

Proof. Every finitely generated substructure of $\mathfrak{M}$ has size at most λ. By Lemma B1.1.5, there are, up to isomorphism, at most $2^{\kappa \oplus \lambda}$ such substructures. Since $\mu > 2^{\kappa \oplus \lambda}$, there exists a subset $I_0 \subseteq \mu$ of size $|I_0| = \mu$ such that

$$\langle \langle\!\langle \bar{a}^\alpha \rangle\!\rangle_{\mathfrak{M}}, \bar{a}^\alpha \rangle \cong \langle \langle\!\langle \bar{a}^\beta \rangle\!\rangle_{\mathfrak{M}}, \bar{a}^\beta \rangle\,, \quad \text{for all } \alpha, \beta \in I_0\,.$$

For each $\alpha \in I_0$, we fix an enumeration $\bar{b}^\alpha = (b_i^\alpha)_{i<\gamma}$, without repetitions, of $\langle\!\langle \bar{a}^\alpha \rangle\!\rangle_{\mathfrak{M}}$ such that, for all $\alpha, \beta \in I_0$, the map $b_i^\alpha \mapsto b_i^\beta$, $i < \gamma$, induces an isomorphism

$$\langle\langle\!\langle \bar a^\alpha \rangle\!\rangle_{\mathfrak M}, \bar a^\alpha\rangle \to \langle\langle\!\langle \bar a^\beta \rangle\!\rangle_{\mathfrak M}, \bar a^\beta\rangle\,.$$

For every $\nu < \mu$, we have $\nu^{<\lambda^+} = \nu^\lambda \leq (2^{\kappa\oplus\lambda})^\lambda = 2^{\kappa\oplus\lambda} < \mu$. Hence, we can use Lemma A4.6.11 to find a subset $U \subseteq M$ and a subset $I_1 \subseteq I_0$ of size $|I_1| = \mu$ such that

$$\bar b^\alpha \cap \bar b^\beta = U\,, \quad \text{for all } \alpha \neq \beta \text{ in } I_1\,.$$

Since $\mu > 2^\lambda$, there exists a subset $I_2 \subseteq I_1$ of size $|I_2| = \mu$ and a set $K \subseteq \lambda$ such that, for all $\alpha \in I_2$,

$$K = \{\, i < \gamma \mid b^\alpha_i \in U \,\}\,.$$

We claim that the sequence $(\bar a^i)_{i\in I_2}$ is totally QF-indiscernible. Let $\varphi(\bar x)$ be an atomic formula. We have to show that

$$\mathfrak M \vDash \varphi(\bar a[\bar\alpha]) \leftrightarrow \varphi(\bar a[\bar\beta])\,, \quad \text{for all } \bar\alpha, \bar\beta \subseteq I_2\,.$$

First, we consider the case where $\varphi = (s = t)$ is an equation. Then there are indices $\xi_0, \dots, \xi_{m-1}, \eta_0, \dots, \eta_{n-1} < \lambda$ and variables x, y such that

$$s(x) = f_{\xi_{m-1}} \cdots f_{\xi_0} x \quad \text{and} \quad t(y) = f_{\eta_{n-1}} \cdots f_{\eta_0} y\,.$$

For each component $i < n$ of $\bar a^\alpha$, there are indices i' and i'' such that $b^\alpha_{i'} = s(a^\alpha_i)$ and $b^\alpha_{i''} = t(a^\alpha_i)$. Since we have chosen $\bar b^\alpha$ without repetitions, it follows that

$$\begin{aligned} \mathfrak M \vDash s(a^\alpha_i) = t(a^\beta_k) \quad &\text{iff} \quad b^\alpha_{i'} = b^\beta_{k''} \\ &\text{iff} \quad i' = k'' \text{ and } (\alpha = \beta \text{ or } i' \in K)\,. \end{aligned}$$

The latter condition is invariant under permutations of I_2.

It remains to consider the case where $\varphi = P_\zeta t$ for $t = f_{\xi_{m-1}} \cdots f_{\xi_0} x$. Again we can find, for every component i of $\bar{a}^\alpha$ an index i' such that $b^\alpha_{i'} = t(a^\alpha_i)$. Therefore,

$$\mathfrak{M} \vDash P_\zeta(t(a^\alpha_i)) \quad \text{iff} \quad b^\alpha_{i'} \in P_\zeta .$$

The latter condition does not depend on α, since the substructures induced by the tuples $\bar{b}^\alpha$ are isomorphic. □

Lemma 6.7. *Let $\mathfrak{M} \in \mathrm{Lf}(\kappa, \lambda)$, let $\mu > \kappa, \lambda$ be an uncountable cardinal, and $n < \omega$. For every sequence $(\bar{a}^\alpha)_{\alpha<\mu}$ of n-tuples $\bar{a}^\alpha \in M^n$ of length μ and every set $U \subseteq M$ of size $|U| < \mu$, there exists a subset $I \subseteq \mu$ of size $|I| = \mu$ such that*

$$\mathrm{atp}(\bar{a}^\alpha/U) = \mathrm{atp}(\bar{a}^\beta/U) , \qquad \textit{for all } \alpha, \beta \in I .$$

Proof. Since each substructure $\langle\!\langle \bar{a}^\alpha \rangle\!\rangle_\mathfrak{M}$ is finite and there are at most $\kappa \oplus \lambda \oplus \aleph_0 < \mu$ finite substructures, there is a subset $I_0 \subseteq \mu$ of size $|I_0| = \mu$ such that

$$\langle \langle\!\langle \bar{a}^\alpha \rangle\!\rangle_\mathfrak{M}, \bar{a}^\alpha \rangle \cong \langle \langle\!\langle \bar{a}^\beta \rangle\!\rangle_\mathfrak{M}, \bar{a}^\beta \rangle , \quad \text{for all } \alpha, \beta \in I_0 .$$

Let $V := \langle\!\langle U \rangle\!\rangle_\mathfrak{M}$. Since there are at most $|V|^{<\omega} = |V| \oplus \aleph_0 < \mu$ finite subsets of V, we can find a subset $I_1 \subseteq I_0$ of size $|I_1| = \mu$ and a finite set $W \subseteq V$ such that

$$\langle\!\langle \bar{a}^\alpha \rangle\!\rangle_\mathfrak{M} \cap V = W , \quad \text{for all } \alpha \in I_1 .$$

Let $\bar{c}$ be an enumeration of W. There exists a subset $I_2 \subseteq I_1$ of size $|I_2| = \mu$ such that

$$\langle \langle\!\langle \bar{a}^\alpha \rangle\!\rangle_\mathfrak{M}, \bar{a}^\alpha, \bar{c} \rangle \cong \langle \langle\!\langle \bar{a}^\beta \rangle\!\rangle_\mathfrak{M}, \bar{a}^\beta, \bar{c} \rangle , \quad \text{for all } \alpha, \beta \in I_2 .$$

It follows that

$$\mathrm{atp}(\bar{a}^\alpha/V) = \mathrm{atp}(\bar{a}^\beta/V) , \quad \text{for all } \alpha, \beta \in I_2 .$$ □

Using these two lemmas we can show that theories with representations in one of the above classes are stable.

Proposition 6.8. *Let T be a complete first-order theory.*

(a) *If T has* $\mathrm{Un}(\kappa, \lambda)$*-representations, for some cardinals κ, λ, it is stable.*

(b) *If T has* $\mathrm{Lf}(\kappa, \kappa)$*-representations, it is λ-stable, for all $\lambda \geq \kappa \oplus \aleph_0$.*

(c) *If T has* $\mathrm{Lf}(\kappa, \lambda)$*-representations, it is superstable with* $\mathrm{st}(T) \leq \kappa \oplus \lambda \oplus \aleph_0$.

Proof. (a) With out loss of generality, we may assume that κ and λ are infinite. For a contradiction, assume that T is unstable. Then we can use Theorem E5.3.13 to find a model $\mathfrak{M}$ containing an infinite indiscernible sequence $(\bar{a}^i)_{i \in I}$ that is not totally indiscernible. By Lemma E5.3.9, we can find such a sequence where $I = \mu$, for $\mu := (2^{\kappa \oplus \lambda})^+$. Let $r : \mathfrak{I} \to \mathfrak{M}$ be a representation of $\mathfrak{M}$ in $\mathrm{Un}(\kappa, \lambda)$ and set $\bar{b}^i := r^{-1}(\bar{a}^i)$. By Lemma 6.6, there exists an infinite subset $I_0 \subseteq I$ such that the subsequence $(\bar{b}^i)_{i \in I_0}$ is totally QF-indiscernible. It follows that $(\bar{a}^i)_{i \in I_0}$ is totally indiscernible. As $\mathrm{Av}(\bar{a}^i)_{i \in I_0} = \mathrm{Av}(\bar{a}^i)_{i \in I}$, the sequence $(\bar{a}^i)_{i \in I}$ is also totally indiscernible. A contradiction.

(b) For a contradiction, suppose that $|S^{\bar{s}}(U)| > \lambda$, for some finite tuple $\bar{s}$ of sorts and some set U of parameters of size $|U| = \lambda \geq \kappa \oplus \aleph_0$. Fix a sequence $(\bar{a}^\alpha)_{\alpha < \lambda^+}$ such that

$$\mathrm{tp}(\bar{a}^\alpha / U) \neq \mathrm{tp}(\bar{a}^\beta / U)\,, \quad \text{for all } \alpha \neq \beta\,.$$

Let $r : \mathfrak{I} \to \mathfrak{M}$ be a representation of $\mathfrak{M}$ in $\mathrm{Lf}(\kappa, \kappa)$. We set

$$V := \langle\!\langle r^{-1}[U] \rangle\!\rangle_{\mathfrak{I}} \quad \text{and} \quad \bar{b}^\alpha := r^{-1}(\bar{a}^\alpha)\,, \quad \text{for } \alpha < \lambda^+.$$

By Lemma 6.7, there exists a subset $I \subseteq \lambda^+$ of size $|I| = \lambda^+$ such that

$$\mathrm{atp}(\bar{b}^\alpha / V) = \mathrm{atp}(\bar{b}^\beta / V)\,, \quad \text{for all } \alpha, \beta \in I\,.$$

Since r is a representation, it follows that

$$\operatorname{tp}(\bar{a}^\alpha/U) = \operatorname{tp}(\bar{a}^\beta/U), \quad \text{for all } \alpha, \beta \in I,$$

in contradiction to our choice of $(\bar{a}^\alpha)_\alpha$.

(c) According to Theorem G1.6.6, a theory T is superstable if, and only if, there exists some cardinal λ such that T is κ-stable, for all $\kappa \geq \lambda$. Consequently, the claim follows from (b). □

Stable theories have representations

For the converse statements, we employ $\sqrt[\mathrm{si}]{}$-stratifications. The following two technical lemmas contain the key argument.

Lemma 6.9. *Let T be stable, $\zeta = (B_\alpha)_{\alpha<\gamma}$ a $\sqrt[\mathrm{si}]{}$-stratification of some set M over $\varnothing$ such that $|B_0| = |B_1| = 1$, and let $(W(a))_{a\in M}$ be a system of bases for ζ.*

Suppose that $p : C \to D$ is a bijective function with $C, D \subseteq M$ satisfying the following conditions:

- $p[B_\alpha] \subseteq B_\alpha$, *for all* $\alpha < \gamma$.
- C *is* W*-closed.*
- $p[W(a)] = W(p(a))$, *for all* $a \in C$.
- $p[\operatorname{tp}(a/W(a))] = \operatorname{tp}(p(a)/W(p(a)))$, *for all* $a \in C$.

Then p is an elementary map.

Proof. Set $C_\alpha := C \cap B_\alpha$ and $D_\alpha := D \cap B_\alpha$. By assumption, $p[C_\alpha] = D_\alpha$. We will show by induction on $\alpha < \gamma$ that

$$p[\operatorname{tp}(\bar{a}/C[<\alpha])] = \operatorname{tp}(p(\bar{a})/D[<\alpha]), \quad \text{for all finite } \bar{a} \subseteq C_\alpha .$$

For $\alpha < 2$, the claim holds trivially since $|B_\alpha| = 1$. For $\alpha \geq 2$, we prove the statement by induction on $|\bar{a}|$. Hence, suppose that $\bar{a} = b\bar{c}$ and that we have already shown that

$$p[\operatorname{tp}(\bar{c}/C[<\alpha])] = \operatorname{tp}(p(\bar{c})/D[<\alpha]) .$$

By assumption, we have

$$\begin{aligned} & W(p(b)) = p[W(b)]\,, \\ \text{and} \quad & \mathrm{tp}\big(p(b) \,/\, W(p(b))\big) = p\big(\mathrm{tp}(b/W(b))\big)\,. \end{aligned}$$

Since $b \mathop{\smile\!\!\!\!\!\mid}^{!}_{W(b)} B[\leq\alpha] \setminus \{b\}$, the type $\mathfrak{p} := \mathrm{tp}(b/C[<\alpha]\bar{c})$ is the unique free extension of $\mathrm{tp}(b/W(b))$. As we have already shown that the map $p \upharpoonright C[<\alpha]\bar{c}$ is elementary, it follows that the image $p(\mathfrak{p})$ does not fork over $p[W(b)] = W(p(b))$. Since

$$\mathrm{tp}(p(b)/W(p(b))) = \mathrm{tp}(p(b)/p[W(b)]) = p(\mathrm{tp}(b/W(b)))$$

and $p(b) \mathop{\smile\!\!\!\!\!\mid}^{!}_{W(p(b))} D[<\alpha]p(\bar{c})$, it follows that

$$p(\mathfrak{p}) = \mathrm{tp}\big(p(b) \,/\, D[<\alpha]p(\bar{c})\big)\,.$$

Consequently,

$$p\big(\mathrm{tp}(b\bar{c}/C[<\alpha])\big) = \mathrm{tp}\big(p(b\bar{c})/D[<\alpha]\big)\,. \qquad \square$$

Lemma 6.10. *Let $\mathfrak{M}$ be a structure, $\zeta = (B_\alpha)_{\alpha<\gamma}$ a $\sqrt[\mathrm{si}]{}$-stratification of M over $\varnothing$ such that $|B_0| = |B_1| = 1$, and let $(W(a))_{a\in M}$ be a system of bases for ζ. If*

$$\mathfrak{I} := \big\langle M,\ (f_\alpha)_{\alpha<\kappa},\ (P_\alpha)_{\alpha<\lambda},\ (Q_\alpha)_{\alpha<\mu}\big\rangle$$

is a structure such that

- $\{\, f_\alpha(a) \mid \alpha < \kappa \,\} = \{a\} \cup W(a)\,, \quad$ *for all* $a \in M$,
- $a \in P_\beta \Leftrightarrow b \in P_\beta$, *for all* $\beta < \lambda$, *implies that* $a \in B_\alpha \Leftrightarrow b \in B_\alpha$, *for all* $\alpha < \gamma$,
- $a \in Q_\beta \Leftrightarrow b \in Q_\beta$, *for all* $\beta < \mu$, *implies that*

$$a(f_\alpha(a))_{\alpha<\kappa} \equiv b(f_\alpha(b))_{\alpha<\kappa}\,,$$

then the identity map $\mathrm{id} : M \to M$ *is a representation of* $\mathfrak{M}$ *in* $\mathfrak{I}$.

Proof. Let $\bar{a}, \bar{b} \subseteq M$ be finite tuples such that

$$\langle\langle\!\langle \bar{a} \rangle\!\rangle_{\mathfrak{I}}, \bar{a}\rangle \cong \langle\langle\!\langle \bar{b} \rangle\!\rangle_{\mathfrak{I}}, \bar{b}\rangle .$$

We have to show that $\bar{a} \equiv \bar{b}$ in $\mathfrak{M}$. Let $p : \langle\!\langle \bar{a} \rangle\!\rangle_{\mathfrak{I}} \to \langle\!\langle \bar{b} \rangle\!\rangle_{\mathfrak{I}}$ be an isomorphism with $p(\bar{a}) = \bar{b}$. It is sufficient to prove that p satisfies the conditions of Lemma 6.9.

By assumption on the predicates P_β we have

$$p(c) \in B_\alpha \quad \text{iff} \quad c \in B_\alpha .$$

Hence, $p[B_\alpha] \subseteq B_\alpha$. Furthermore, if $c \in \mathrm{dom}(p) = \langle\!\langle \bar{a} \rangle\!\rangle_{\mathfrak{I}}$, then

$$\{c\} \cup W(c) = \{\, f_\alpha(c) \mid \alpha < \kappa \,\} \subseteq \langle\!\langle \bar{a} \rangle\!\rangle_{\mathfrak{I}} = \mathrm{dom}(p) ,$$

and
$$\begin{aligned} p[W(c)] &= \{\, p(f_\alpha(c)) \mid \alpha < \kappa,\ f_\alpha(c) \neq c \,\} \\ &= \{\, f_\alpha(p(c)) \mid \alpha < \kappa,\ f_\alpha(p(c)) \neq p(c) \,\} \\ &= W(p(c)) . \end{aligned}$$

Hence, $\mathrm{dom}(p)$ is W-closed and $p[W(c)] = W(p(c))$.

Finally, for $c \in \mathrm{dom}(p)$, it follows by assumption on the predicates Q_α that

$$c\bar{d} \equiv p(c\bar{d}) ,$$

where $\bar{d}$ is an enumeration of $W(c)$. Hence,

$$p[\mathrm{tp}(c/W(c))] = \mathrm{tp}(p(c)/W(p(c))) . \qquad \square$$

Using these two lemmas we can construct representations for stable theories.

Proposition 6.11. *Let T be a stable theory and set $\kappa := \min\{\mathrm{st}(T), |T|\}$ and $\lambda := \min\{\mathrm{fc}(\downarrow^{!}), |T|\}$.*

(a) *T has $\mathrm{Wf}(\kappa, \lambda)$-representations.*

(b) *If $\mathrm{fc}(\downarrow^{!}) \leq \aleph_0$, T has $\mathrm{Lf}(\kappa, \aleph_0)$-representations.*

Proof. Let $\mathfrak{M}$ be a model of T. By Theorem 5.8, there exists a $\sqrt[\mathrm{i}]{\ }$-stratification $\zeta = (B_n)_{n<\gamma}$ of M over $\varnothing$ of length $\gamma \leq \mathrm{fc}(\downarrow\!\!|^{!})$. Since

$$\mathrm{fc}(\downarrow\!\!|^{!}) \leq \mathrm{fc}(\downarrow\!\!|^{\mathrm{f}}) \oplus \mathrm{mult}(\downarrow\!\!|^{\mathrm{f}})^{+} \leq \mathrm{st}(T)^{+} \quad \text{and} \quad \mathrm{fc}(\downarrow\!\!|^{!}) \leq |T|^{+},$$

it follows that $\gamma \leq \kappa^{+}$. Taking a suitable refinement of ζ we may assume by Lemma 5.10 that $|B_0| = |B_1| = 1$. By Lemma 5.12, there exists a system of bases $(W(a))_{a\in M}$ with

$$|W(a)| < \mathrm{loc}_0(\downarrow\!\!|^{!}) \leq \mathrm{fc}(\downarrow\!\!|^{!}) \leq |T|^{+}.$$

Consequently, $|W(a)| \leq \lambda$.

We define $\mathfrak{I} := \langle M, (f_\alpha)_{\alpha<\lambda}, (P_\alpha)_{\alpha<\kappa}, (Q_\alpha)_{\alpha<\kappa}\rangle$ as follows. We choose the functions $(f_\alpha)_{\alpha<\lambda}$ such that, for every $a \in M$, $(f_\alpha(a))_{\alpha<\lambda}$ is an enumeration of $\{a\} \cup W(a)$. Fixing an injective function $h : 2^{\kappa} \to \mathscr{P}(\kappa)$, we define

$$P_\alpha := \bigcup \{\, B_\beta \mid \alpha \in h(\beta) \,\}.$$

(Note that $\gamma \leq \kappa^{+} \leq 2^{\kappa}$.) As there are at most $\mathrm{st}(T)$ many types over the empty set and there are only $|T|$ formulae, we can fix an enumeration $(\varphi_\alpha(x))_{\alpha<\kappa}$ of all formulae (up to logical equivalence) of the form

$$\varphi(x; f_{\beta_0}x, \dots, f_{\beta_{n-1}}x),$$

where φ is a formula over the signature of T, $n < \omega$, and $\beta_0, \dots, \beta_{n-1} < \kappa$. We set

$$Q_\alpha := \{\, a \in M \mid \langle \mathfrak{M}, (f_\alpha)_\alpha\rangle \vDash \varphi_\alpha(a) \,\}.$$

It is straightforward to check that the structure $\mathfrak{I}$ satisfies the conditions of Lemma 6.10. Hence, the identity function $\mathrm{id} : M \to M$ is a representation of $\mathfrak{M}$ in $\mathfrak{I}$.

Finally, note that $\mathfrak{I} \in \mathrm{Wf}(\kappa, \lambda)$ since,

$$a \in B_\alpha \quad \text{implies} \quad f_\beta(a) \in \{a\} \cup B[<\alpha].$$

If $\mathrm{fc}(\downarrow\!\!|^{!}) \leq \aleph_0$, the sets $\{\, f_\alpha(a) \mid \alpha < \kappa \,\} = W(a) \cup \{a\}$ are finite, for all $a \in M$. By the Lemma of Kőnig, it follows that $\mathfrak{I} \in \mathrm{Lf}(\kappa, \aleph_0)$. ☐

The following two theorems summarise the results of this section.

Theorem 6.12 (Cohen, Shelah)**.** *Let T be a complete first-order theory. The following conditions are equivalent:*

(1) *T is stable.*

(2) *T has* $\mathrm{Un}(\kappa, \lambda)$*-representations, for some cardinals κ and λ.*

(3) *T has* $\mathrm{Wf}(0, |T|)$*-representations.*

(4) *T has* $\mathrm{Wf}(|T|, |T|)$*-representations.*

Proof. (2) ⇒ (1) has been shown in Proposition 6.8 (a), the implications (4) ⇒ (3) ⇒ (2) follow from Lemmas 6.5 and 6.2, and (1) ⇒ (4) follows by Proposition 6.11. □

Theorem 6.13 (Cohen, Shelah)**.** *Let T be a complete first-order theory. The following conditions are equivalent:*

(1) *T is $\aleph_0$-stable.*

(2) *T has* $\mathrm{Lf}(\aleph_0, \aleph_0)$*-representations.*

Proof. (2) ⇒ (1) follows by Proposition 6.8 (b) and (1) ⇒ (2) follows by Proposition 6.11. □

Recommended Literature

Set theory

M. D. Potter, *Sets. An Introduction,* Oxford University Press 1990.

A. Lévy, *Basic Set Theory,* Springer 1979, Dover 2002.

K. Kunen, *Set Theory. An Introduction to Independence Proofs,* North-Holland 1983.

T. J. Jech, *Set Theory,* 3rd ed., Springer 2003.

Algebra and Category Theory

G. M. Bergman, *An Invitation to General Algebra and Universal Constructions,* 2nd ed., Springer 2015.

P. M. Cohn, *Universal Algebra,* 2nd ed., Springer 1981.

P. M. Cohn, *Basic Algebra,* Springer 2003.

S. Lang, *Algebra,* 3rd ed., Springer 2002.

F. Borceux, *Handbook of Categorical Algebra,* Cambridge University Press 1994.

S. MacLane, *Categories for the Working Mathematician,* 2nd ed., Springer 1998.

J. Adámek, J. Rosický, and M. Vitale, *Algebraic Theories,* Cambridge University Press 2011.

J. Adámek and J. Rosický, *Locally Presentable and Accessible Categories,* Cambridge University Press 1994.

Topology and lattice theory

R. Engelking, *General Topology,* 2nd ed., Heldermann 1989.

C.-A. Faure, A. Frölicher, *Modern Projective Geometry,* Kluwer 2000.

P. T. Johnstone, *Stone Spaces,* Cambridge University Press 1982.

G. Gierz, K. H. Hofmann, K. Keimel, J. D. Lawson, M. Mislove, and D. S. Scott, *Continuous Lattices and Domains,* Cambridge University Press 2003.

Model theory

K. Tent and M. Ziegler, *A Course in Model Theory,* Cambridge University Press 2012.

W. Hodges, *Model Theory,* Cambridge University Press 1993.

B. Poizat, *A Course in Model Theory,* Springer 2000.

C. C. Chang and H. J. Keisler, *Model Theory,* 3rd ed., North-Holland 1990.

General model theory

J. Barwise and S. Feferman, eds., *Model-Theoretic Logics,* Springer 1985.

J. T. Baldwin, *Categoricity,* AMS 2010.

R. Diaconescu, *Institution-Independent Model Theory,* Birkhäuser 2008.

H.-D. Ebbinghaus and J. Flum, *Finite Model Theory,* Springer 1995.

Stability theory

S. Buechler, *Essential Stability Theory,* Springer 1996.

E. Casanovas, *Simple Theories and Hyperimaginaries,* Cambridge University Press 2011.

A. Pillay, *Geometric Stability Theory,* Oxford Science Publications 1996.

F. O. Wagner, *Simple Theories,* Kluwer Academic Publishers 2000.

S. Shelah, *Classification Theory,* 2nd ed., North-Holland 1990.

Symbol Index

Chapter A1

Chapter A2

Chapter A3

Chapter A4

Chapter B1

Chapter B2

Chapter B3

Chapter B4

Chapter B5

Chapter B6

Chapter C1

Chapter C2

Chapter C3

Chapter C4

Chapter C5

Chapter D1

Chapter D2

Chapter D3

Chapter E2

Chapter E3

Chapter E4

Chapter E5

Chapter F3

Chapter F4

Chapter F5

Chapter G1

Chapter G2

Index

The Roman and Fraktur alphabets

A	a	$\mathfrak{A}$	$\mathfrak{a}$	N	n	$\mathfrak{N}$	$\mathfrak{n}$
B	b	$\mathfrak{B}$	$\mathfrak{b}$	O	o	$\mathfrak{O}$	$\mathfrak{o}$
C	c	$\mathfrak{C}$	$\mathfrak{c}$	P	p	$\mathfrak{P}$	$\mathfrak{p}$
D	d	$\mathfrak{D}$	$\mathfrak{d}$	Q	q	$\mathfrak{Q}$	$\mathfrak{q}$
E	e	$\mathfrak{E}$	$\mathfrak{e}$	R	r	$\mathfrak{R}$	$\mathfrak{r}$
F	f	$\mathfrak{F}$	$\mathfrak{f}$	S	s	$\mathfrak{S}$	ſ $\mathfrak{s}$
G	g	$\mathfrak{G}$	$\mathfrak{g}$	T	t	$\mathfrak{T}$	$\mathfrak{t}$
H	h	$\mathfrak{H}$	$\mathfrak{h}$	U	u	$\mathfrak{U}$	$\mathfrak{u}$
I	i	$\mathfrak{I}$	$\mathfrak{i}$	V	v	$\mathfrak{V}$	$\mathfrak{v}$
J	j	$\mathfrak{J}$	$\mathfrak{j}$	W	w	$\mathfrak{W}$	$\mathfrak{w}$
K	k	$\mathfrak{K}$	$\mathfrak{k}$	X	x	$\mathfrak{X}$	$\mathfrak{x}$
L	l	$\mathfrak{L}$	$\mathfrak{l}$	Y	y	$\mathfrak{Y}$	$\mathfrak{y}$
M	m	$\mathfrak{M}$	$\mathfrak{m}$	Z	z	$\mathfrak{Z}$	$\mathfrak{z}$

The Greek alphabet

A	α	alpha	N	ν	nu
B	β	beta	Ξ	ξ	xi
Γ	γ	gamma	O	o	omicron
Δ	δ	delta	Π	π	pi
E	ε	epsilon	P	ρ	rho
Z	ζ	zeta	Σ	σ	sigma
H	η	eta	T	τ	tau
Θ	ϑ	theta	Υ	υ	upsilon
I	ι	iota	Φ	ϕ	phi
K	κ	kappa	X	χ	chi
Λ	λ	lambda	Ψ	ψ	psi
M	μ	mu	Ω	ω	omega