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

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 $A, B, C, \dots$.

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

$$A_s \subseteq B_s\,, \qquad \text{for each sort } s \in S\,,$$
$$R^{\mathfrak{A}} = R^{\mathfrak{B}} \cap A^n, \qquad \text{for every } n\text{-ary relation symbol } R \in \Sigma\,,$$
$$\text{and} \quad f^{\mathfrak{A}} = f^{\mathfrak{B}} \cap A^{n+1}, \qquad \text{for every } n\text{-ary function symbol } f \in \Sigma\,.$$

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

$$R^{\mathfrak{A}_0} := R^{\mathfrak{A}} \cap X^n, \qquad \text{for every } n\text{-ary relation } R \in \Sigma\,,$$
$$f^{\mathfrak{A}_0} := f^{\mathfrak{A}} \cap X^{n+1}, \qquad \text{for every } n\text{-ary function } f \in \Sigma\,. \quad \square$$

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 $s \in S \smallsetminus T$ and all relations and function from $\Sigma \smallsetminus \Gamma$.

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 $A_t := \varnothing$, for $t \in T \smallsetminus S$. 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.*

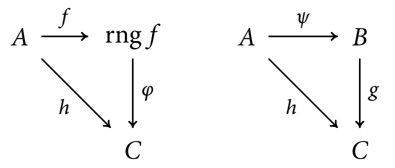
Proof. Suppose that $h(\bar{a}) \in R^{\mathfrak{B}}$. Then there is some tuple $\bar{a}' \in R^{\mathfrak{A}}$ with $h(\bar{a}') = h(\bar{a})$. Since h is injective, it follows that $\bar{a}' = \bar{a}$ and, hence, $\bar{a} \in R^{\mathfrak{A}}$. □

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 $\langle a, a' \rangle \in \ker h$.

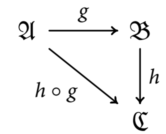
($\Leftarrow$) For $b \in \operatorname{rng} f$, select an arbitrary element $a \in f^{-1}(b)$ and set $\varphi(b) := g(a)$. We claim that $\varphi \circ f = g$. Let $a \in A$ and set $b := f(a)$. By definition of φ, we have $\varphi(b) = g(a')$, for some element $a' \in A$ with $f(a') = b$. Hence, $\langle a, a' \rangle \in \ker f \subseteq \ker g$, 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 $g(\psi(a)) = g(b) = h(a)$. □

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

$$\begin{aligned} f^{\mathfrak{C}}(h(\bar{b})) = f^{\mathfrak{C}}\big(h(g(\bar{a}))\big) &= (h \circ g)(f^{\mathfrak{A}}(\bar{a})) \\ &= h\big(f^{\mathfrak{B}}(g(\bar{a}))\big) = h(f^{\mathfrak{B}}(\bar{b})). \end{aligned}$$

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 $g(\bar{a}) = \bar{b} \in R^{\mathfrak{B}}$.

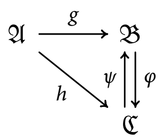
(ii) Since h is injective we have $h^{-1}[h[X]] = X$, for every $X \subseteq B$. Furthermore, injective semi-strict homomorphisms are strict. Therefore, we have

$$g[R^{\mathfrak{A}}] = h^{-1}[h[g[R^{\mathfrak{A}}]]] = h^{-1}[R^{\mathfrak{C}}] = R^{\mathfrak{B}}.$$

(iii) As in (ii) we have

$$g^{-1}[R^{\mathfrak{B}}] = g^{-1}[h^{-1}[h[R^{\mathfrak{B}}]]] = (h \circ g)^{-1}[R^{\mathfrak{C}}] = R^{\mathfrak{A}}.$$ □

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 $\psi \circ \varphi = \text{id}$. In the same way we obtain $\varphi \circ \psi = \text{id}$. 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* $\operatorname{rng} h$ *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 $\mathcal{C}^{\mathrm{obj}}$ of *objects*,
- for each pair of objects $\mathfrak{a}, \mathfrak{b} \in \mathcal{C}^{\mathrm{obj}}$, a set $\mathcal{C}(\mathfrak{a}, \mathfrak{b})$ of *morphisms* from $\mathfrak{a}$ to $\mathfrak{b}$, and
- for all $\mathfrak{a}, \mathfrak{b}, \mathfrak{c} \in \mathcal{C}^{\mathrm{obj}}$, 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 $f(a) = b$.

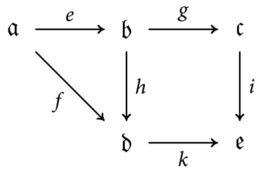
(e) Similarly, in the category $\mathfrak{Set}^2$ the objects are pairs $\langle A, A_0 \rangle$ of sets with $A_0 \subseteq A$ and a morphism $f : \langle A, A_0 \rangle \to \langle B, B_0 \rangle$ is a function $f : A \to B$ such that $f[A_0] \subseteq B_0$.

(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

$$f_m \circ \dots \circ f_0 = g_n \circ \dots \circ g_0$$

holds, where $f_0, \dots, f_m$ and $g_0, \dots, g_n$ 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 $f \upharpoonright \mathbb{Z}$. 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 $\mathcal{C}^{op}$ to $\mathcal{D}$.

(d) The *opposite* of a functor $F : \mathcal{C} \to \mathcal{D}$ is the functor $F^{op} : \mathcal{C}^{op} \to \mathcal{D}^{op}$ with

$$F^{op}(\mathfrak{a}) := F(\mathfrak{a}), \quad \text{for } \mathfrak{a} \in \mathcal{C}^{obj},$$
$$F^{op}(f^{op}) := F(f)^{op}, \quad \text{for } f \in \mathcal{C}^{mor}.$$

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 $g_0 : \mathfrak{b} \to \mathfrak{a}$ with $F(g_0) = g$. Hence,

$$F(\mathrm{id}_{\mathfrak{a}}) = \mathrm{id}_{F(\mathfrak{a})} = F(g_0) \circ F(f) = F(g_0 \circ f).$$

Since F is faithful, this implies that $g_0 \circ f = \mathrm{id}_{\mathfrak{a}}$. 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

- $\mathcal{C}^{obj} \subseteq \mathcal{D}^{obj}$ and $\mathcal{C}^{mor} \subseteq \mathcal{D}^{mor}$,
- 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

$$\mathcal{C}(\mathfrak{a}, \mathfrak{b}) = \mathcal{D}(\mathfrak{a}, \mathfrak{b}), \quad \text{for all } \mathfrak{a}, \mathfrak{b} \in \mathcal{C}^{obj}.$$

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

$$(\mathcal{C} \times \mathcal{D})^{\mathrm{obj}} := \mathcal{C}^{\mathrm{obj}} \times \mathcal{D}^{\mathrm{obj}},$$

and $$(\mathcal{C} \times \mathcal{D})(\langle \mathfrak{a}_0, \mathfrak{a}_1 \rangle, \langle \mathfrak{b}_0, \mathfrak{b}_1 \rangle) := \mathcal{C}(\mathfrak{a}_0, \mathfrak{b}_0) \times \mathcal{D}(\mathfrak{a}_1, \mathfrak{b}_1)\,,$$

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:

$$\langle f_0, f_1 \rangle \circ \langle g_0, g_1 \rangle := \langle f_0 \circ g_0,\ f_1 \circ g_1 \rangle\,.$$

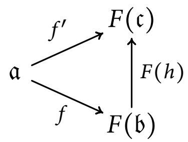
With each product $\mathcal{C} \times \mathcal{D}$ are associated two *projection functors*

$$P_0 : \mathcal{C} \times \mathcal{D} \to \mathcal{C} \quad \text{and} \quad P_1 : \mathcal{C} \times \mathcal{D} \to \mathcal{D}\,,$$

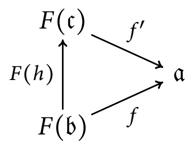
where P_i maps an object $\langle \mathfrak{a}_0, \mathfrak{a}_1 \rangle$ to $\mathfrak{a}_i$ and a morphism $\langle f_0, f_1 \rangle$ to f_i.

(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

$$f' = F(h) \circ f\,.$$

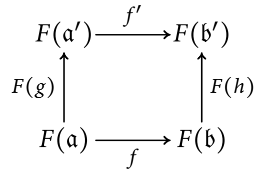


Similarly, we can define the *comma category* $(F \downarrow \mathfrak{a})$ consisting of all pairs $\langle \mathfrak{b}, f \rangle$ consisting of an object $\mathfrak{b} \in \mathcal{C}$ and a morphism $f : F(\mathfrak{b}) \to \mathfrak{a}$ of $\mathcal{D}$, where a morphism $h : \langle \mathfrak{b}, f \rangle \to \langle \mathfrak{c}, f' \rangle$ consists of a morphism $h \in \mathcal{C}^{\mathrm{mor}}$ 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

$$F(h) \circ f = f' \circ F(g)\,.$$



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 $\eta = (\eta_{\mathfrak{a}})_{\mathfrak{a} \in \mathcal{C}^{\mathrm{obj}}}$ of morphisms

$$\eta_{\mathfrak{a}} \in \mathcal{D}(F(\mathfrak{a}), G(\mathfrak{a}))\,, \qquad \text{for } \mathfrak{a} \in \mathcal{C}^{\mathrm{obj}}\,,$$

such that, for every morphism $f : \mathfrak{a} \to \mathfrak{b}$ of $\mathcal{C}$, the diagram

$$\begin{array}{ccc} F(\mathfrak{a}) & \xrightarrow{\;F(f)\;} & F(\mathfrak{b}) \\ {\scriptstyle \eta_{\mathfrak{a}}}\big\downarrow & & \big\downarrow{\scriptstyle \eta_{\mathfrak{b}}} \\ G(\mathfrak{a}) & \xrightarrow[\;G(f)\;]{} & G(\mathfrak{b}) \end{array}$$

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 $\mathcal{D}^{\mathrm{op}}$, 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) $\Rightarrow$ (1) is trivial.

(1) $\Rightarrow$ (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

$$\psi_{\mathfrak{a},\mathfrak{b}} := F \restriction \mathcal{C}(\mathfrak{a},\mathfrak{b}) : \mathcal{C}(\mathfrak{a},\mathfrak{b}) \to \mathcal{D}(F(\mathfrak{a}), F(\mathfrak{b})) , \quad \text{for } \mathfrak{a}, \mathfrak{b} \in \mathcal{C} .$$

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, $\mathrm{Cong}_w(\mathfrak{A})$ is the set of all weak congruences and

$$\mathfrak{Cong}_w(\mathfrak{A}) := \langle \mathrm{Cong}_w(\mathfrak{A}), \subseteq \rangle$$

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 $\mathrm{Cong}_w(\mathfrak{A})$ 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* $\mathfrak{Cong}_w(\mathfrak{A})$.

Proof. Let $\approx \in \mathrm{Cong}(\mathfrak{A})$ and $\sim \in \mathrm{Cong}_w(\mathfrak{A})$ 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

$$\begin{aligned}
&\big\langle f(\bar a), f(b_0, a_1, a_2, \dots, a_{n-1})\big\rangle \in E_+\,,\\
&\big\langle f(\bar a), f(b_0, b_1, a_2, \dots, a_{n-1})\big\rangle \in E_+\,,\\
&\qquad\dots\\
&\big\langle f(\bar a), f(b_0, \dots, b_{n-2}, a_{n-1})\big\rangle \in E_+\,,\\
&\big\langle f(\bar a), f(b_0, \dots, b_{n-2}, b_{n-1})\big\rangle \in E_+\,.
\end{aligned}$$

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*

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

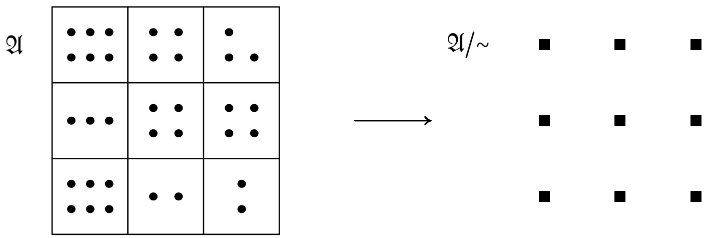
Proof. We have seen in Corollary A2.4.17 that the partial order of equivalence relations on A is complete. Consequently, the claim follows from Lemma 4.3 and Corollary A2.3.11. □

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

$$R^{\mathfrak{A}/\sim} := \big\{\, \langle [a_0]_\sim, \dots, [a_{n-1}]_\sim\rangle \;\big|\; \langle a_0, \dots, a_{n-1}\rangle \in R^{\mathfrak{A}} \,\big\}\,,$$



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

$$\bar{a} \in R^{\mathfrak{A}} \quad \text{iff} \quad h(\bar{a}) \in R^{\mathfrak{B}} \quad \text{iff} \quad h(\bar{b}) \in R^{\mathfrak{B}} \quad \text{iff} \quad \bar{b} \in R^{\mathfrak{A}}. \qquad \square$$

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*

$$\mathcal{G} : \mathfrak{Hom}_{\mathrm{s}}(\Sigma) \to \mathfrak{Hom}(\subseteq) : \mathfrak{A} \mapsto \mathfrak{Cong}(\mathfrak{A})$$

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

$$\mathcal{F}(\mathrm{id}_{\mathfrak{A}})(\sim) = \sim, \quad \text{for all } \sim\, \in \mathrm{Cong}_{\mathrm{w}}(\mathfrak{A}),$$

which implies that $\mathcal{F}(\mathrm{id}_{\mathfrak{A}}) = \mathrm{id}_{\mathfrak{Cong}_{\mathrm{w}}(\mathfrak{A})}$. Finally, if $f : \mathfrak{A} \to \mathfrak{B}$ and $g : \mathfrak{B} \to \mathfrak{C}$ are homomorphisms then we have

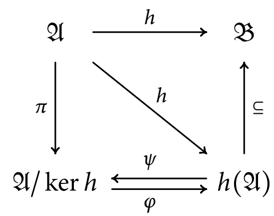
$$\begin{aligned}\mathcal{F}(g \circ f)(\sim) &= (g \circ f)^{-1}(\sim) \\ &= f^{-1}(g^{-1}(\sim)) = (\mathcal{F}(f) \circ \mathcal{F}(g))(\sim).\end{aligned}$$

(b) is shown in exactly the same way replacing 'homomorphism' by 'strict homomorphism' and 'weak congruence' by 'congruence'. $\square$

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.

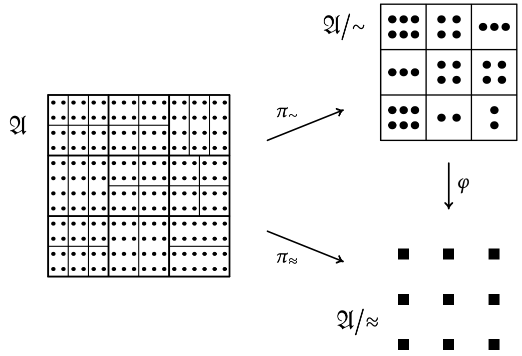
Definition 4.15. For weak congruences $\sim \subseteq \approx$ on $\mathfrak{A}$ we define

$$\approx/\sim \; := \big\{ \langle [a]_\sim, [b]_\sim \rangle \in A/\sim \times A/\sim \;\big|\; a \approx b \big\} .$$

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

$$\begin{aligned}
\varphi[a]_\sim = \varphi[b]_\sim \quad &\text{iff} \quad (\varphi \circ \pi_\sim)(a) = (\varphi \circ \pi_\sim)(b) \\
&\text{iff} \quad \pi_\approx(a) = \pi_\approx(b) \\
&\text{iff} \quad a \approx b \\
&\text{iff} \quad [a]_\sim \approx/\sim [b]_\sim .
\end{aligned}$$

□

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 $\sim \in \mathrm{Cong}(\mathfrak{A})$. The function*

$$h : \Uparrow{\sim} \to \mathrm{Cong}(\mathfrak{A}/{\sim}) \quad \textit{with} \quad h(\approx) := {\approx}/{\sim}$$

defines an isomorphism between $\mathrm{Cong}(\mathfrak{A}/{\sim})$ *and the final segment* $\Uparrow{\sim}$ *of* $\mathrm{Cong}(\mathfrak{A})$.

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 $\langle a, b\rangle \in \rho \smallsetminus \sigma$. 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 $|x| := \mathrm{dom}\, x$.

(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

$$XY := \{\, xy \mid x \in X,\ y \in Y \,\} \quad \text{and} \quad xY := \{\, xy \mid y \in Y \,\}\,.$$

(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* $\leq_{\mathrm{lex}}$ on $A^{<\alpha}$ by

$$x \leq_{\mathrm{lex}} y \quad :\text{iff} \quad x \preceq y \text{ or there are } z \in A^{<\alpha} \text{ and } a \sqsubset b \in A \text{ such that } za \preceq x \text{ and } zb \preceq y\,.$$

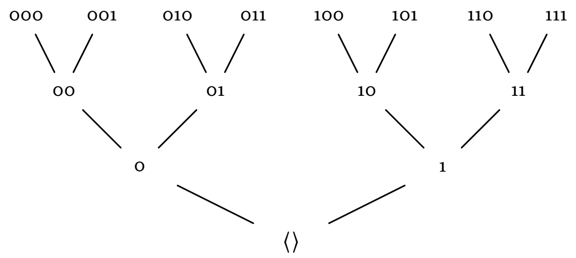


Figure 1.. $\langle 2^{<4}, \preceq \rangle$

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, $x \preceq xy$.

(b) We have $x \cdot \langle\rangle = x = \langle\rangle \cdot x$, for all $x \in A^{<\alpha}$.

(c) The prefix order $\langle 2^{<4}, \preceq \rangle$ is depicted in Figure 1, while the lexicographic ordering $\langle 2^{<4}, \leq_{\text{lex}} \rangle$ 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 $\langle 2^{<4}, \preceq \rangle$.

Exercise 1.1. Prove that $x \preceq y$ 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 $\mathfrak{T} = \langle A^{<\alpha}, \preceq \rangle$. 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

$$x_\eta := h_\gamma(u)\,, \quad \text{for the vertex } u \leq v \text{ with } |u| = \eta\,.$$ □

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 $\langle A^{<\alpha}, \preceq \rangle$ is $|A|$.

Remark. (a) Note that there are finitely branching trees of branching degree $\aleph_0$. For instance, the tree $\langle T, \preceq \rangle$ 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

$$f(X) := \{\, v \in T \mid \Uparrow v \subseteq X \,\}\,.$$

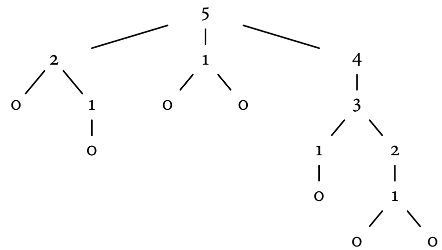
The *rank* $\mathrm{frk}(\mathfrak{T})$ of $\mathfrak{T}$ is the rank of its root.

Remark. We have $\mathrm{frk}(v) = 0$ 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:

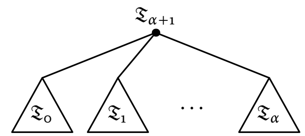
- $\mathrm{frk}(v) \geq 0$, for all $v \in T$.
- For successor ordinals, we have $\mathrm{frk}(v) \geq \alpha + 1$ if and only if there is some $u > v$ with $\mathrm{frk}(u) \geq \alpha$.
- If δ is a limit ordinal then $\mathrm{frk}(v) \geq \delta$ iff $\mathrm{frk}(v) \geq \alpha$, for all $\alpha < \delta$.

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*

$$\mathrm{frk}(u) > \mathrm{frk}(v) \quad \textit{or} \quad \mathrm{frk}(u) = \mathrm{frk}(v) = \infty\,.$$

Lemma 1.13. *Let $\mathfrak{T}$ be a tree and $v \in T$.*

(a) $\operatorname{frk}(v) = \sup \{ \operatorname{frk}(u) + 1 \mid u$ *is a successor of* $v \}$.

(b) *We have* $\operatorname{frk}(v) = \infty$ *if and only if* $\Uparrow v$ *contains an infinite path.*

Proof. (a) Let F be the fixed point induction used to define $\operatorname{frk}(v)$. If u is a successor of v then $u \in F(\operatorname{frk}(u) + 1) \setminus F(\operatorname{frk}(u))$ and $u \in \uparrow v$ implies that $v \notin F(\operatorname{frk}(u) + 1)$. Hence, $\operatorname{frk}(v) \geq \operatorname{frk}(u) + 1$. For the converse, suppose that $\operatorname{frk}(v) > \alpha$, 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 $\alpha < \operatorname{frk}(v)$, there exists some successor u with $\operatorname{frk}(u) \geq \alpha$.

(b) If $\operatorname{frk}(v) = \infty$ then (a) implies that there is some successor u of v with $\operatorname{frk}(u) = \infty$. Hence, we can inductively construct an infinite path $v = v_0 < v_1 < \dots$ such that $\operatorname{frk}(v_n) = \infty$, 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 $\operatorname{frk}(v_n) = \infty$. □

Corollary 1.14. *Let $\mathfrak{T} = \langle T, \leq \rangle$. We have* $\operatorname{frk}(\mathfrak{T}) < \infty$ *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* $\operatorname{frk}(T) < \infty$ *then* $\operatorname{frk}(T) < \kappa^+$.

Proof. Suppose, for a contradiction that $\kappa^+ \leq \operatorname{frk}(T) < \infty$. 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 $\operatorname{frk}(v) \geq \kappa^+$. Let S be the set of successors of v. By maximality and Lemma 1.13, it follows that

$$\kappa^+ = \operatorname{frk}(v) = \sup \{ \operatorname{frk}(u) + 1 \mid u \in S \},$$

where $\operatorname{frk}(u) < \kappa^+$. 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 $\bot$ 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*

$$a \sqcap b \sqsubseteq a \sqcap c \quad \textit{and} \quad a \sqcup b \sqsubseteq a \sqcup c\,.$$

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*

$$a \sqsubseteq b \quad \text{iff} \quad a \sqcap b = a$$

and $$a \sqcap a = a \qquad a \sqcup a = a \qquad \textit{(idempotence)}$$

$$a \sqcap b = b \sqcap a \qquad a \sqcup b = b \sqcup a \qquad \textit{(commutativity)}$$

$$a \sqcap (a \sqcup b) = a \qquad a \sqcup (a \sqcap b) = a \qquad \textit{(absorption)}$$

$$a \sqcap (b \sqcap c) = (a \sqcap b) \sqcap c \qquad \textit{(associativity)}$$
$$a \sqcup (b \sqcup c) = (a \sqcup b) \sqcup c$$

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

$$a \sqcap (b \sqcup c) = (a \sqcap b) \sqcup (a \sqcap c)$$
and $$a \sqcup (b \sqcap c) = (a \sqcup b) \sqcap (a \sqcup c)\,.$$

(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, $\langle \wp(A), \cap, \cup, {}^* \rangle$ forms a boolean algebra with $X^* := A \smallsetminus X$.

Remark. Note that every sublattice of a power-set lattice $\langle \wp(A), \subseteq \rangle$ 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.

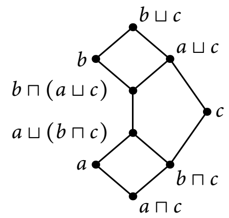


Figure 2.. The general situation

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

$$a \sqsubseteq b \quad \text{and } b \sqcap c \sqsubseteq b \qquad \text{implies} \quad a \sqcup (b \sqcap c) \sqsubseteq b\,,$$
and $$a \sqsubseteq a \sqcup c \text{ and } b \sqcap c \sqsubseteq c \sqsubseteq a \sqcup c \quad \text{implies} \quad a \sqcup (b \sqcap c) \sqsubseteq a \sqcup c\,. \square$$

In general the distributive laws also hold only in one direction.

Lemma 2.7. *In every lattice $\mathfrak{L}$, we have*

$$a \sqcap (b \sqcup c) \sqsupseteq (a \sqcap b) \sqcup (a \sqcap c)$$
and $$a \sqcup (b \sqcap c) \sqsubseteq (a \sqcup b) \sqcap (a \sqcup c)\,,$$

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

$$a \sqcup (b \sqcap c) = (a \sqcup b) \sqcap (a \sqcup c) = b \sqcap (a \sqcup c)\,. \qquad \square$$

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

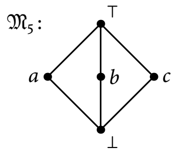
$$b \sqcap c \sqsubseteq a \sqcup (b \sqcap c) \sqsubseteq x \sqsubseteq b\,,$$

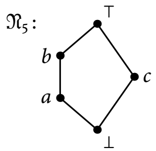
which implies that $x \sqcap c = b \sqcap c$. Hence,

$$a \sqcup (b \sqcap c) = a \sqcup (x \sqcap c) = x\,. \qquad \square$$

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

$$\begin{aligned} a &:= x \sqcup (y \sqcap z)\,, \qquad & d &:= b \sqcup z\,, \\ b &:= y \sqcap (x \sqcup z)\,, & e &:= a \sqcap z\,. \end{aligned}$$

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 $\mathfrak{M}_5 \to \mathfrak{L}$.

(⇒) 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.

(⇐) 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 $\langle \wp(A), \subseteq \rangle$. 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* $\mathfrak{L}^{op} = \langle L, \sqcup, \sqcap, \sqsupseteq \rangle$ where the order is reversed. Obviously, this functions maps filters of $\mathfrak{L}$ to ideals of $\mathfrak{L}^{op}$ 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

$$\mathrm{cl}_\downarrow(X) := \{ b \in L \mid b \sqsubseteq a_0 \sqcup \dots \sqcup a_n \text{ for some } a_0, \dots, a_n \in X, n < \omega \},$$
$$\mathrm{cl}_\uparrow(X) := \{ b \in L \mid b \sqsupseteq a_0 \sqcap \dots \sqcap a_n \text{ for some } a_0, \dots, a_n \in X, n < \omega \}.$$

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 $\mathfrak{a} \cap \mathfrak{u} = \emptyset$ then the set*

$$\mathcal{I} := \{ \mathfrak{b} \mid \mathfrak{b} \text{ a proper ideal with } \mathfrak{a} \subseteq \mathfrak{b} \text{ and } \mathfrak{b} \cap \mathfrak{u} = \emptyset \}$$

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 $\mathfrak{a} \cap \mathfrak{u} = \emptyset$, for all $\mathfrak{a} \in \mathcal{I}$, we have $\mathfrak{c} \cap \mathfrak{u} = \emptyset$. 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 $a_{n+1} \in \mathfrak{a} \setminus \Downarrow(a_0 \sqcup \dots \sqcup a_n) \neq \emptyset$. 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:

$$\mathfrak{u}_n := \Uparrow\{n\} , \qquad \text{for } n < \omega ,$$
$$\mathfrak{u}_\infty := \{ X \subseteq \omega \mid \omega \setminus X \text{ is finite} \} .$$

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

$$\begin{aligned} h(x) \sqcup h(y) &= \langle h_0(x), h_1(x) \rangle \sqcup \langle h_0(y), h_1(y) \rangle \\ &= \langle h_0(x) \sqcup h_0(y),\ h_1(x) \sqcup h_1(y) \rangle = h(x \sqcup y) , \end{aligned}$$

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 $x \in \mathfrak{b} \smallsetminus \mathfrak{a}$. 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 $\langle \wp(A), \subseteq \rangle$.

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 $\wp(A)$, there are at most $|\wp(\wp(A))| = 2^{2^\kappa}$ 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

$$\left| X_0 \cap \dots \cap X_{m-1} \cap (A \smallsetminus Y_0) \cap \dots \cap (A \smallsetminus Y_{n-1}) \right| = |A| .$$

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

$$S_K := K \cup \{ A \smallsetminus X \mid X \in F \smallsetminus K \} .$$

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 $X \in K \smallsetminus L$. Then $X \in S_K \subseteq \mathfrak{u}_K$ and $A \smallsetminus X \in S_L \subseteq \mathfrak{u}_L$. 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 $a \in s \smallsetminus t$, it follows that

$$\langle \{a\}, \{\{a\}\} \rangle \in P_s \smallsetminus P_t , \quad \text{which implies that} \quad P_s \neq P_t .$$

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

$$\big|P_{s_0} \cap \dots \cap P_{s_{m-1}} \cap (W \setminus P_{t_0}) \cap \dots \cap (W \setminus P_{t_{n-1}})\big| \geq |Q| = \kappa = |W|\,.$$

□

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

$$\operatorname{ht}(a) := \sup\{\, |C| \mid C \subseteq {\downarrow} a \text{ is a chain} \,\}\,.$$

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.

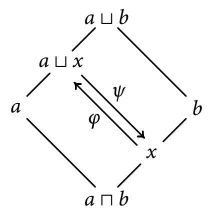
The notion of height is mainly meaningful for modular lattices where it is well-behaved, at least for elements of finite height.

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. □

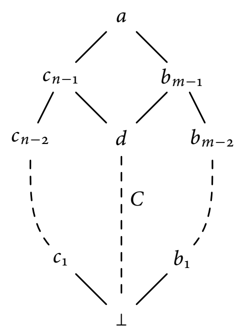


Figure 3.. Proof of Lemma 5.3

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 $b_0 \sqsubset \dots \sqsubset b_m$ is a maximal chain with $b_m = a$, then $m = n$. Since a has height n, there exists a chain $c_0 \sqsubset \dots \sqsubset c_n$ 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 $|C| = \operatorname{ht}(d) + 1 < \operatorname{ht}(c_{n-1}) + 1 = n$.

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 $\mathfrak{I}_0 := \langle I_0, \sqsubseteq \rangle$ and $\mathfrak{I}_1 := \langle I_1, \sqsubseteq \rangle$ are modular lattices and, by Lemma 5.2, there exists an isomorphism $\varphi : \mathfrak{I}_0 \to \mathfrak{I}_1$. By Lemma 5.4, the height of the top element of $\mathfrak{I}_0$ is $\mathrm{ht}(a) - \mathrm{ht}(a \sqcap b) + 1$ and the height of the top element of $\mathfrak{I}_1$ is $\mathrm{ht}(a \sqcup b) - \mathrm{ht}(b) + 1$. Since $\mathfrak{I}_0 \cong \mathfrak{I}_1$ 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 $c_0 \sqsubset \cdots \sqsubset c_n$ 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 $c_0 \sqsubset \cdots \sqsubset c_n$ 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

$$\mathrm{ht}(X) = \begin{cases} |X| & \text{if } X \text{ is finite}, \\ \aleph_0 & \text{otherwise}. \end{cases}$$

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:

- $\mathrm{rk_P}(a) = -1$ iff $a = \perp$.
- $\mathrm{rk_P}(a) \geq 0$ iff $a \neq \perp$.
- $\mathrm{rk_P}(a) \geq \alpha + 1$ iff there exists an infinite partition P of a such that $\mathrm{rk_P}(p) \geq \alpha$, for all $p \in P$.
- For limit ordinals δ, we set $\mathrm{rk_P}(a) \geq \delta$ iff $\mathrm{rk_P}(a) \geq \alpha$, 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 $\mathrm{rk_P}(a) = 1$.

Lemma 5.10. $a \sqsubseteq b$ *implies* $\mathrm{rk_P}(a) \leq \mathrm{rk_P}(b)$.

Lemma 5.11. *If $\mathfrak{L}$ is a distributive lattice then*

$$\mathrm{rk_P}(a \sqcup b) = \max\{\mathrm{rk_P}(a), \mathrm{rk_P}(b)\}.$$

Proof. By the preceding lemma, we have $\mathrm{rk_P}(a \sqcup b) \geq \mathrm{rk_P}(a), \mathrm{rk_P}(b)$. It remains to show that $\mathrm{rk_P}(a \sqcup b) \geq \alpha$ implies $\mathrm{rk_P}(a) \geq \alpha$ or $\mathrm{rk_P}(b) \geq \alpha$. 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 $\mathrm{rk_P}(a \sqcup b) \geq \alpha + 1$. Then there exists an infinite partition P of $a \sqcup b$ such that $\mathrm{rk_P}(p) \geq \alpha$, 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

$$P_a := \{ p \in P \mid \mathrm{rk_P}(a_p) \geq \alpha \}$$

and $$P_b := \{ p \in P \mid \mathrm{rk_P}(b_p) \geq \alpha \}.$$

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*

$$\mathrm{rk_P}(a) \leq \mathrm{rk_P}(h(a)), \quad \text{for all } a \in A.$$

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 $\perp = h(a) \sqcap h(b)^* = h(a \sqcap b^*)$. Since h is injective it follows that $a \sqcap b^* = \perp$. 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

$$\perp \sqsubset a_w \sqsubset a_u \quad \text{for all } u \prec w\,,$$
$$a_u \sqcap a_w = \perp \quad \text{for all } u, w \text{ with } u \npreceq w \text{ and } w \npreceq u\,.$$

(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) $\mathrm{rk}_{\mathrm{P}}(a) = \infty$.

(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) ⇒ (2) is trivial.

(1) ⇒ (3) Let $\kappa := |L|^+$. We construct the family $(b_w)_w$ by induction on w such that $\mathrm{rk}_{\mathrm{P}}(b_w) = \infty$. We start with $b_{\langle\rangle} = a$. If b_w has been defined then $\mathrm{rk}_{\mathrm{P}}(b_w) \geq \kappa + 1$ implies that there exists an infinite partition P of b_w with $\mathrm{rk}_{\mathrm{P}}(p) \geq \kappa$, for all $p \in P$. By Lemma 5.13, it follows that $\mathrm{rk}_{\mathrm{P}}(p) = \infty$, 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 $\mathrm{rk}_{\mathrm{P}}(b_{wi}) = \infty$, as desired.

(2) ⇒ (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 $\mathrm{rk}_{\mathrm{P}}(b_w) \geq \alpha$, for all w. Since $b_{w0} \sqsubset b_w$ we have $b_w \neq \perp$ and $\mathrm{rk}_{\mathrm{P}}(b_w) \geq 0$. For limit ordinals, the claim follows immediately from the inductive hypothesis. Hence, it remains to consider the successor step. Suppose that $\mathrm{rk}_{\mathrm{P}}(b_w) \geq \alpha$, 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, $\mathrm{rk}_{\mathrm{P}}(b_w) \geq \alpha + 1$. □

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*

$$a = \sup P \quad \text{and} \quad \mathrm{rk}_{\mathrm{P}}(p) = \mathrm{rk}_{\mathrm{P}}(a)\,, \qquad \text{for all } p \in P\,.$$

Furthermore, if Q is any other partition of a with

$$\mathrm{rk}_{\mathrm{P}}(q) = \mathrm{rk}_{\mathrm{P}}(a)\,, \qquad \text{for all } q \in Q\,,$$

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 $w_n := \beta \restriction n$ 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

$$c_k \sqcap c_n = b_{w_k} \sqcap b^*_{w_{k+1}} \sqcap b_{w_n} \sqcap b^*_{w_{n+1}} = \perp\,.$$

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 $\mathrm{rk}_{\mathrm{P}}(p_n) = \alpha$, $p_k \sqcap p_n = \perp$, 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* $\deg_\mathrm{P}(a)$ 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 $\deg_\mathrm{P}(a) := \infty$.

(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

$$\deg_\mathrm{P}(\mathfrak{u}) := \min \{ \deg_\mathrm{P}(a) \mid a \in \mathfrak{u} \text{ with } \mathrm{rk_P}(a) = \mathrm{rk_P}(\mathfrak{u}) \}.$$

We say that an element $a \in \mathfrak{u}$ has *minimal rank and degree* if

$$\mathrm{rk_P}(a) = \mathrm{rk_P}(\mathfrak{u}) \quad \text{and} \quad \deg_\mathrm{P}(a) = \deg_\mathrm{P}(\mathfrak{u}).$$

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

$$\deg_\mathrm{P}(X) = \begin{cases} |X| & \text{if } X \text{ is finite}, \\ 1 & \text{otherwise}. \end{cases}$$

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

$$\begin{aligned} \mathrm{rk_P}(\mathfrak{u}_a) &= 0 & \deg_\mathrm{P}(\mathfrak{u}_a) &= 1, \\ \mathrm{rk_P}(\mathfrak{u}_\infty) &= 1 & \deg_\mathrm{P}(\mathfrak{u}_\infty) &= 1. \end{aligned}$$

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. $\mathrm{rk}_P(a \sqcap c) = \mathrm{rk}_P(a) = \mathrm{rk}_P(a \sqcap c^*) < \infty$ *implies that* $\deg_P(a \sqcap c) < \deg_P(a)$.

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* $\mathrm{rk}_P(\mathfrak{u}), \mathrm{rk}_P(\mathfrak{v}) < \infty$. *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 $c \in \mathfrak{u} \setminus \mathfrak{v}$. It follows that $a \sqcap c \in \mathfrak{u}$ and

$$\mathrm{rk}_P(a \sqcap c) \leq \mathrm{rk}_P(a) = \mathrm{rk}_P(\mathfrak{u}) .$$

Since a is of minimal rank we therefore have

$$\mathrm{rk}_P(a \sqcap c) = \mathrm{rk}_P(a) .$$

Analogously, we can conclude that

$$\mathrm{rk}_P(b \sqcap c^*) = \mathrm{rk}_P(b) .$$

If $a = b$ then it would follow that

$$\mathrm{rk}_P(a \sqcap c) = \mathrm{rk}_P(a) = \mathrm{rk}_P(a \sqcap c^*) .$$

This implies that $\deg_P(a \sqcap c) < \deg_P(a)$ 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* $\mathrm{rk}_P(\mathfrak{u}) < \infty$.

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 $\mathrm{dom}\, t := T$. 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 $t : T \to \Sigma_{\mathrm{fun}} \cup X$ satisfying the following properties:

- All inner vertices $v \in \mathrm{dom}\, t$ are labelled by elements of Σ_{fun}.
- If the function symbol $t(v) = f \in \Sigma_{\mathrm{fun}}$ 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 $t(u_i) = g \in \Sigma_{\mathrm{fun}}$ 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 $\bar{r}$.

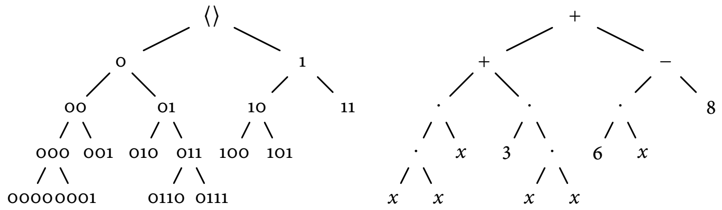


Figure 1.. Domain and labelling of t.

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

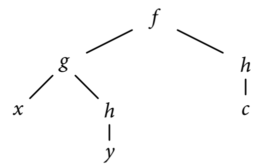
$$y(t) := t(x_0) \cdots t(x_n)$$

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

$$s_1 = t_1\,,\ \dots,\ s_{n-1} = t_{n-1} \quad \text{and} \quad u = v\,.$$ □

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. □

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

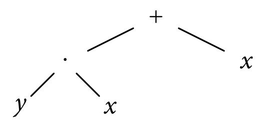
$$\mathrm{free}(t) := \mathrm{rng}\, t \cap X\,.$$

(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 $\mathrm{free}(t) \subseteq X_0 \subseteq X$. 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*

$$\beta \restriction \mathrm{free}(t) = \gamma \restriction \mathrm{free}(t)$$

then $t^{\mathfrak{A}}[\beta] = t^{\mathfrak{A}}[\gamma]$.

Remark. We write $t(x_0, \dots, x_{n-1})$ to indicate that

$$\operatorname{free}(t) \subseteq \{x_0, \dots, x_{n-1}\}.$$

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 $\operatorname{rng} h = \langle\!\langle \operatorname{rng} \beta \rangle\!\rangle_{\mathfrak{A}}$.

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

$$\begin{aligned} h(f t_0 \dots t_{n-1}) &= (f t_0 \dots t_{n-1})^{\mathfrak{A}}[\beta] \\ &= f^{\mathfrak{A}}(t_0^{\mathfrak{A}}[\beta], \dots, t_{n-1}^{\mathfrak{A}}[\beta]) \\ &= f^{\mathfrak{A}}(h(t_0), \dots, h(t_{n-1})), \end{aligned}$$

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

$$\begin{aligned} g(f t_0 \dots t_{n-1}) &= f^{\mathfrak{A}}(g(t_0), \dots, g(t_{n-1})) \\ &= f^{\mathfrak{A}}(h(t_0), \dots, h(t_{n-1})) = h(f t_0 \dots t_{n-1}). \end{aligned}$$

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

$$h(f t_0 \dots t_{n-1}) = f^{\mathfrak{A}}(h(t_0), \dots, h(t_{n-1})) = f^{\mathfrak{A}}(a_0, \dots, a_{n-1}).$$

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. $\square$

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

$$\mathfrak{Hom}_s(\Sigma)(\mathfrak{T}[\Sigma, X], \mathfrak{A}) \to \mathfrak{Set}_S(X, A) : h \mapsto h \restriction X,$$

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 $h \restriction X = \mathrm{id}_X$.*

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 $s = f s_0 \dots s_{n-1}$ then we have by inductive hypothesis

$$\begin{aligned} (f s_0 \dots s_{n-1})[x/t]^{\mathfrak{A}}[\beta] &= f^{\mathfrak{A}}(s_0[x/t]^{\mathfrak{A}}[\beta], \dots, s_{n-1}[x/t]^{\mathfrak{A}}[\beta]) \\ &= f^{\mathfrak{A}}(s_0^{\mathfrak{A}}[\beta'], \dots, s_{n-1}^{\mathfrak{A}}[\beta']) \\ &= (f s_0 \dots s_{n-1})^{\mathfrak{A}}[\beta'] . \end{aligned}$$ □

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 $\varphi(\circ) = +, \quad \varphi({}^{-1}) = -, \quad \varphi(e) = 0, \quad \text{and} \quad \psi(x) = y.$

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

$$R^{\mathfrak{B}} = \big\{\, \bar a \in B^n \bigm| \bar a^i \in R^{\mathfrak{A}^i} \text{ for all } i \in I \,\big\}\,,$$

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 $\mathfrak{u} \subseteq \mathscr{P}(I)$ 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 $\qquad \bar a \sim_{\mathfrak{u}} \bar b \quad :\text{iff} \quad [\![\bar a^i = \bar b^i]\!]_i \in \mathfrak{u}\,.$

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 $\mathfrak{u} \subseteq \mathscr{P}(I)$ 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 $\mathfrak u \subseteq \wp(I)$ 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

$$R^{\mathfrak B} := \big\{\, [\bar a]_{\mathfrak u} \in B^n \;\big|\; [\![\bar a^i \in R]\!]_i \in \mathfrak u \,\big\}\,,$$

and, for each function $f \in \Sigma$,

$$f^{\mathfrak B}\big([\bar a]_{\mathfrak u}\big) := [(b_i)_i]_{\mathfrak u} \quad \text{where} \quad b_i := f^{\mathfrak A^i}(\bar a^i)\,.$$

(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*

$$s^{\mathfrak B}[\beta] = t^{\mathfrak B}[\beta] \quad \text{iff} \quad \big\{\, i \in I \;\big|\; s^{\mathfrak A^i}[\beta_i] = t^{\mathfrak A^i}[\beta_i] \,\big\} \in \mathfrak u\,.$$

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*

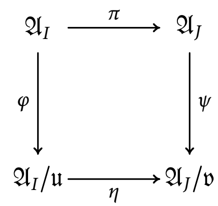
$$\prod_{i\in I} \mathfrak{A}^i/\mathfrak{u} \cong \prod_{j\in J} \mathfrak{A}^j/\mathfrak{u}|_J\,.$$

Proof. To simplify notation set $\mathfrak{v} := \mathfrak{u}|_J$ and define

$$\begin{aligned} &\mathfrak{A}_I := \textstyle\prod_{i\in I} \mathfrak{A}^i\,, \qquad &&\mathfrak{A}_I/\mathfrak{u} := \textstyle\prod_{i\in I} \mathfrak{A}^i/\mathfrak{u}\,, \\ \text{and}\quad &\mathfrak{A}_J := \textstyle\prod_{j\in J} \mathfrak{A}^j\,, &&\mathfrak{A}_J/\mathfrak{v} := \textstyle\prod_{j\in J} \mathfrak{A}^j/\mathfrak{v}\,. \end{aligned}$$

For sequences $(\bar{a}^i)_{i\in I}$ set $\bar{a} \upharpoonright J := (\bar{a}^j)_{j\in J}$. Let

$$\begin{aligned} &\varphi : \mathfrak{A}_I \to \mathfrak{A}_I/\mathfrak{u} : (a^i)_i \mapsto [(a^i)_i]_{\mathfrak{u}} \\ &\psi : \mathfrak{A}_J \to \mathfrak{A}_J/\mathfrak{v} : (a^j)_j \mapsto [(a^j)_j]_{\mathfrak{v}} \\ &\pi : \mathfrak{A}_I \to \mathfrak{A}_j : \bar{a} \mapsto \bar{a} \upharpoonright J \end{aligned}$$



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.,

$$\eta([\bar{a}]_{\mathfrak{u}}) = [\bar{a} \upharpoonright J]_{\mathfrak{v}}\,.$$

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

$$\begin{aligned} \eta\big(f^{\mathfrak{A}_I/\mathfrak{u}}([\bar{a}]_{\mathfrak{u}})\big) &= \eta\big([f^{\mathfrak{A}_I}(\bar{a})]_{\mathfrak{u}}\big) \\ &= \big[f^{\mathfrak{A}_J}(\bar{a} \upharpoonright J)\big]_{\mathfrak{v}} \\ &= f^{\mathfrak{A}_J/\mathfrak{v}}([\bar{a} \upharpoonright J]_{\mathfrak{v}}) = f^{\mathfrak{A}_J/\mathfrak{v}}\big(\eta([\bar{a}]_{\mathfrak{u}})\big)\,, \end{aligned}$$

and, for a relation symbol R, we have

$$\begin{aligned} [\bar{a}]_{\mathfrak{u}} \in R^{\mathfrak{A}_I/\mathfrak{u}} \quad &\text{iff} \quad [\![\bar{a}^i \in R]\!]_i \in \mathfrak{u} \\ &\text{iff} \quad [\![\bar{a}^i \in R]\!]_i \cap J \in \mathfrak{u} \\ &\text{iff} \quad \eta([\bar{a}]_{\mathfrak{u}}) = [\bar{a} \upharpoonright J]_{\mathfrak{v}} \in R^{\mathfrak{A}_J/\mathfrak{v}}. \end{aligned}$$

□

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 $\mathfrak{J} = \langle I, \leq \rangle$ *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 $\Uparrow i \cap D \neq \varnothing$, 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 $\omega \leq \alpha < \kappa$. $\square$

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. $\square$

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 $X_\alpha := \bigcup H_\alpha \in \operatorname{fix} c$. Since $(X_\alpha)_{\alpha<\kappa}$ is a chain it follows that $\bigcup I = \bigcup_\alpha X_\alpha \in \operatorname{fix} c$, 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 $\bigcup I \in \operatorname{fix} c$. 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

$$\sup f[I] = \sup f[C] = f(\sup C) = f(\sup I) .$$ □

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

$$\begin{aligned} \langle i, a \rangle \sim \langle j, b \rangle \quad &:\text{iff} \quad D(i,k)(a) = D(j,k)(b) \\ &\qquad \text{for some } k \geq i, j . \end{aligned}$$

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

$$R := \{ \langle [i,a_0], \dots, [i,a_{n-1}] \rangle \mid \langle a_0, \dots, a_{n-1} \rangle \in R^{D(i)} \} ,$$

$$\text{and} \quad f([i,a_0], \dots, [i,a_{n-1}]) := [i, f^{D(i)}(a_0, \dots, a_{n-1})] .$$

(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 $\mathfrak{J}$ has a greatest element k, then we have $\varinjlim D \cong D(k)$.

Exercise 3.1. Let $D : \mathfrak{J} \to \mathfrak{Hom}(\Sigma)$ be a directed diagram and $S \subseteq I$ dense. Prove that

$$\varinjlim D \cong \varinjlim\,(D \restriction S)\,,$$

where $D \restriction S : \mathfrak{J}|_S \to \mathfrak{Hom}(\Sigma)$ is the restriction of D to S.

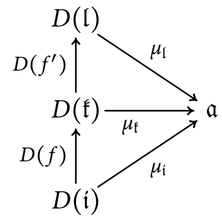
Directed colimits can also be characterised in category-theoretical terms via so-called limiting cocones. We use this property to define directed colimits in an arbitrary category.

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 $\mu = (\mu_{\mathfrak{i}})_{\mathfrak{i} \in \mathcal{I}^{\mathrm{obj}}}$ of morphisms $\mu_{\mathfrak{i}} : D(\mathfrak{i}) \to \mathfrak{a}$ such that

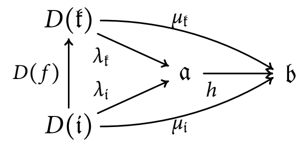
$$\mu_{\mathfrak{k}} \circ D(f) = \mu_{\mathfrak{i}}\,,$$

for all $f : \mathfrak{i} \to \mathfrak{k}$ in $\mathcal{I}^{\mathrm{mor}}$.



(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.

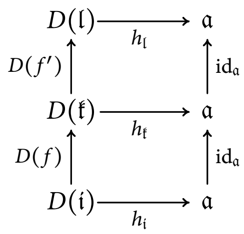
Example. Let $\mathfrak{L}$ be a partial order and $D : \mathcal{I} \to \mathfrak{L}$ a diagram.

(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

$$\Delta(\mathfrak{a})(\mathfrak{i}) = \mathfrak{a}\,, \qquad \text{for all } \mathfrak{i} \in \mathcal{I}^{\mathrm{obj}},$$
$$\text{and} \quad \Delta(\mathfrak{a})(f) = \mathrm{id}_{\mathfrak{a}}\,, \qquad \text{for all } f \in \mathcal{I}^{\mathrm{mor}}.$$



(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

$$\lambda_i : D(i) \to \mathfrak{A} : a \mapsto [i, a]\,, \quad \text{for } i \in I\,,$$

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

$$\begin{aligned} h[i, a] = \mu_i(a) &= (\mu_k \circ D(i,k))(a) \\ &= (\mu_k \circ D(j,k))(b) = \mu_j(b) = h[j, b]\,. \end{aligned} \qquad \Box$$

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

$$\mathfrak{A} \cong \varinjlim D\,. \qquad \Box$$

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

$$\varprojlim D := \big(\textstyle\prod_i \mathfrak{A}^i\big)\big|_U$$

obtained from the product of the $\mathfrak{A}^i$ by restriction to the set

$$U := \big\{ (a_i)_i \in \textstyle\prod_i A^i \;\big|\; a_i = D(i,j)(a_j) \text{ for all } i \le j \big\} \,.$$

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,

$$\varprojlim D \cong D(k)|_C$$

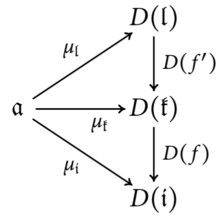
where $C := \bigcap_i D(i)$ and $k \in I$ is arbitrary.
(b) Let $\mathfrak{K}$ be a field and $D(n) := \mathfrak{K}[x]/(x^n)$, for $n < \omega$, the ring of polynomials over $\mathfrak{K}$ of degree less than n. The directed limit $\varprojlim D \cong \mathfrak{K}[[x]]$ is isomorphic to the ring of formal 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

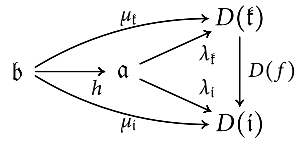
$$D(f) \circ \mu_{\mathfrak{k}} = \mu_{\mathfrak{i}} \,,$$

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

$$f * \mu := (f \circ \mu_{\mathfrak{i}})_{\mathfrak{i} \in \mathcal{I}} .$$

(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, $F(\mu_i) = F(\mu'_i)$ implies that $\mu_i = \mu'_i$, for all $i \in \mathcal{I}$.

For surjectivity, let $\mu \in \mathrm{Cone}(F \circ D, F(\mathfrak{a}))$. As F is full, we can find morphisms $\lambda_i : D(i) \to \mathfrak{a}$, for every $i \in \mathcal{I}$, such that $F(\lambda_i) = \mu_i$. Then $F[\lambda] = \mu$ where $\lambda := (\lambda_i)_{i \in \mathcal{I}}$. Hence, it remains to prove that λ is a cocone of D. Let $f : i \to j$ be a morphism of $\mathcal{I}$. Then

$$F(\lambda_j \circ D(f)) = F(\lambda_j) \circ F(D(f)) = \mu_j \circ F(D(f)) = \mu_i = F(\lambda_i)$$

implies, by faithfulness of F, that $\lambda_j \circ D(f) = \lambda_i$.

(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*

$$\zeta : \mathrm{Cone}(D, -) \cong \mathrm{Cone}(E, -) : \mu \mapsto (\mu_i \circ \eta_i^{-1})_{i \in \mathcal{I}}\,.$$

Proof. We define ζ and its inverse ξ by

$$\begin{aligned}\zeta(\mu) &:= (\mu_i \circ \eta_i^{-1})_{i \in \mathcal{I}}\,, &&\text{for } \mu \in \mathrm{Cone}(D, \mathfrak{a})\,, \\ \xi(\mu) &:= (\mu_i \circ \eta_i)_{i \in \mathcal{I}}\,, &&\text{for } \mu \in \mathrm{Cone}(E, \mathfrak{a})\,.\end{aligned}$$

To show that ζ and ξ are well-defined, let $\mu \in \mathrm{Cone}(D, \mathfrak{a})$ and let $f : i \to j$ be a morphism of $\mathcal{I}$. Then

$$\begin{aligned}\zeta(\mu)_j \circ E(f) &= \mu_j \circ \eta_j^{-1} \circ E(f) \\ &= \mu_j \circ D(f) \circ \eta_i^{-1} = \mu_i \circ \eta_i^{-1} = \zeta(\mu)_i\,.\end{aligned}$$

Hence, $\zeta(\mu)$ is a cocone of E. In the same way, one can check that

$$\xi(\mu)_j \circ D(f) = \xi(\mu)_i\,, \quad \text{for } \mu \in \mathrm{Cone}(E, \mathfrak{a}) \text{ and } f : i \to j\,.$$

Furthermore, ζ is a natural transformation since, for $\mu \in \mathrm{Cone}(D, \mathfrak{a})$ and $f : \mathfrak{a} \to \mathfrak{b}$,

$$\begin{aligned}\zeta(\mathrm{Cone}(D,f)(\mu)) &= \zeta((f \circ \mu_i)_{i\in\mathcal{I}}) \\ &= (f \circ \mu_i \circ \eta_i^{-1})_{i\in\mathcal{I}} \\ &= \mathrm{Cone}(E,f)((\mu_i \circ \eta_i^{-1})_{i\in\mathcal{I}}) \\ &= \mathrm{Cone}(E,f)(\zeta(\mu)) .\end{aligned}$$

Finally, note that

$$\xi(\zeta(\mu)) = \xi((\mu_i \circ \eta_i^{-1})_{i\in\mathcal{I}}) = (\mu_i \circ \eta_i^{-1} \circ \eta_i)_{i\in\mathcal{I}} = \mu ,$$

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.

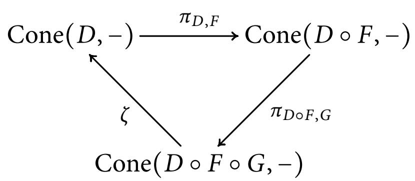
Proof. By Theorem B1.3.14, there exist a functor $G : \mathcal{J} \to \mathcal{I}$ and natural isomorphisms $\rho : G \circ F \cong \mathrm{id}_{\mathcal{I}}$ and $\eta : \mathrm{id}_{\mathcal{J}} \cong F \circ G$ such that

$$F(\rho_i) = \eta_{F(i)}^{-1} \quad \text{and} \quad G(\eta_j) = \rho_{G(j)}^{-1}.$$

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

$$\zeta : \mathrm{Cone}(D \circ F \circ G,-) \to \mathrm{Cone}(D,-) : \mu \mapsto (\mu_j \circ D(\eta_j))_{j\in\mathcal{J}} .$$

We claim that $\zeta \circ \pi_{D\circ F,G}$ is an inverse of $\pi_{D,F}$.



For $\mu \in \mathrm{Cone}(D,\mathfrak{a})$, $\mu_{F(G(j))} \circ D(\eta_j) = \mu_j$ implies that

$$\begin{aligned}(\zeta \circ \pi_{D\circ F,G} \circ \pi_{D,F})(\mu) &= (\zeta \circ \pi_{D\circ F,G})\big((\mu_{F(i)})_{i\in\mathcal{I}}\big) \\ &= \zeta\big((\mu_{F(G(j))})_{j\in\mathcal{J}}\big) \\ &= \big(\mu_{F(G(j))} \circ D(\eta_j)\big)_{j\in\mathcal{J}} = (\mu_j)_{j\in\mathcal{J}} .\end{aligned}$$

Similarly, let $\mu \in \mathrm{Cone}(D \circ F,\mathfrak{a})$. Then $\mu_i \circ D(F(\rho_i)) = \mu_{G(F(i))}$ implies that

$$\begin{aligned}(\pi_{D,F} \circ \zeta \circ \pi_{D\circ F,G})(\mu) &= (\pi_{D,F} \circ \zeta)\big((\mu_{G(j)})_{j\in\mathcal{J}}\big) \\ &= \pi_{D,F}\big((\mu_{G(j)} \circ D(\eta_j))_{j\in\mathcal{J}}\big) \\ &= \big(\mu_{G(F(i))} \circ D(\eta_{F(i)})\big)_{i\in\mathcal{I}} \\ &= \big(\mu_{G(F(i))} \circ D(F(\rho_i)^{-1})\big)_{i\in\mathcal{I}} \\ &= (\mu_i)_{i\in\mathcal{I}} .\end{aligned}$$

□

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 $\mathcal{C}_o \subseteq \mathcal{C}$ such that

- every object of $\mathcal{C}$ is isomorphic to some object of $\mathcal{C}_o$,
- no two objects of $\mathcal{C}_o$ 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 $\mathcal{C}_0$. 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

$$I(Q(f)) \circ \rho_\mathfrak{a} = \pi_\mathfrak{b} \circ f \circ \pi_\mathfrak{a}^{-1} \circ \pi_\mathfrak{a} = \rho_\mathfrak{a} \circ f\,.$$

□

By Corollary 4.10, we obtain the following result.

Corollary 4.14. *Let $\mathcal{I}_0 \subseteq \mathcal{I}$ be a skeleton of $\mathcal{I}$ and $F : \mathcal{I}_0 \to \mathcal{I}$ the inclusion functor. Then*

$$\varinjlim D = \varinjlim (D \circ F)\,, \quad \text{for every diagram } D : \mathcal{I} \to \mathcal{C}\,.$$

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, $D : \mathfrak{J} \to \mathcal{C}$ 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 $\alpha < \kappa$,

$$C(\alpha) = \varinjlim (D \restriction H_\alpha)\,, \quad \text{for some directed subset } H_\alpha \subseteq I \text{ of size } |H_\alpha| < |I|\,.$$

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

$$\begin{aligned} \pi_\alpha &: \mathrm{Cone}(D, -) \to \mathrm{Cone}(D \restriction H_\alpha, -)\,, \\ \pi_{\alpha,\beta} &: \mathrm{Cone}(D \restriction H_\beta, -) \to \mathrm{Cone}(D \restriction H_\alpha, -)\,, \end{aligned}$$

be the projections along the inclusion functors $H_\alpha \to I$ and $H_\alpha \to H_\beta$, respectively. We define C^{obj} by

$$C(\alpha) := \varinjlim (D \upharpoonright H_\alpha)\,, \quad \text{for } \alpha < \kappa\,.$$

To define C^{mor}, let $\alpha < \beta$. Since λ^α is limiting and $\pi_{\alpha,\beta}(\lambda^\beta)$ is a cocone of $D \upharpoonright H_\alpha$, there exists a unique morphism

$$C(\alpha,\beta) : \varinjlim (D \upharpoonright H_\alpha) \to \varinjlim (D \upharpoonright H_\beta)\,,$$

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

$$\tau_\alpha : \mathrm{Cone}(D \upharpoonright H_\alpha, -) \cong \mathcal{C}(C(\alpha), -)\,, \quad \text{for } \alpha < \kappa\,,$$

such that

$$\begin{aligned} \mu &= \tau_\alpha(\mu) * \lambda^\alpha\,, && \text{for cocones } \mu \text{ of } D \upharpoonright H_\alpha\,,\\ f &= \tau_\alpha(f * \lambda^\alpha)\,, && \text{for all } f : C(\alpha) \to \mathfrak{a}\,. \end{aligned}$$

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

$$\begin{aligned}\eta(\zeta(\mu)) &= \eta\big((\mu_{\alpha_i} \circ \lambda_i^{\alpha_i})_{i \in I}\big) \\ &= \big(\tau_\beta\big(\pi_\beta\big((\mu_{\alpha_i} \circ \lambda_i^{\alpha_i})_{i \in I}\big)\big)\big)_{\beta < \kappa} \\ &= \big(\tau_\beta\big((\mu_{\alpha_i} \circ \lambda_i^{\alpha_i})_{i \in H_\beta}\big)\big)_{\beta < \kappa} \\ &= \big(\tau_\beta\big((\mu_\beta \circ \lambda_i^{\beta})_{i \in H_\beta}\big)\big)_{\beta < \kappa} \\ &= \big(\tau_\beta(\mu_\beta * \lambda^\beta)\big)_{\beta < \kappa} = (\mu_\beta)_{\beta < \kappa}\,. \qquad \square\end{aligned}$$

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 $\varinjlim D = \varinjlim C \in \mathcal{K}$. $\square$

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.

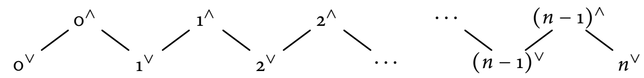
Definition 5.1. Let $\mathcal{C}$ be a category.

(a) For $n < \omega$, we denote by $\mathfrak{Z}_n = \langle Z_n, \leq \rangle$ the partial order on the elements

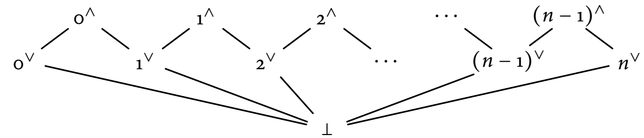
$$Z_n := \{0^\vee, \dots, n^\vee, 0^\wedge, \dots, (n-1)^\wedge\}$$

that is defined by

$$x < y \quad :\text{iff} \quad x = i^\vee \text{ and } y = k^\wedge \text{ for } k \leq i \leq k+1\,.$$



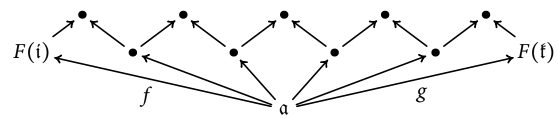
And we write $\mathfrak{Z}_n^\perp$ for the extension of $\mathfrak{Z}_n$ by a bottom element.



(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

$$f \mathrel{\wedge\!\!\wedge}_F g \quad :\text{iff} \quad (\mathfrak{a} \downarrow F) \text{ contains an alternating path from } f \text{ to } g.$$

If $f \mathrel{\wedge\!\!\wedge}_F g$, we call f and g *alternating-path equivalent*, or *a.p.-equivalent* for short. We denote the a.p.-equivalence class of f by $[f]_F^{\wedge\!\!\wedge}$.

(b) For families $f = (f_i)_{i \in I}$ and $g = (g_i)_{i \in I}$ of morphisms, we set

$$f \mathrel{\wedge\!\!\wedge}_F g \quad :\text{iff} \quad f_i \mathrel{\wedge\!\!\wedge}_F g_i \quad \text{for all } i \in I.$$

Again, we denote the a.p.-equivalence class of f by $[f]_F^{\wedge\!\!\wedge}$.

The following lemma collects the basic properties of the relation $\mathrel{\wedge\!\!\wedge}_F$.

Lemma 5.3. *Let $F : \mathcal{I} \to \mathcal{C}$ be a functor and $f, g \in (\mathfrak{a} \downarrow F)$.*

(a) *$\mathrel{\wedge\!\!\wedge}_F$ is an equivalence relation.*

(b) *For every morphism $h : \mathfrak{b} \to \mathfrak{a}$,*

$$f \mathrel{\wedge\!\!\wedge}_F g \quad \textit{implies} \quad f \circ h \mathrel{\wedge\!\!\wedge}_F g \circ h.$$

(c) *For all functors $D : \mathcal{C} \to \mathcal{D}$,*

$$f \mathrel{\wedge\!\!\wedge}_F g \quad \textit{implies} \quad D(f) \mathrel{\wedge\!\!\wedge}_{D \circ F} D(g).$$

(d) *For all functors $G : \mathcal{J} \to \mathcal{I}$ and morphisms $h, h' \in \mathcal{I}^{\mathrm{mor}}$,*

$$F(h) \circ f \mathrel{\wedge\!\!\wedge}_{F \circ G} F(h') \circ g \quad \textit{implies} \quad f \mathrel{\wedge\!\!\wedge}_F g.$$

Proof. (a) $\mathrel{\wedge\!\!\wedge}_F$ 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 $f \mathrel{\wedge\!\!\wedge}_F g$ and $g \mathrel{\wedge\!\!\wedge}_F h$. 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

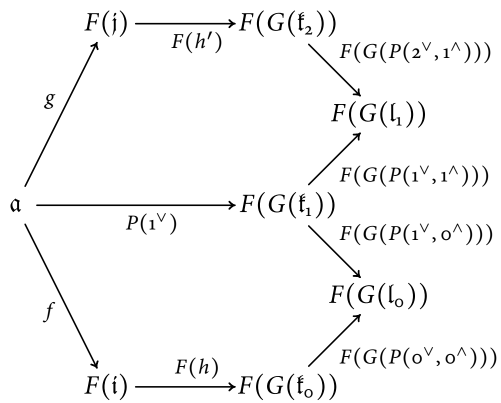
$$Q(x) := P(x) \circ h \quad \text{and} \quad Q(x, y) := P(x, y), \quad \text{for } x, y \in Z_n.$$

(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

$$\mu_{\mathfrak{i}} \circ f = \mu_{\mathfrak{k}} \circ P(1^\vee) = \mu_{\mathfrak{j}} \circ g\,,$$

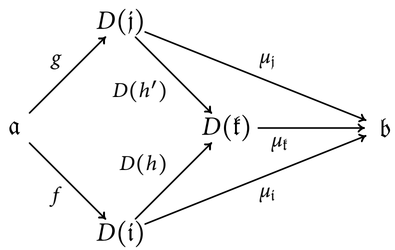
where $\mathfrak{k} \in I$ is the index such that $P(1^\vee) : \mathfrak{a} \to D(\mathfrak{k})$.

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.

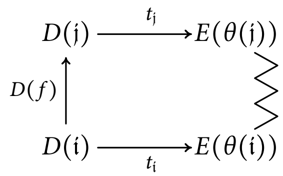
Definition 5.5. Let $D : \mathcal{I} \to \mathcal{C}$ and $E : \mathcal{J} \to \mathcal{C}$ be diagrams. A *link* from D to E is a family $t = (t_{\mathfrak{i}})_{\mathfrak{i} \in \mathcal{I}^{\mathrm{obj}}}$ of morphisms

$$t_{\mathfrak{i}} : D(\mathfrak{i}) \to E(\theta(\mathfrak{i}))\,, \quad \text{for some function } \theta : \mathcal{I}^{\mathrm{obj}} \to \mathcal{J}^{\mathrm{obj}},$$

satisfying

$$t_{\mathfrak{i}} \wedge\!\!\wedge_E t_{\mathfrak{j}} \circ D(f)\,,$$

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 $s \wedge\!\!\wedge_E s'$ and $t \wedge\!\!\wedge_F t'$, then $s * t \wedge\!\!\wedge_F s' * t'$.*

(c) *For morphisms $f : \mathfrak{a} \to D(\mathfrak{i})$ and $g : \mathfrak{a} \to D(\mathfrak{j})$,*

$$f \wedge\!\!\wedge_D g \quad \text{implies} \quad t_{\mathfrak{i}} \circ f \wedge\!\!\wedge_E t_{\mathfrak{j}} \circ g .$$

(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*

$$\mathrm{in}_E * t = t \quad \text{and} \quad s * \mathrm{in}_E = s .$$

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

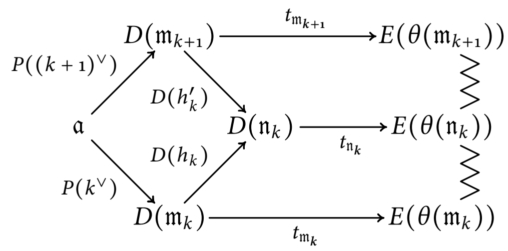
$$P(k^\vee, k^\wedge) = h_k : \mathfrak{m}_k \to \mathfrak{n}_k$$
$$\text{and} \quad P((k+1)^\vee, k^\wedge) = h'_k : \mathfrak{m}_{k+1} \to \mathfrak{n}_k .$$

As t is a link, we have

$$t_{\mathfrak{m}_k} \wedge\!\!\wedge_E t_{\mathfrak{n}_k} \circ D(h_k) \quad \text{and} \quad t_{\mathfrak{m}_{k+1}} \wedge\!\!\wedge_E t_{\mathfrak{n}_k} \circ D(h'_k) ,$$

which implies that

$$t_{\mathfrak{m}_k} \circ P(k^\vee) \wedge\!\!\wedge_E t_{\mathfrak{n}_k} \circ D(h_k) \circ P(k^\vee) = t_{\mathfrak{n}_k} \circ D(h'_k) \circ P((k+1)^\vee) \wedge\!\!\wedge_E t_{\mathfrak{m}_{k+1}} \circ P((k+1)^\vee) .$$



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 $s_\mathfrak{i} \mathrel{\wedge\!\!\wedge}_E t_\mathfrak{i}$, 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,

$$\begin{aligned} & \pi_{\mathrm{in}_E}(\mu) = \mu * \mathrm{in}_E = \mu , \\ \text{and} \quad & \pi_{t * s}(\mu) = \mu * t * s = \pi_s(\pi_t(\mu)) . \end{aligned}$$ □

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) *$\pi_{D,F} = \pi_{D[t]} \circ \pi_{D,G}$.*

$$\begin{array}{ccc} & \mathrm{Cone}(D, -) & \\ {}^{\pi_{D,F}} \swarrow & & \searrow {}^{\pi_{D,G}} \\ \mathrm{Cone}(D \circ F, -) & \xleftarrow[\pi_{D[t]}]{} & \mathrm{Cone}(D \circ G, -) \end{array}$$

(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

$$t_\mathfrak{j} \circ F(g) \mathrel{\wedge\!\!\wedge}_G t_\mathfrak{i} ,$$

which, by Lemma 5.3 (c), implies that

$$D(t_\mathfrak{j}) \circ D(F(g)) \mathrel{\wedge\!\!\wedge}_{D \circ G} D(t_\mathfrak{i}) .$$

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 $[t]^{\Lsh}_E$ 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 $[s]^{\Lsh}_E : D \to E$ be an isomorphism with inverse $[t]^{\Lsh}_D : E \to D$. By Lemma 5.8,

$$\begin{aligned}&t * s \approx_D \mathrm{in}_D \quad \text{implies} \quad \pi_s \circ \pi_t = \pi_{t*s} = \pi_{\mathrm{in}_D} = \mathrm{id}\,,\\ \text{and} \quad &s * t \approx_E \mathrm{in}_E \quad \text{implies} \quad \pi_t \circ \pi_s = \pi_{s*t} = \pi_{\mathrm{in}_E} = \mathrm{id}\,.\end{aligned}$$

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 5.1. 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

$$\big\{\, [f]^{\Lsh}_E \bigm| f : \mathfrak{a} \to E(\mathfrak{j}) \text{ for some } \mathfrak{j} \in \mathcal{J} \,\big\}\,.$$

(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 $[\mathrm{in}_F]^{\Lsh}_{\mathrm{id}_{\mathcal{C}}} : F \to \mathrm{id}_{\mathcal{C}}$ of $\mathrm{Ind}_{\mathrm{all}}(\mathcal{C})$ has a left inverse.*

Proof. (⇒) 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 $[t]_F^{\curlywedge} \circ [\mathrm{in}_F]_{\mathrm{id}_\mathcal{C}}^{\curlywedge} = \mathrm{id}$.

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, $t_{F(\mathfrak{i})} \circ \mathrm{id}_{F(\mathfrak{i})} \curlywedge_F \mathrm{id}_{F(\mathfrak{i})}$.

(⇐) Let $[t]_F^{\curlywedge}$ be a left inverse of $[\mathrm{in}_F]_{\mathrm{id}_\mathcal{C}}^{\curlywedge}$. 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 $g : \mathfrak{a} \to F(\mathfrak{j})$. Since $[t]_F^{\curlywedge} \circ [\mathrm{in}_F]_{\mathrm{id}_\mathcal{C}}^{\curlywedge} = \mathrm{id}$, we have

$$t_{F(\mathfrak{i})} = t_{F(\mathfrak{i})} \circ \mathrm{id}_{F(\mathfrak{i})} \curlywedge_F \mathrm{id}_{F(\mathfrak{i})}\,,$$
$$t_{F(\mathfrak{j})} = t_{F(\mathfrak{j})} \circ \mathrm{id}_{F(\mathfrak{j})} \curlywedge_F \mathrm{id}_{F(\mathfrak{j})}\,,$$

which implies that

$$t_{F(\mathfrak{i})} \circ f \curlywedge_F \mathrm{id}_{F(\mathfrak{i})} \circ f = f\,,$$
$$t_{F(\mathfrak{j})} \circ g \curlywedge_F \mathrm{id}_{F(\mathfrak{j})} \circ g = g\,.$$

As t is a link from $\mathrm{id}_\mathcal{C}$ to F, it follows that

$$f \curlywedge_F t_{F(\mathfrak{i})} \circ f \curlywedge_F t_\mathfrak{a} \curlywedge_F t_{F(\mathfrak{j})} \circ g \curlywedge_F g\,. \qquad \square$$

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, $[\mathrm{in}_F]_{\mathrm{id}_\mathcal{I}}^{\curlywedge} : F \to \mathrm{id}_\mathcal{I}$ has a left inverse $[t]_F^{\curlywedge} : \mathrm{id}_\mathcal{I} \to F$. 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

$$t_{F(\mathfrak{i})} = t_{F(\mathfrak{i})} \circ \mathrm{id}_{F(\mathfrak{i})} \curlywedge_F \mathrm{id}_{F(\mathfrak{i})}\,,$$

which implies, by Lemma 5.3 (c), that

$$D(t_{F(\mathfrak{i})}) \curlywedge_{D \circ F} D(\mathrm{id}_{F(\mathfrak{i})})\,.$$

For $\mu \in \mathrm{Cone}(D \circ F, \mathfrak{a})$, it therefore follows by Lemma 5.4 that

$$\begin{aligned}
\pi_{D,F}(\pi_{D[t]}(\mu)) &= \pi_{D,F}\big((\mu_{\theta(\mathfrak{i})} \circ D(t_\mathfrak{i}))_{\mathfrak{i} \in \mathcal{I}}\big) \\
&= \big(\mu_{\theta(F(\mathfrak{i}))} \circ D(t_{F(\mathfrak{i})})\big)_{\mathfrak{i} \in \mathcal{S}} \\
&= \big(\mu_\mathfrak{i} \circ D(\mathrm{id}_{F(\mathfrak{i})})\big)_{\mathfrak{i} \in \mathcal{S}} \\
&= \mu\,.
\end{aligned} \qquad \square$$

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.

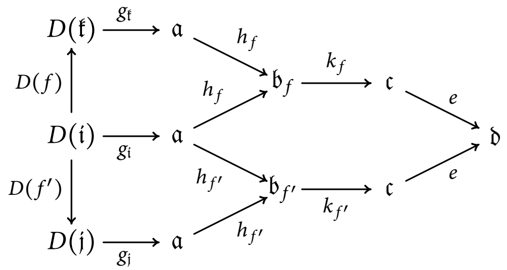
Lemma 1.2. *A category $\mathcal{C}$ is κ-filtered if, and only if, there is a cocone for every diagram $D : \mathcal{I} \to \mathcal{C}$ of size less than κ.*

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*

$$f \approx_D g \quad \text{iff} \quad \text{there are } h : \mathfrak{i} \to \mathfrak{k} \text{ and } h' : \mathfrak{j} \to \mathfrak{k} \text{ in } \mathcal{I} \text{ such that } D(h) \circ f = D(h') \circ g\,.$$

Proof. ($\Leftarrow$) 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.

($\Rightarrow$) 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

$$\begin{aligned} Q(g) &:= \mathfrak{k}\,, && \text{for objects } g : \mathfrak{a} \to D(\mathfrak{k})\,, \\ Q(h) &:= h\,, && \text{for morphisms } h : g \to g'\,. \end{aligned}$$

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, $f \approx_{D \circ Q \circ P} g$ and, applying Lemma B3.5.4 to the diagram $D \circ Q \circ P$, we obtain

$$D(\mu_0) \circ f = D(\mu_n) \circ g\,.$$ □

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'. $\square$

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 $D : \mathfrak{I} \to \mathcal{C}$.*

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 $D : \mathfrak{I} \to \mathcal{J}$. Then the composition $P \circ D$ is the desired dense κ-directed diagram.

As index set we use the partial order $\mathfrak{I} := \langle \mathcal{I}, \subseteq \rangle$ 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 $\mathfrak{I}$ 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

$$\bigcup_{\mathcal{B} \in X} \mathcal{B}^{\mathrm{mor}} .$$

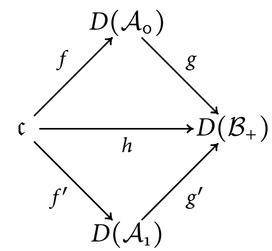
Since κ is regular, $\mathcal{A}$ still has less than κ morphisms. By Lemma 1.6, there exists a subcategory $\mathcal{A} \subseteq \mathcal{A}_+ \subseteq \mathcal{J}$ with a unique terminal object. Hence, $\mathcal{A}_+ \in \mathcal{I}$ is an upper bound of X.

Let $D : \mathfrak{I} \to \mathcal{J}$ be the functor mapping a subcategory $\mathcal{A} \in \mathfrak{I}$ 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, $\mathcal{A} \in \mathfrak{I}$ 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 $\mathcal{P}^\lambda_\kappa$ 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 $\mathrm{Ind}^\lambda_\kappa(\mathcal{C})$ 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 $D : \mathcal{I} \to \mathrm{Ind}^\lambda_\kappa(\mathcal{C})$. 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 $D : \mathcal{I} \to \mathrm{Ind}^\lambda_\kappa(\mathcal{C})$ 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 $D(f) = [t(f)]^{\wedge\!\!\wedge}_{D(\mathfrak{j})}$. 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

$$U(\langle \mathfrak{i},\mathfrak{k}\rangle) := D(\mathfrak{i})(\mathfrak{k})\,, \quad \text{for } \langle \mathfrak{i},\mathfrak{k}\rangle \in \mathcal{J}^{\mathrm{obj}},$$
$$U(f) := f\,, \quad \text{for } f \in \mathcal{J}^{\mathrm{mor}}.$$

(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 $t^\mathfrak{i} = (t^\mathfrak{i}_\mathfrak{k})_{\mathfrak{k}\in\mathcal{K}(\mathfrak{i})}$ be a link such that $\mu_\mathfrak{i} = [t^\mathfrak{i}]^{\wedge\!\!\wedge}_E$. The *union* of μ is the a.p.-equivalence class $[t]^{\wedge\!\!\wedge}_E$ of the family

$$t := (t^\mathfrak{i}_\mathfrak{k})_{\langle \mathfrak{i},\mathfrak{k}\rangle\in\mathcal{J}}\,.$$

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 $[u]^{\wedge\!\!\wedge}_E$ of $\mu \in \mathrm{Cone}(D, E)$ is a morphism $[u]^{\wedge\!\!\wedge}_E : U \to E$ 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

$$[u^\mathfrak{i}]^{\wedge\!\!\wedge}_E = \mu_\mathfrak{i} = [w^\mathfrak{i}]^{\wedge\!\!\wedge}_E\,.$$

Then $[u^\mathfrak{i}_\mathfrak{k}]^{\wedge\!\!\wedge}_E = [w^\mathfrak{i}_\mathfrak{k}]^{\wedge\!\!\wedge}_E$, for all $\langle \mathfrak{i},\mathfrak{k}\rangle \in \mathcal{J}$, which implies that the corresponding links $u = (u^\mathfrak{i}_\mathfrak{k})_{\langle \mathfrak{i},\mathfrak{k}\rangle\in\mathcal{J}}$ and $w = (w^\mathfrak{i}_\mathfrak{k})_{\langle \mathfrak{i},\mathfrak{k}\rangle\in\mathcal{J}}$ are a.p.-equivalent and induce the same value $[u]^{\wedge\!\!\wedge}_E = [w]^{\wedge\!\!\wedge}_E$.

(b) Let $\mu \in \mathrm{Cone}(D, E)$ be a cocone where $\mu_\mathfrak{i} = [u^\mathfrak{i}]^{\wedge\!\!\wedge}_E$, and let $[u]^{\wedge\!\!\wedge}_E$ 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

$$u^{\mathfrak{j}}_{\mathfrak{l}} \circ U(h) \curlywedge_E u^{\mathfrak{i}}_{\mathfrak{k}}$$

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,

$$u^{\mathfrak{i}}_{\mathfrak{l}} \circ D(\mathfrak{i})(g) \curlywedge_E u^{\mathfrak{i}}_{\mathfrak{k}} .$$

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 $[u^{\mathfrak{j}}]^{\curlywedge}_E \circ [t(f)]^{\curlywedge}_{D(\mathfrak{j})} = [u^{\mathfrak{i}}]^{\curlywedge}_E$. Hence,

$$u^{\mathfrak{j}}_{\theta(\mathfrak{k})} \circ t(f)_{\mathfrak{k}} \curlywedge_E u^{\mathfrak{i}}_{\mathfrak{k}} ,$$

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

$$\mu_{\mathfrak{i}} = [u^{\mathfrak{i}}]^{\curlywedge}_E , \quad \mu'_{\mathfrak{i}} = [w^{\mathfrak{i}}]^{\curlywedge}_E , \quad \text{and} \quad \eta_E(\mu) = [t]^{\curlywedge}_E .$$

Then $[u^{\mathfrak{i}}_{\mathfrak{k}}]^{\curlywedge}_E = [t_{\langle \mathfrak{i},\mathfrak{k} \rangle}]^{\curlywedge}_E = [w^{\mathfrak{i}}_{\mathfrak{k}}]^{\curlywedge}_E$ for all indices $\mathfrak{i}, \mathfrak{k}$. Consequently,

$$\mu_{\mathfrak{i}} = [u^{\mathfrak{i}}]^{\curlywedge}_E = [w^{\mathfrak{i}}]^{\curlywedge}_E = \mu'_{\mathfrak{i}} , \quad \text{for all } \mathfrak{i} \in \mathcal{I} ,$$

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 $\mu := ([s^{\mathfrak{i}}]^{\curlywedge}_E)_{\mathfrak{i} \in \mathcal{I}}$. As $\eta_E(\mu) = [s]^{\curlywedge}_E$ 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 $s_{\mathfrak{j},\mathfrak{l}} \circ D(\mathfrak{i})(g) \curlywedge_E s_{\mathfrak{i},\mathfrak{k}}$, 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,

$$s_{\mathfrak{j},\theta(\mathfrak{k})} \circ U(t(f)_{\mathfrak{k}}) \curlywedge_E s_{\mathfrak{i},\mathfrak{k}} , \quad \text{for every } \mathfrak{k} \in \mathcal{K}(\mathfrak{i}) .$$

Consequently,

$$\mu_{\mathfrak{j}} \circ D(f) = [s^{\mathfrak{j}}]^{\curlywedge}_E \circ [t(f)]^{\curlywedge}_{D(\mathfrak{j})} = [s^{\mathfrak{i}}]^{\curlywedge}_E = \mu_{\mathfrak{i}} .$$

(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 $D : \mathcal{I} \to \mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$ be a κ-filtered diagram of size less than λ with union U.*

(a) *$U \in \mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$.*

(b) *$U = \varinjlim D$ and a limiting cocone $\mu = (\mu_{\mathfrak{i}})_{\mathfrak{i} \in \mathcal{I}}$ from D to U is given by*

$$\mu_{\mathfrak{i}} = [\mathrm{in}_{D(\mathfrak{i})}]^{\curlywedge}_U : D(\mathfrak{i}) \to U .$$

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 $U_0 := U \restriction \mathcal{J}_0$ is a union of some restriction $D \restriction \mathcal{I}_0$, 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

$$|\mathcal{I}_0^{\mathrm{mor}}| < \kappa \quad \text{and} \quad |\mathcal{J}_0^{\mathrm{mor}}| < \kappa\,.$$

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 $[u]^{\Lsh}_{D(\mathfrak{m})}$ 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

$$\begin{aligned} \mu_\mathfrak{j} \circ D(f) &= [\mathrm{in}_{D(\mathfrak{j})}]^{\Lsh}_U \circ [t(f)]^{\Lsh}_{D(\mathfrak{j})} \\ &= [\mathrm{in}_{D(\mathfrak{j})} * t(f)]^{\Lsh}_U = [\mathrm{in}_{D(\mathfrak{i})}]^{\Lsh}_U = \mu_\mathfrak{i}\,. \end{aligned}$$

It remains to show that μ is limiting. Let $\mu' \in \mathrm{Cone}(D, E)$ be a cocone where $\mu'_\mathfrak{i} = [w^\mathfrak{i}]^{\Lsh}_E$, and let $[w]^{\Lsh}_E$ be the union of μ'. We have seen in Lemma 1.12 (b) that $[w]^{\Lsh}_E$ is a morphism $U \to E$. Furthermore,

$$[w]^{\Lsh}_E * \mu = \big([w^\mathfrak{i}]^{\Lsh}_E \circ [\mathrm{in}_{D(\mathfrak{i})}]^{\Lsh}_U\big)_{\mathfrak{i} \in \mathcal{I}} = \big([w^\mathfrak{i}]^{\Lsh}_E\big)_{\mathfrak{i} \in \mathcal{I}} = (\mu'_\mathfrak{i})_{\mathfrak{i} \in \mathcal{I}} = \mu'\,.$$

Hence, the function $[w]^{\Lsh}_E \mapsto [w]^{\Lsh}_E * \mu$ 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 $I(\mathfrak{a})$,*

$$[s]^{\Lsh}_{I(\mathfrak{a})} = [t]^{\Lsh}_{I(\mathfrak{a})} : D \to I(\mathfrak{a}) \quad \text{implies} \quad s = t\,.$$

(c) *I is full and faithful.*

(d) *For every $D \in \mathrm{Ind}_{\mathcal{P}}(\mathcal{C})$, the inclusion $[\mathrm{in}_D]^{\Lsh}_U : D \to U$ 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 $f \circ I(\mathrm{id}_{\mathfrak{a}}) \Lsh_{I(\mathfrak{b})} f$.

(b) Let $\mathfrak{i} \in \mathcal{I}$. Since $[s]^{\Lsh}_{I(\mathfrak{a})} = [t]^{\Lsh}_{I(\mathfrak{a})}$, 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 $[f]^{\Lsh}_{I(\mathfrak{b})} : I(\mathfrak{a}) \to I(\mathfrak{b})$ 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 $I(f_0) = [f]^{\Lsh}_{I(\mathfrak{b})}$.

To prove that I is faithful, suppose that $I(f) = I(g)$ for morphisms $f, g : \mathfrak{a} \to \mathfrak{b}$. Then $[f]^{\Lsh}_{I(\mathfrak{b})} = [g]^{\Lsh}_{I(\mathfrak{b})}$ 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 $[\mathrm{in}_D]^{\Lsh}_U : D \to U$ only consists of identity morphisms $\mathrm{id}_{D(\mathfrak{i})} : D(\mathfrak{i}) \to U(\langle \mathfrak{i}, 0 \rangle)$, it has an inverse $[t]^{\Lsh}_D : U \to D$ 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

$$\tau_{I,D} : \mathrm{Cone}(D, \mathfrak{a}) \to \mathrm{Cone}(I \circ D, I(\mathfrak{a})) : \mu \mapsto I[\mu]$$
$$\text{and} \quad \eta_{I(\mathfrak{a})} : \mathrm{Cone}(I \circ D, I(\mathfrak{a})) \to \mathrm{Ind}_{\mathrm{all}}(\mathcal{C})(D, I(\mathfrak{a}))\,,$$

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 $t_{\mathfrak{j}} \circ D(f) \Lsh_{I(\mathfrak{a})} t_{\mathfrak{i}}$.

($\Rightarrow$) 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 $I[\mu] = [t]^{\Lsh}_{I(\mathfrak{a})}$. 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* $\mathrm{Ind}^{\lambda}_{\kappa}(\mathfrak{I})$.

Proof. We have to prove that between any two objects $D : \mathcal{J} \to \mathfrak{I}$ and $E : \mathcal{K} \to \mathfrak{I}$ of $\mathrm{Ind}^{\lambda}_{\kappa}(\mathfrak{I})$, there is at most one morphism. Consider two links s and t from D to E. We claim that $s \wedge\!\!\wedge_{E}\, t$. 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 $s_{\mathfrak{j}} \wedge\!\!\wedge_{E}\, t_{\mathfrak{j}}$. □

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* $\mathrm{Ind}^{\lambda}_{\kappa}(\mathfrak{I})$.

Proof. It is sufficient to prove that the function

$$h : \mathrm{Ind}^{\lambda}_{\kappa}(\mathfrak{I}) \to \mathfrak{I}^{+} : D \mapsto \Downarrow \mathrm{rng}\, D^{\mathrm{obj}}$$

is a surjective strict homomorphism. Then h induces a full and faithful functor $\mathrm{Ind}^{\lambda}_{\kappa}(\mathfrak{I}) \to \mathfrak{I}^{+}$. 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 $[t]^{\approx}_{E} : D \to E$. 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 $S = \Downarrow S_{\mathrm{o}}$, for a κ-directed set $S_{\mathrm{o}} \subseteq I$ of size $|S_{\mathrm{o}}| < \lambda$. Let $D : \mathfrak{J} \restriction S_{\mathrm{o}} \to \mathfrak{J}$ be the inclusion functor. Then $D \in \mathrm{Ind}^{\lambda}_{\kappa}(\mathfrak{J})$ and $h(D) = \Downarrow S_{\mathrm{o}} = S$. ☐

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 $\mathrm{Cn}_{\aleph_{\mathrm{o}}}$ be the class of all infinite cardinals.*

(a) *$\trianglelefteq$ is a partial order on $\mathrm{Cn}_{\aleph_{\mathrm{o}}}$.*

(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 $\langle \mathrm{Cn}_{\aleph_{\mathrm{o}}}, \trianglelefteq\rangle$ 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 $z_n \in \aleph_{n+1} \setminus \bigcup f[\downarrow\aleph_n]$. 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

$$z_n \in \aleph_{n+1} \setminus \bigcup f[\downarrow\aleph_n] \supseteq \aleph_{n+1} \setminus Y$$

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 $\mathfrak{I}$ be a κ-directed partial order and let $\mathfrak{I}^+$ be its (κ, λ)-completion. We have seen in Proposition 2.3 that the categories $\mathrm{Ind}^\lambda_\kappa(\mathfrak{I})$ and $\mathfrak{I}^+$ 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 $\mathfrak{I}^+$ of a given κ-directed partial order $\mathfrak{I}$, we can use it to extend κ-directed diagrams $D : \mathfrak{I} \to \mathcal{C}$ to a λ-directed diagram $D^+ : \mathfrak{I}^+ \to \mathcal{C}$. This construction is defined via a detour through the inductive completion $\mathrm{Ind}^\lambda_\kappa(\mathcal{C})$. We construct two diagrams $\mathfrak{I}^+ \to \mathrm{Ind}^\lambda_\kappa(\mathcal{C})$ and $\mathrm{Ind}^\lambda_\kappa(\mathcal{C}) \to \mathcal{C}$ whose composition is the extension $\mathfrak{I}^+ \to \mathcal{C}$ 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

$$\begin{aligned} D^+(S) &:= D \restriction S\,, && \text{for objects } S \in F\,, \\ D^+(S,T) &:= [\mathrm{in}_{D\restriction S}]^{\ggg}_{D\restriction T}\,, && \text{for pairs } S \subseteq T\,. \end{aligned}$$

(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 $D^+(S,T) = [u^{S,T}]^{\ggg}_{D\restriction T}$. 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 $[s]^{\ggg}_D : U \to D$ is an inverse of $[t]^{\ggg}_U : D \to U$.

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

$$\begin{aligned} u^{\theta(\mathfrak{j}),S}_{\mathfrak{j}} \circ t_{\mathfrak{j}} \circ D(f) &= \mathrm{id}_{D(\mathfrak{j})} \circ \mathrm{id}_{D(\mathfrak{j})} \circ D(f) \\ &= D(f) \circ \mathrm{id}_{D(\mathfrak{i})} \circ \mathrm{id}_{D(\mathfrak{i})} \\ &= U(D(f)) \circ u^{\theta(\mathfrak{i}),S}_{\mathfrak{i}} \circ t_{\mathfrak{i}}\,. \end{aligned}$$

Hence, $u_j^{\theta(j),S}$ and $D(f) \circ u_i^{\theta(i),S}$ form an alternating path from $t_j \circ D(f)$ to t_i in $(D(i) \downarrow U)$.

For s, note that $\mathcal{J}$ is generated by morphisms of the form $D(f)$ and $u_i^{S,T}$, for $f \in \mathcal{I}^{\text{mor}}$, $S \subseteq T$, and $i \in \mathcal{I}^{\text{obj}}$. Hence, it is sufficient to check that

$$s_{\langle T,j\rangle} \circ U(h) \mathrel{\wedge\!\!\!\wedge}_D s_{\langle S,i\rangle} \quad \text{for such morphisms } h\,.$$

For $h = u_i^{S,T}$, we have

$$s_{\langle T,i\rangle} \circ U(u_i^{S,T}) = \mathrm{id}_{D(i)} \circ \mathrm{id}_{D(i)} = \mathrm{id}_{D(i)} = s_{\langle S,i\rangle}\,.$$

For $h = D(f)$ with $f : i \to j$ in $\mathcal{I}$,

$$\begin{aligned} D(\mathrm{id}_j) \circ s_{\langle S,j\rangle} \circ U(D(f)) &= D(\mathrm{id}_j) \circ \mathrm{id}_{D(j)} \circ D(f) \\ &= D(f) \circ \mathrm{id}_{D(i)} \\ &= D(f) \circ s_{\langle S,i\rangle} \end{aligned}$$

implies that $s_{\langle S,j\rangle} \circ U(D(f)) \mathrel{\wedge\!\!\!\wedge}_D s_{\langle S,i\rangle}$.

It remains to prove that $[s]_D^{\wedge\!\!\wedge}$ is an inverse of $[t]_U^{\wedge\!\!\wedge}$. Since

$$s * t = (s_{\langle\theta(i),i\rangle} \circ t_i)_{i\in\mathcal{I}} = (\mathrm{id}_{D(i)})_{i\in\mathcal{I}}\,,$$

s is a left inverse of t. To show that it is also a right inverse, let $\langle S, i\rangle \in \mathcal{J}$ and fix a set $T \in F$ with $\theta(i) \cup S \subseteq T$. Then

$$\begin{aligned} U(u_i^{\theta(i),T}) \circ (t * s)_{\langle S,i\rangle} &= \mathrm{id}_{D(i)} \circ t_i \circ s_{\langle S,i\rangle} \\ &= \mathrm{id}_{D(i)} \circ \mathrm{id}_{D(i)} \circ \mathrm{id}_{D(i)} \\ &= U(\mathrm{id}_{D(i)}) \circ \mathrm{id}_{U(\langle S,i\rangle)} \end{aligned}$$

implies that $(t * s)_{\langle S,i\rangle} \mathrel{\wedge\!\!\!\wedge}_U \mathrm{id}_{U(\langle S,i\rangle)}$. □

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 $[t]_E^{\wedge\!\!\wedge} : D \to E$, we choose for $Q^{\text{mor}}([t]_E^{\wedge\!\!\wedge})$ 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 $s \mathrel{\wedge\!\!\!\wedge}_E t$. Then

$$\lambda^E * s \mathrel{\wedge\!\!\!\wedge}_{I(\mathfrak{a})} \lambda^E * t$$

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 $\mu \in \mathrm{Cone}(Q \circ D, \mathfrak{b})$ 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

$$\nu_{\mathfrak{i}} := [\mu_{\mathfrak{i}} * \lambda^{D(\mathfrak{i})}]^{\ggg}_{I(\mathfrak{b})} .$$

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 $D(f) = [t]^{\ggg}_{D(\mathfrak{j})}$. 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

$$\begin{aligned} \nu_{\mathfrak{j}} \circ D(f) &= \big[\mu_{\mathfrak{j}} * \lambda^{D(\mathfrak{j})}\big]^{\ggg}_{I(\mathfrak{b})} \circ D(f) \\ &= \big[\mu_{\mathfrak{j}} * \lambda^{D(\mathfrak{j})} * t\big]^{\ggg}_{I(\mathfrak{b})} \\ &= \big[\mu_{\mathfrak{j}} * Q(D(f)) * \lambda^{D(\mathfrak{i})}\big]^{\ggg}_{I(\mathfrak{b})} \\ &= \big[\mu_{\mathfrak{i}} * \lambda^{D(\mathfrak{i})}\big]^{\ggg}_{I(\mathfrak{b})} = \nu_{\mathfrak{i}} , \end{aligned}$$

as desired.

As ν is a cocone of D and λ^* is limiting, there exists a unique morphism $[t]^{\ggg}_{I(\mathfrak{b})} : E \to I(\mathfrak{b})$ such that

$$\nu = [t]^{\ggg}_{I(\mathfrak{b})} * \lambda^* .$$

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 $\lambda^*_{\mathfrak{i}} = [s^{\mathfrak{i}}]^{\ggg}_E$. Then

$$Q(\lambda^*_{\mathfrak{i}}) * \lambda^{D(\mathfrak{i})} = \lambda^E * s^{\mathfrak{i}}$$

implies that

$$\big[Q(\lambda^*_{\mathfrak{i}}) * \lambda^{D(\mathfrak{i})}\big]^{\ggg}_{I(\mathfrak{a}_E)} = \big[\lambda^E\big]^{\ggg}_{I(\mathfrak{a}_E)} * \lambda^*_{\mathfrak{i}} .$$

For every $i \in \mathcal{I}$, it follows that

$$\begin{aligned} \big[\varphi * Q(\lambda^*_{\mathfrak{i}}) * \lambda^{D(\mathfrak{i})}\big]^{\ggg}_{I(\mathfrak{b})} &= \big[\varphi * \lambda^E\big]^{\ggg}_{I(\mathfrak{b})} * \lambda^*_{\mathfrak{i}} \\ &= [t]^{\ggg}_{I(\mathfrak{b})} * \lambda^*_{\mathfrak{i}} = \nu_{\mathfrak{i}} = [\mu_{\mathfrak{i}} * \lambda^{D(\mathfrak{i})}]^{\ggg}_{I(\mathfrak{b})} . \end{aligned}$$

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

$$\begin{aligned} \big[\psi * \lambda^E\big]^{\ggg}_{I(\mathfrak{b})} * \lambda^*_{\mathfrak{i}} &= \big[\psi * Q(\lambda^*_{\mathfrak{i}}) * \lambda^{D(\mathfrak{i})}\big]^{\ggg}_{I(\mathfrak{b})} \\ &= [\mu_{\mathfrak{i}} * \lambda^{D(\mathfrak{i})}]^{\ggg}_{I(\mathfrak{b})} = \nu_{\mathfrak{i}} = [t]^{\ggg}_{I(\mathfrak{b})} * \lambda^*_{\mathfrak{i}} , \end{aligned}$$

and it follows by Lemma B3.4.2 that

$$[\psi * \lambda^E]^{\ggg}_{I(\mathfrak{b})} = [t]^{\ggg}_{I(\mathfrak{b})} .$$

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 $D^+ : \mathfrak{J}^+ \to \mathrm{Ind}^{\lambda}_{\kappa}(\mathcal{C})$ 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, $Q \circ D^+ : \mathfrak{J}^+ \to \mathcal{C}$ 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 $t \circ t \mathrel{⩕}_D t$.

(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 $[t]^{⩕}_*$, and let $D^+ : \mathcal{I}^+ \to \mathcal{C}$ be the union of F where we choose t as representative of $[t]^{⩕}_D$. 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

$$\psi = \psi \circ (\mu_* \circ s) = \mu'_* \circ s = \varphi \circ \mu_* \circ s = \varphi\,.$$ ☐

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. ($\Rightarrow$) 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 $s \circ r = D(e) = p$.

($\Leftarrow$) 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 $[t]^{\wedge\!\!\wedge}_D : D \to D$ forms a cocone from $F : \circlearrowleft \to \mathrm{Ind}_{\mathrm{all}}(\mathcal{C})$ to D whose union is just $[t]^{\wedge\!\!\wedge}_D$. Therefore, Lemma 1.12 (b) implies that t is a link from D^+ to D.

(b) ($\Rightarrow$) 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

$$\mu * t = (\mu_{\theta(\mathrm{i})} \circ t_{\mathrm{i}})_{\mathrm{i} \in \mathcal{I}} = (\mu_{\mathrm{i}})_{\mathrm{i} \in \mathcal{I}}\,.$$

($\Leftarrow$) If $\mu * t = \mu$, then it follows by (a) and Lemma B3.5.8 that

$$\mu = \mu * t = \pi_t(\mu) \in \mathrm{Cone}(D^+, \mathfrak{a})\,.$$

☐

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) $\Rightarrow$ (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) $\Rightarrow$ (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,

$$\begin{aligned} \mu = \mu * t &= \varphi * \lambda * t \\ &= \varphi * (e \circ r) * \lambda \\ &= (\varphi \circ e) * (r * \lambda) \\ &= (\varphi \circ e) * \lambda^+ = \eta_\mathfrak{c}(\varphi \circ e) \in \mathrm{rng}\, \eta_\mathfrak{c} . \end{aligned}$$

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 $\mu \in \mathrm{Cone}(D, \mathfrak{b})$ 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.

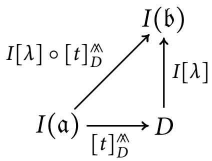
(b) For a singular cardinal κ, it follows by Lemma 1.4 that an object is κ-presentable if, and only if, it is κ^+-presentable.

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 $\lambda \in \mathrm{Cone}(D, \mathfrak{b})$, the function

$$\begin{aligned} &\mathrm{Ind}_{\mathrm{all}}(\mathcal{C})(I(\mathfrak{a}), I[\lambda]) \\ &\quad : \mathrm{Ind}_{\mathrm{all}}(\mathcal{C})(I(\mathfrak{a}), D) \to \mathrm{Ind}_{\mathrm{all}}(\mathcal{C})(I(\mathfrak{a}), I(\mathfrak{b})) \\ &\quad : [t]^{\approx}_D \mapsto I[\lambda] \circ [t]^{\approx}_D \end{aligned}$$

is bijective. (I denotes the inclusion functor $\mathcal{C} \to \mathrm{Ind}_{\mathrm{all}}(\mathcal{C})$.)



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 $f'_0 \mathrel{\wedge\!\!\wedge}_D f_0$. □

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) $\Rightarrow$ (3) is trivial.

(3) ⇒ (1) Let $D : \mathfrak{J} \to \mathcal{C}$ 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 $f_o \in \mathcal{C}(\mathfrak{a}, D(i))$ with

$$f = \mathcal{C}(\mathfrak{a}, \lambda_i)(f_o) = \lambda_i \circ f_o \,.$$

Hence, f factorises through λ. For essential uniqueness, suppose that there is a second index $j \in I$ and a morphism $f'_o : \mathfrak{a} \to D(j)$ such that $f = \lambda_j \circ f'_o$. Then

$$\mathcal{C}(\mathfrak{a}, \lambda_j)(f'_o) = \lambda_j \circ f'_o = \lambda_i \circ f_o = \mathcal{C}(\mathfrak{a}, \lambda_i)(f_o) \,.$$

Hence, $f_o \in \mathcal{C}(\mathfrak{a}, D(i))$ and $f'_o \in \mathcal{C}(\mathfrak{a}, D(j))$ 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

$$\mathcal{C}(\mathfrak{a}, D(i,k))(f_o) = \mathcal{C}(\mathfrak{a}, D(j,k))(f'_o) \,.$$

Consequently,

$$D(i,k) \circ f_o = D(j,k) \circ f'_o \,,$$

which implies that $f_o \approx_D f'_o$.

(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 $f = \mu_k \circ f_o$, for some $k \in K$ and $f_o : \mathfrak{a} \to D(F(k))$.

We claim that $\lambda_{F(k)} \circ f_o$ is an essentially unique factorisation of f through λ. Note that $\lambda_{F(k)} \circ f_o = \mu_k \circ f_o = f$ implies that it is a factorisation of f. Hence, it remains to prove essential uniqueness.

Suppose that $f = \lambda_{\mathfrak{i}} \circ f'_o$ 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,

$$\mu_k \circ f_o \quad \text{and} \quad \mu_l \circ D(g) \circ f'_o$$

are two factorisations of f through μ and, by essential uniqueness, we obtain

$$f_o \approx_{D \circ F} D(g) \circ f'_o \,.$$

By Lemma B3.5.3 (d), this implies that $f_o \approx_D f'_o$.

(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 $f = \lambda_{\mathfrak{j}} \circ f_0$ 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

$$D(h) \circ g(f) = D(h') \circ f_0\,.$$

For a cocone μ of $\mathcal{C}(\mathfrak{a}, -) \circ D$, it therefore follows that

$$\begin{aligned}\mu_{\mathfrak{i}(f)}(g(f)) &= \big(\mu_k \circ \mathcal{C}(\mathfrak{a}, D(h))\big)(g(f)) \\ &= \mu_k\big(D(h) \circ g(f)\big) \\ &= \mu_k(D(h') \circ f_0) \\ &= \big(\mu_k \circ \mathcal{C}(\mathfrak{a}, D(h'))\big)(f_0) = \mu_{\mathfrak{j}}(f_0)\,.\end{aligned}$$

Consequently,

$$\begin{aligned}\eta(\zeta(\mu)) = \zeta(\mu) * \lambda' &= \big(\zeta(\mu) \circ \mathcal{C}(\mathfrak{a}, \lambda_{\mathfrak{j}})\big)_{\mathfrak{j} \in \mathcal{I}} \\ &= \big(f_0 \mapsto \mu_{\mathfrak{i}(\lambda_{\mathfrak{j}} \circ f_0)}(g(\lambda_{\mathfrak{j}} \circ f_0))\big)_{\mathfrak{j} \in \mathcal{I}} \\ &= (f_0 \mapsto \mu_{\mathfrak{j}}(f_0))_{\mathfrak{j} \in \mathcal{I}} \\ &= (\mu_{\mathfrak{j}})_{\mathfrak{j} \in \mathcal{I}}\,. \qquad \square\end{aligned}$$

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,

$$\begin{aligned}D(g_i) \circ f_i &= D(g_{ij}) \circ D(h_{ij}) \circ f_i \\ &= D(g_{ij}) \circ D(h'_{ij}) \circ f_j = D(g_j) \circ f_j\,. \qquad \square\end{aligned}$$

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,*

$$s \curlywedge_E t \quad \text{iff} \quad \mu * s = \mu * t .$$

(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

$$s_{\mathfrak{i}} \curlywedge_E t_{\mathfrak{i}} \quad \text{iff} \quad \mu_{\rho(\mathfrak{i})} \circ s_{\mathfrak{i}} = \mu_{\theta(\mathfrak{i})} \circ t_{\mathfrak{i}} ,$$

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)

$$\mu * t' = f * \lambda = \mu * t \quad \text{implies} \quad t' \curlywedge_E t .$$

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 $t_{\mathfrak{i}} \curlywedge_D t_{\mathfrak{j}} \circ D(g)$. □

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 $D : \mathfrak{J} \to \mathcal{C}$ 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 $D^+ : \mathfrak{J}^+ \to \mathcal{C}$ 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 $t \circ t \mathrel{\approx}_D t$. 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 $D : \mathfrak{J} \to \mathcal{C}$ 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 $g \mathrel{\approx}_D h \circ r$. By Lemma B3.5.3 (b), it follows that

$$g \circ e \mathrel{\approx}_D h \circ r \circ e = h ,$$

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 $\mu \in \operatorname{Cone}(D, \mathfrak{B})$ 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 $|I| = |\operatorname{rng} g| \leq |A| < \lambda$.

(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 $\mathfrak{J} \in \mathfrak{Dir}(\kappa)$, let $\mathcal{S}$ be the family of all substructures of $\mathfrak{J}$ 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 $\mathfrak{J} = \varinjlim D$. Let $\mathcal{S}_0 \subseteq \mathcal{S}$ be the subfamily of all substructures of $\mathfrak{J}$ 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 $\mathfrak{J}$ is κ-directed, the set $A \subseteq I$ has an upper bound $b \in I$. Consequently, $\mathfrak{J}|_{A \cup \{b\}}$ 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 $\mathfrak{J}$ 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, $J(\mathfrak{J}) = \mathfrak{J} = \varinjlim (J \circ D_0)$ implies that $\mathfrak{J} = \varinjlim D_0$, 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 $[f]^{\Lsh}_U$ is essentially unique.

Let $[f]^{\Lsh}_U = \mu_\mathfrak{j} \circ [g]^{\Lsh}_{D(\mathfrak{j})}$ be a second factorisation where $[g]^{\Lsh}_{D(\mathfrak{j})} : I(\mathfrak{a}) \to D(\mathfrak{j})$. Then $g : \mathfrak{a} \to D(\mathfrak{j})(\mathfrak{l})$, for some index $\mathfrak{l}$, and, as above, it follows that

$$[f]^{\Lsh}_U = \mu_\mathfrak{j} \circ [g]^{\Lsh}_{D(\mathfrak{j})} = [\mathrm{id}_{D(\mathfrak{j})(\mathfrak{l})} \circ g]^{\Lsh}_U = [g]^{\Lsh}_U .$$

Hence, $f \Lsh_U g$ 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

$$h = h_{u-1} \circ \dots \circ h_0 \quad \text{and} \quad h' = h'_{v-1} \circ \dots \circ h'_0$$

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 $D(\mathfrak{x}, \mathfrak{y}) = [t(\mathfrak{x}, \mathfrak{y})]^{\Lsh}_{D(\mathfrak{y})}$. By induction on u and v it follows that

$$[h_{u-1} \circ \dots \circ h_0 \circ f]^{\Lsh}_{D(\mathfrak{m})} \Lsh_D [f]^{\Lsh}_{D(\mathfrak{i})}$$
$$\text{and} \quad [h'_{v-1} \circ \dots \circ h'_0 \circ g]^{\Lsh}_{D(\mathfrak{m})} \Lsh_D [g]^{\Lsh}_{D(\mathfrak{j})} .$$

Hence, $h \circ f = h' \circ g$ implies that

$$[f]^{\Lsh}_{D(\mathfrak{i})} \Lsh_D [h_{u-1} \circ \dots \circ h_0 \circ f]^{\Lsh}_{D(\mathfrak{m})}$$
$$= [h'_{v-1} \circ \dots \circ h'_0 \circ g]^{\Lsh}_{D(\mathfrak{m})} \Lsh_D [g]^{\Lsh}_{D(\mathfrak{j})} .$$ ☐

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 $\mathrm{Ind}_\kappa^\infty(\mathcal{C})$ is a κ-filtered colimit of κ-presentable objects.

Consequently, it remains to prove that, up to isomorphism, the κ-presentable objects of $\mathrm{Ind}_\kappa^\infty(\mathcal{C})$ 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 $\mathrm{Ind}_\kappa^\infty(\mathcal{C}_0)$, for some small category $\mathcal{C}_0$.*

Proof. ($\Leftarrow$) We have seen in Proposition 4.7 that $\mathrm{Ind}_\kappa^\infty(\mathcal{C}_0)$ is κ-accessible. Hence, all categories $\mathcal{C}$ equivalent to $\mathrm{Ind}_\kappa^\infty(\mathcal{C}_0)$ 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 $\mathcal{C}_0$ be a skeleton of $\mathcal{C}_1$. We claim that $\mathcal{C}$ is equivalent to $\mathrm{Ind}_\kappa^\infty(\mathcal{C}_0)$.

Let $Q_0 : \mathrm{Ind}_\kappa^\infty(\mathcal{C}_0) \to \mathcal{C}$ be the restriction of the canonical projection $Q : \mathrm{Ind}_\kappa^\infty(\mathcal{C}) \to \mathcal{C}$ to $\mathrm{Ind}_\kappa^\infty(\mathcal{C}_0)$. We claim that Q_0 is the desired equivalence. By Theorem B1.3.14, it is sufficient to prove that Q_0 is full and faithful and that every object of $\mathcal{C}$ is isomorphic to some object in rng $\mathrm{Q}_0^{\mathrm{obj}}$.

Let $D : \mathcal{I} \to \mathcal{C}_0$ and $E : \mathcal{J} \to \mathcal{C}_0$ be objects of $\mathrm{Ind}_\kappa^\infty(\mathcal{C}_0)$ and let λ^D and λ^E be the limiting cocones used to define $\mathrm{Q}_0(D)$ and $\mathrm{Q}_0(E)$.

To show that Q_0 is faithful, let $[f]_E^{\pitchfork}, [g]_E^{\pitchfork} : D \to E$ be morphisms of $\mathrm{Ind}_\kappa^\infty(\mathcal{C}_0)$ with $\mathrm{Q}_0([f]_E^{\pitchfork}) = \mathrm{Q}_0([g]_E^{\pitchfork})$. Then

$$\lambda^E * f = \mathrm{Q}_0([f]_E^{\pitchfork}) * \lambda^D = \mathrm{Q}_0([g]_E^{\pitchfork}) * \lambda^D = \lambda^E * g .$$

By Lemma 3.6, this implies that $f \pitchfork_E g$. Hence, $[f]_E^{\pitchfork} = [g]_E^{\pitchfork}$.

To prove that Q_0 is full, let $f : \mathrm{Q}_0(D) \to \mathrm{Q}_0(E)$ 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 $\mathrm{Q}_0^{\mathrm{mor}}$, this implies that $\mathrm{Q}_0([t]_E^{\pitchfork}) = f$.

Hence, it remains to prove that every object $\mathfrak{a} \in \mathcal{C}$ is isomorphic to some object in rng $\mathrm{Q}_0^{\mathrm{obj}}$. Let $D : \mathfrak{J} \to \mathcal{C}$ 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 $E : \mathfrak{J} \to \mathcal{C}_0$. 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

$$\mathrm{Q}_0(E) = \varinjlim E \cong \varinjlim D = \mathfrak{a} ,$$

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 $D : \mathfrak{J} \to \mathcal{C}$ of κ-presentable objects, there exists a λ-directed diagram $D^+ : \mathfrak{J}^+ \to \mathcal{C}$ 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 $D : \mathfrak{J} \to \mathfrak{Dir}(\kappa)$ of partial orders of size $|D(i)| < \lambda$.*

Proof. (1) ⇒ (3) Let $D : \mathfrak{J} \to \mathcal{C}$ be a κ-directed diagram of κ-presentable objects. By (1) and Proposition 2.11, there exists a λ-directed diagram

$D^+ : \mathfrak{J}^+ \to \mathcal{C}$ 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 $D : \mathfrak{J} \to \mathcal{C}$ 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 $\langle \mathscr{P}_\kappa(X), \subseteq \rangle$ 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 $\mathscr{P}_\kappa(X)$ as the colimit of a λ-directed diagram $D : \mathfrak{J} \to \mathfrak{Dir}(\kappa)$ 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 $H \subseteq \mathscr{P}_\kappa(X)$ of size $|H| < \lambda$. By (4), there exists a λ-directed diagram $D : \mathfrak{J} \to \mathfrak{Dir}(\kappa)$ of partial orders of size less than λ with $\varinjlim D = \mathscr{P}_\kappa(X)$. 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 $\mathfrak{J}$ 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 $Y \in \mathscr{P}_\kappa(X)$. 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 $E : \mathfrak{J} \to \mathcal{C}$ 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 $\mathfrak{J}$ 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

$$\varinjlim D = \varinjlim (D \circ F) = \varinjlim E = \mathfrak{a}\,.$$ □

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 $X \setminus O$ 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

$$\|\bar{a}\| := \sqrt{a_0^2 + \cdots + a_{n-1}^2}\,,$$

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 $\mathcal{C} = \mathscr{P}(X)$ 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 $\mathscr{P}(X)$.

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

$$\mathcal{C} = \{ \bigcap Z \mid Z \subseteq \mathcal{B} \} .$$

(By convention, we set $\bigcap \varnothing := X$.)

(b) An *open base* of $\mathcal{C}$ is a system $\mathcal{B} \subseteq \wp(X)$ such that

$$\mathcal{C} = \{ X \smallsetminus \bigcup Z \mid Z \subseteq \mathcal{B} \} .$$

(c) A *closed subbase* of $\mathcal{C}$ is a system $\mathcal{B} \subseteq \wp(X)$ such that the set

$$\{ B_0 \cup \dots \cup B_{n-1} \mid n < \omega,\ B_i \in \mathcal{B} \}$$

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

$$f^{-1}[B_0 \cup \dots \cup B_{n-1}] = f^{-1}[B_0] \cup \dots \cup f^{-1}[B_{n-1}]$$

is closed, for all $B_0, \dots, B_{n-1} \in \mathcal{B}$. Hence, there is a close base

$$\mathcal{B}_+ := \{ B_0 \cup \dots \cup B_{n-1} \mid n < \omega,\ B_0, \dots, B_{n-1} \in \mathcal{B} \}$$

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 $X = C_0 \cup \dots \cup C_{n-1}$. If each restriction $f \restriction C_i$ is continuous then so is f.*

Proof. Let $A \subseteq Y$ be closed. Since $f \restriction C_i$ is continues it follows that the sets $f^{-1} \restriction C_i[A]$ are closed. Hence,

$$f^{-1}[A] = f^{-1} \restriction C_0[A] \cup \dots \cup f^{-1} \restriction C_{n-1}[A]$$

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*

$$(f \sqcup g)(x) := f(x) \sqcup g(x) \quad \text{and} \quad (f \sqcap g)(x) := f(x) \sqcap g(x) \,.$$

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*

$$f \sqsubseteq g \quad :\text{iff} \quad f(x) \sqsubseteq g(x) \,, \quad \text{for all } x \in X \,,$$

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 $b \ngeq c$. 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

$$\langle W \rangle := \{ x \in 2^\omega \mid w \leq x \text{ for some } w \in W \}$$

with $W \subseteq 2^{<\omega}$. ($\leq$ denotes the prefix order.)

Remark. The Cantor discontinuum can be regarded as the set of all branches of the infinite binary tree $\langle 2^{<\omega}, \leq \rangle$. 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 $x \in X \setminus \bigcup \mathcal{V}$ 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 $\mathcal{U} := \{ X \setminus C \mid C \in \mathcal{C} \}$. 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 $\mathcal{C}_0 := \{ X \setminus U \mid U \in \mathcal{U}_0 \} \subseteq \mathcal{C}$.

(3) ⇒ (1) Let $\mathcal{U}$ be an open cover of X and set $\mathcal{C} := \{ X \setminus U \mid U \in \mathcal{U} \}$. 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 $\{ X \setminus C \mid C \in \mathcal{C}_0 \}$ 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 $\langle 2^{<\omega} \setminus T, \preceq \rangle$ 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 $K \cap U_x = \emptyset$. 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 $\mathrm{int}(U)$ 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

$$\mathcal{W} := \{ \mathrm{int}(W_x) \mid x \in K \}$$

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

$$\begin{aligned} \mathrm{cl}(U) = \mathrm{cl}(\bigcup \mathcal{W}_0) &= \bigcup \{ \mathrm{cl}(\mathrm{int}(W_x)) \mid \mathrm{int}(W_x) \in \mathcal{W}_0 \} \\ &\subseteq \bigcup \{ W_x \mid \mathrm{int}(W_x) \in \mathcal{W}_0 \} \subseteq O . \end{aligned}$$

Finally, $\mathrm{cl}(U)$ 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 $x \in C \cap \mathrm{int}(K)$ and $C \cap \mathrm{int}(K)$ 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 $\mathrm{int}(K)$ and $Y \setminus K$ are disjoint open sets with $x \in \mathrm{int}(K)$ 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

$$\mathcal{B} := \big\{ \operatorname{pr}_k^{-1}[O] \bigm| k \in I,\ O \subseteq X_k \text{ open} \big\}.$$

Since $\mathcal{B}$ is an open subbase of $\mathcal{C}$ it follows that $\operatorname{pr}_k^{-1}[O]$ is open, for every open $O \subseteq X_k$. Hence, $\operatorname{pr}_k : \prod_i X_i \to X_k$ 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 $\operatorname{pr}_k^{-1}[O]$ 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) (⇐) 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, $U_* := \operatorname{pr}_i^{-1}[U]$ and $V_* := \operatorname{pr}_i^{-1}[V]$ are disjoint open sets with $(x_i)_i \in U_*$ and $(y_i)_i \in V_*$.

(⇒) 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) (⇒) 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 $\mathfrak{X}_k$.

(⇐) 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

$$\operatorname{pr}_{k_0}^{-1}[B_0] \cap \dots \cap \operatorname{pr}_{k_n}^{-1}[B_n]$$

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. (⇒) Let $\mathcal{U}$ be an open cover of X_i. Then

$$\mathcal{V} := \{ \operatorname{pr}_i^{-1}[U] \mid U \in \mathcal{U} \}$$

is an open cover of $\prod_i X_i$. Consequently, there exists a finite subcover $\mathcal{V}_0 \subseteq \mathcal{V}$ and $\{ U \in \mathcal{U} \mid \operatorname{pr}_i^{-1}[U] \in \mathcal{V}_0 \}$ is a finite subcover of $\mathcal{U}$.

(⇐) 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 $\operatorname{pr}_i^{-1}(U)$ where $i \in I$ and $U \subseteq X_i$ is open. For $i \in I$, let

$$\mathcal{U}_i := \{ U \subseteq X_i \mid \operatorname{pr}_i^{-1}[U] \in \mathcal{U} \}.$$

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 $f : \mathfrak{Y}_0 \times \dots \times \mathfrak{Y}_{n-1} \to \mathfrak{Z}$ and $g_i : \mathfrak{X}_i \to \mathfrak{Y}_i$, for $i < n$, be functions and define $h : \mathfrak{X}_0 \times \dots \times \mathfrak{X}_{n-1} \to \mathfrak{Z}$ 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 $k : \mathfrak{X}_0 \times \dots \times \mathfrak{X}_{n-1} \to \mathfrak{Y}_0 \times \dots \times \mathfrak{Y}_{n-1}$ 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 $O \subseteq X_0 \times \dots \times X_{n-1}$ be a basic open set. Then $O = U_0 \times \dots \times U_{n-1}$ where each U_i is open. Since g_i is continuous it follows that $g_i^{-1}[U_i]$ is also open. Consequently,

$$k^{-1}[O] = g_0^{-1}[U_0] \times \dots \times g_{n-1}^{-1}[U_{n-1}]$$

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 $U \times V \cap \Delta = \varnothing$. It follows that $X \times X \setminus \Delta$ 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 $A \cap O \neq \varnothing$, 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) $\mathrm{int}(A) = \varnothing$ *if and only if $X \setminus A$ is dense.*

Proof. (a) ($\Leftarrow$) Let O be a nonempty open set. Then $C := X \setminus O \neq X$. Since $\mathrm{cl}(A) = X$ it follows that $C \nsupseteq A$. This implies that $O \cap A \neq \varnothing$.

($\Rightarrow$) Let $C \supseteq A$ be closed and set $O := X \setminus C$. If $O \neq \varnothing$ then we have $O \cap A \neq \varnothing$ since A is dense. It follows that $A \setminus C \neq \varnothing$. 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 $O \cap (X \setminus A) = \varnothing$ then $O \subseteq A$ which implies that $\operatorname{int}(A) \neq \varnothing$. Conversely, if $O \subseteq A$ then $O \cap (X \setminus A) = \varnothing$ and $X \setminus A$ 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 $O \subseteq \operatorname{cl}_{\mathfrak{X}}(B)$ implies $O \cap D \subseteq \operatorname{cl}_{\mathfrak{D}}(B \cap D)$. Since $O \cap D$ is open in D we have $O \cap D \subseteq \operatorname{int}_{\mathfrak{D}}(\operatorname{cl}_{\mathfrak{D}}(B \cap D))$. 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 $\mathscr{P}(X)$ 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 $X \setminus A$ is dense.*

Proof. (1) ⇒ (2) If A_n is open and dense then $X \setminus A_n$ is a closed set with empty interior. By (1), it follows that $B = \bigcup_n (X \setminus A_n)$ has empty interior. Consequently, $\bigcap_{n<\omega} A_n = X \setminus B$ 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 $\operatorname{cl}(B_n)$ is empty it follows that $O_n := X \setminus \operatorname{cl}(B_n)$ is dense and open. (2) implies that the set $X \setminus A = \bigcap_n O_n$ 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 $X \setminus A$ is not dense. Then $\operatorname{int}(A) \neq \varnothing$ and there exists a nonempty open subset $O = \operatorname{int}(A) \subseteq A$ 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 $X \setminus B$ is dense. Consequently, we have $\operatorname{int}(B) = \varnothing$. □

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 $X \setminus A$ has the property of Baire. In particular, $X \setminus A$ is not meagre.*

Proof. Let A be a meagre subset of X. By Lemma 5.5 (4), it follows that $X \setminus A$ is dense. According to Lemma 5.4 (c), if B is a meagre set in $X \setminus A$ then B is also meagre in $\mathfrak{X}$. By Lemma 5.4 it follows that $A \cup B$ is also meagre. Consequently, $C = (X \setminus A) \setminus B = X \setminus (A \cup B)$ is dense in $\mathfrak{X}$ and, therefore, C is also dense in $X \smallsetminus A$. By Lemma 5.5, it follows that $X \smallsetminus A$ 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 $O_0 \cap \bigcap_n A_n \neq \emptyset$. 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 $C \neq \emptyset$. 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 $x \in \mathrm{cl}(A \smallsetminus \{x\})$. 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 $x \notin A$.

- $\mathrm{rk_{CB}}(x/A) \geq 0$ iff $x \in A$.
- $\mathrm{rk_{CB}}(x/A) \geq \alpha + 1$ if $\mathrm{rk_{CB}}(x/A) \geq \alpha$ and x is an accumulation point of the set $\{ a \in A \mid \mathrm{rk_{CB}}(a/A) \geq \alpha \}$.
- For limit ordinals δ, we set $\mathrm{rk_{CB}}(x/A) \geq \delta$ if $\mathrm{rk_{CB}}(x/A) \geq \alpha$, for all $\alpha < \delta$.

The Cantor-Bendixson rank of A is

$$\mathrm{rk_{CB}}(A) := \sup \{ \mathrm{rk_{CB}}(a/A) \mid a \in A \} .$$

Remark. A point a is an isolated point of A if and only if $\mathrm{rk_{CB}}(a/A) = 0$.

Proposition 5.12. *Let $\mathfrak{X}$ be a topological space. For $\alpha \in \mathrm{On} \cup \{\infty\}$, define*

$$X^{<\alpha} := \{ x \in X \mid \mathrm{rk_{CB}}(x/X) < \alpha \}$$

and set $X^{\geq\alpha} := X \setminus X^{<\alpha}$ and $X^{\alpha} := X^{\geq\alpha} \cap X^{<\alpha+1}$.

(a) $\mathrm{rk_{CB}}(X) \geq |X|^+$ *implies* $\mathrm{rk_{CB}}(X) = \infty$.

(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 $\mathrm{rk_{CB}}(X) \leq \alpha < \kappa^+$. Otherwise, $\mathrm{rk_{CB}}(X) = \infty$.

(b) Suppose that there is some element $x \in \mathrm{cl}(X^{\geq\alpha}) \setminus X^{\geq\alpha}$. Let $\beta := \mathrm{rk_{CB}}(x/X) < \alpha$. 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 $\mathrm{int}(X^{\infty}) = \varnothing$. Conversely, let $O \subseteq X$ be a nonempty open set. Choose some $a \in O$ such that $\alpha := \mathrm{rk_{CB}}(a/X) < \infty$ 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, $a \in O \cap X^{\circ} \neq \varnothing$, 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*

$$\mathrm{rk_{CB}}(c/C) = \mathrm{rk_{CB}}(c/X) .$$

Proof. We prove by induction on α that

$$\mathrm{rk_{CB}}(c/C) = \alpha \quad \text{iff} \quad \mathrm{rk_{CB}}(c/X) = \alpha .$$

Set

$$X^{\alpha} := \{ x \in X \mid \mathrm{rk_{CB}}(x/X) < \alpha \} ,$$
$$C^{\alpha} := \{ x \in C \mid \mathrm{rk_{CB}}(x/C) < \alpha \} .$$

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

$$\begin{aligned} \mathrm{rk_{CB}}(c/C) = \alpha \quad &\text{iff} \quad c \text{ is isolated in } C \setminus C^{\alpha} \\ &\text{iff} \quad c \text{ is isolated in } X \setminus X^{\alpha} \\ &\text{iff} \quad \mathrm{rk_{CB}}(c/X) = \alpha . \end{aligned}$$

□

Lemma 5.14. *Let $f : \mathfrak{X} \to \mathfrak{Y}$ be injective and continuous. For every $x \in X$, we have*

$$\mathrm{rk}_{\mathrm{CB}}(x/X) \le \mathrm{rk}_{\mathrm{CB}}(f(x)/Y)\,.$$

Proof. We prove by induction on α that

$$\mathrm{rk}_{\mathrm{CB}}(x/X) \ge \alpha \quad \text{implies} \quad \mathrm{rk}_{\mathrm{CB}}(f(x)/Y) \ge \alpha\,.$$

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 $\mathrm{rk}_{\mathrm{CB}}(x/X) \ge \alpha + 1$. Set

$$X^{\ge\alpha} := \{\, x \in X \mid \mathrm{rk}_{\mathrm{CB}}(x/X) \ge \alpha \,\}\,,$$
$$Y^{\ge\alpha} := \{\, y \in Y \mid \mathrm{rk}_{\mathrm{CB}}(y/Y) \ge \alpha \,\}\,.$$

By inductive hypothesis, we know that $f[X^{\ge\alpha}] \subseteq Y^{\ge\alpha}$. For a contradiction, suppose that $\mathrm{rk}_{\mathrm{CB}}(f(x)/Y) = \alpha$. Then $f(x)$ is an isolated point of $Y^{\ge\alpha}$ and we can find an open neighbourhood O of $f(x)$ such that $Y^{\ge\alpha} \cap O = \{f(x)\}$. Hence,

$$\begin{aligned}\{x\} &= f^{-1}[\{f(x)\}] = f^{-1}[Y^{\ge\alpha} \cap O] = f^{-1}[Y^{\ge\alpha}] \cap f^{-1}[O] \\ &\supseteq X^{\ge\alpha} \cap f^{-1}[O] \supseteq \{x\}\,.\end{aligned}$$

It follows that $X^{\ge\alpha} \cap f^{-1}[O] = \{x\}$ and x is an isolated point of $X^{\ge\alpha}$. 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 $X^\alpha = X^{\ge\alpha}$ 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 $\mathrm{rk}_{\mathrm{CB}}(X) \le \gamma < \alpha$. Contradiction. □

Lemma 5.16. *Let $\mathfrak{X}$ be a locally compact Hausdorff space. If $\mathrm{rk}_{\mathrm{CB}}(X) = \infty$ then $|X| \ge 2^{\aleph_0}$.*

Proof. Let $A := \{\, x \in X \mid \mathrm{rk}_{\mathrm{CB}}(x/X) = \infty \,\}$. We prove that $|A| \ge 2^{\aleph_0}$. 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 $v \le w$,
- if $v \nleq w$ and $w \nleq v$ then $U_v \cap U_w = \emptyset$.

By assumption $A \neq \emptyset$. 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 $C_\sigma \neq \emptyset$. Furthermore, we have $C_\sigma \cap C_\rho = \emptyset$, for $\sigma \neq \rho$. Consequently,

$$|A| \ge \sum_{\sigma\in 2^\omega} |C_\sigma| \ge 2^{\aleph_0}\,.$$

□

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 $L \smallsetminus \langle x \rangle = \langle x^* \rangle$ of a basic closed set is closed. □

Example. Let A be an infinite set. For the lattice $\mathfrak{F} = \langle F, \subseteq \rangle$ with

$$F := \{\, X \subseteq A \mid X \text{ or } A \smallsetminus X \text{ is finite} \,\}\,,$$

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 $\operatorname{rk}_{\mathrm{CB}}(\operatorname{spec}(\mathfrak{F})) = 1$.

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) : spec$(\mathfrak{K}) \to$ 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

$$\mathfrak{u} = f[f^{-1}[\mathfrak{u}]] = f[f^{-1}[\mathfrak{v}]] = \mathfrak{v} .$$ ☐

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 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 rng f induces a subalgebra of $\mathfrak{B}$. Hence, if 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 spec(f) is not injective.

(d) For a contradiction, suppose that 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 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

$$a \in \mathfrak{u} = f^{-1}[\mathfrak{v}] \quad \Rightarrow \quad f(a) \in \mathfrak{v}$$
$$\Rightarrow \quad f(b) \in \mathfrak{v} \quad \Rightarrow \quad b \in f^{-1}[\mathfrak{v}] = \mathfrak{u}\,.$$

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

$$\langle W \rangle := \{\, x \in 2^\omega \mid w \leq x \text{ for some } w \in W \,\}$$

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 $\langle x \rangle = \operatorname{spec}(\mathfrak{B}) \smallsetminus \langle x^* \rangle$. 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 $x \in \mathfrak{u} \smallsetminus \mathfrak{v}$. This implies that $x^* \in \mathfrak{v} \smallsetminus \mathfrak{u}$. 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

$$U = \langle x_0 \rangle \cup \dots \cup \langle x_n \rangle = \langle x_0 \sqcup \dots \sqcup x_n \rangle$$

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 $\bigcap \mathfrak{u} \neq \emptyset$. 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 $\langle \wp(X), \cap, \cup, {}^*, \emptyset, X \rangle$.*

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 $\operatorname{rk}_{\mathrm{CB}}(\mathfrak{S}) = 0$ 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 $\operatorname{rk}_{\mathrm{P}}(a) < \infty$, for every $a \in B$, then then $|\operatorname{spec}(\mathfrak{B})| \leq |B|$.*

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 $X_\alpha \cap X_\beta = \varnothing$.

Hence, it is sufficient to prove that $X_\alpha \neq \varnothing$, for all $\alpha \in \lambda^{\kappa}$. For finitely many elements $w_0 \prec \dots \prec w_n \prec \alpha$, we have

$$\langle a_{w_0} \rangle \cap \dots \cap \langle a_{w_n} \rangle = \langle a_{w_0} \sqcap \dots \sqcap a_{w_n} \rangle = \langle a_{w_n} \rangle \neq \varnothing .$$

Thus, the family $\langle a_w \rangle_{w \prec \alpha}$ has the finite intersection property and, by compactness, it follows that $X_\alpha = \bigcap_{w \prec \alpha} \langle a_w \rangle \neq \varnothing$. □

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 $\operatorname{rk_P}(a) < \infty$ then there exists an ultrafilter $\mathfrak{u} \in \langle a \rangle$ with $\operatorname{rk_P}(\mathfrak{u}) = \operatorname{rk_P}(a)$.*

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

$$\operatorname{rk_P}(a) = \operatorname{rk_P}(a \sqcap c_{\mathfrak{u}_i}) .$$

This implies that

$$\operatorname{rk_P}(\mathfrak{u}_i) \leq \operatorname{rk_P}(a) = \operatorname{rk_P}(a \sqcap c_{\mathfrak{u}_i}) \leq \operatorname{rk_P}(c_{\mathfrak{u}_i}) = \operatorname{rk_P}(\mathfrak{u}_i) .$$ □

Corollary 7.7. *Let $\mathfrak{B}$ be a boolean algebra and $a \in B$.*

$$\operatorname{rk_P}(a) = \sup \{ \operatorname{rk_P}(\mathfrak{u}) \mid \mathfrak{u} \in \langle a \rangle \} .$$

Proof. If $\mathfrak{u} \in \langle a \rangle$, then $a \in \mathfrak{u}$ implies that $\operatorname{rk_P}(\mathfrak{u}) \leq \operatorname{rk_P}(a)$. Conversely, we can use Lemma 7.6 to find some ultrafilter $\mathfrak{u} \in \langle a \rangle$ with $\operatorname{rk_P}(\mathfrak{u}) = \operatorname{rk_P}(a)$. □

Theorem 7.8. *Let $\mathfrak{B}$ be a boolean algebra. For every $\mathfrak{u} \in \operatorname{spec}(\mathfrak{B})$, we have*

$$\operatorname{rk_P}(\mathfrak{u}) = \operatorname{rk_{CB}}(\mathfrak{u} / \operatorname{spec}(\mathfrak{B})) .$$

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

$$\begin{aligned}
\mathrm{rk}_{\mathrm{CB}}(\langle a \rangle) &= \sup\big\{\, \mathrm{rk}_{\mathrm{CB}}(\mathfrak{u}/\langle a \rangle) \bigm| \mathfrak{u} \in \langle a \rangle \,\big\} \\
&= \sup\big\{\, \mathrm{rk}_{\mathrm{CB}}(\mathfrak{u}/\operatorname{spec}(\mathfrak{B})) \bigm| \mathfrak{u} \in \langle a \rangle \,\big\} \\
&= \sup\{\, \mathrm{rk}_{\mathrm{P}}(\mathfrak{u}) \mid \mathfrak{u} \in \langle a \rangle \,\} \\
&= \mathrm{rk}_{\mathrm{P}}(a)\,.
\end{aligned}$$ □

Corollary 7.10. *Let $\mathfrak{S}$ be a Stone space and $C \subseteq S$ closed.*

$$\mathrm{rk}_{\mathrm{CB}}(C) = \mathrm{rk}_{\mathrm{P}}(C/\operatorname{clop}(C))$$

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 $\mathfrak{B} := \operatorname{clop}(\mathfrak{C})$. Then $\operatorname{spec}(\mathfrak{B}) \cong \mathfrak{C}$ and Corollary 7.9 implies that

$$\mathrm{rk}_{\mathrm{CB}}(C) = \mathrm{rk}_{\mathrm{CB}}(\operatorname{spec}(\mathfrak{B})) = \mathrm{rk}_{\mathrm{P}}(\top/\mathfrak{B}) = \mathrm{rk}_{\mathrm{P}}(C/\operatorname{clop}(\mathfrak{C}))\,.$$ □

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* 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 $\langle A \rangle \setminus C = \langle b^* \rangle \cap \langle A \rangle$ 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

$$C = \langle D \rangle \cap \langle A \rangle \quad \text{and} \quad \langle A \rangle \setminus C = \langle E \rangle \cap \langle A \rangle\,.$$

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 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 $D \in \mathrm{clop}(C)$. Then*

$$\mathrm{clop}(D) = \{\, E \in \mathrm{clop}(C) \mid E \subseteq D \,\}\,.$$

Proof. Let $\mathfrak{B} := \mathrm{clop}(\mathfrak{S})$. By Lemma 7.11, there is some $A \in B$ such that $D = A \cap C$. By the same lemma it follows that

$$\begin{aligned} E \in \mathrm{clop}(D) \quad &\text{iff} \quad E = A' \cap D \text{ for some } A' \in B \\ &\text{iff} \quad E = A' \cap A \cap C \text{ for some } A' \in B \\ &\text{iff} \quad E = A'' \cap C \text{ for some } A'' \in B \text{ with } A'' \subseteq A \\ &\text{iff} \quad E \in \mathrm{clop}(C) \text{ and } E \subseteq D\,. \end{aligned}$$

□

Corollary 7.13. *Let $\mathfrak{S}$ be a Stone space, $C \subseteq S$ closed, and $D \in \mathrm{clop}(C)$. Then*

$$\mathrm{rk}_{\mathrm{P}}(D/\mathrm{clop}(D)) = \mathrm{rk}_{\mathrm{P}}(D/\mathrm{clop}(C))\,.$$

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 $C \subseteq T$,*

$$\mathrm{rk}_{\mathrm{CB}}(C/\mathfrak{T}) \leq \mathrm{rk}_{\mathrm{CB}}(f^{-1}[C]/\mathfrak{S})\,.$$

Proof. We prove by induction on α that

$$\mathrm{rk}_{\mathrm{CB}}(C/\mathfrak{T}) \geq \alpha \quad \text{implies} \quad \mathrm{rk}_{\mathrm{CB}}(f^{-1}[C]/\mathfrak{S}) \geq \alpha\,.$$

For $\alpha = 0$, surjectivity of f implies that

$$\begin{aligned} \mathrm{rk}_{\mathrm{CB}}(C/\mathfrak{T}) \geq 0 \quad &\text{iff} \quad C \neq \emptyset \\ &\text{iff} \quad f^{-1}[C] \neq \emptyset \\ &\text{iff} \quad \mathrm{rk}_{\mathrm{CB}}(f^{-1}[C]/\mathfrak{S}) \geq 0\,. \end{aligned}$$

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

$$\mathrm{rk}_{\mathrm{P}}\big(f^{-1}[D_n] \big/ \mathrm{clop}(f^{-1}[C])\big) \geq \alpha\,,$$

it follows that

$$\mathrm{rk}_{\mathrm{P}}\big(f^{-1}[C] \big/ \mathrm{clop}(f^{-1}[C])\big) \geq \alpha + 1\,.$$

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:

$$a \circ (b \circ c) = (a \circ b) \circ c \qquad \text{(associativitiy)}$$
$$a \circ e = a = e \circ a \qquad \text{(neutral element)}$$

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

$$a \circ a^{-1} = e \qquad \text{(inverse)}$$

(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*

$$\mathfrak{Aut}\,\mathfrak{M} = \langle \mathrm{Aut}\,\mathfrak{M}, \circ, {}^{-1}, \mathrm{id}_M \rangle$$

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

$$|G| = \Big|\bigcup(G/U)\Big| = \sum_{gU \in G/U} |gU| = \sum_{gU \in G/U} |U| = |G/U| \otimes |U|. \qquad \square$$

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*

$$\kappa \otimes |G| = \sum_{g \in G} |\text{fix}\, g| \quad \textit{where} \quad \text{fix}\, g := \{\, a \in \Omega \mid ga = a \,\} .$$

Proof. For each orbit of $\mathfrak{G}$, fix one representative $a_i \in \Omega$, $i < \kappa$. It follows that

$$\begin{aligned} \kappa \otimes |G| &= \sum_{i<\kappa} |G| = \sum_{i<\kappa} |G(a_i)| \otimes |G_{(a_i)}| = \sum_{i<\kappa} \sum_{b \in G(a_i)} |G_{(a_i)}| \\ &= \sum_{i<\kappa} \sum_{b \in G(a_i)} |G_{(b)}| = \sum_{b \in \Omega} |G_{(b)}| \\ &= \big|\{\, \langle g, b\rangle \in G \times \Omega \;\big|\; gb = b \,\}\big| = \sum_{g \in G} |\text{fix}\, g| . \end{aligned} \qquad \square$$

Corollary 2.8. *If $\mathfrak{G}$ is a finite group acting on Ω then the number of orbits is*

$$\frac{1}{|G|} \sum_{g \in G} |\text{fix}\, g| .$$

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 $\quad G \neq \bigcup_{i \in I} a_i H_i$, *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

$$\Big|\bigcap_{i \in I} H_i / H\Big| \leq (n - |I|)! , \quad \text{for all nonempty } I \subseteq [n] .$$

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

$$H = \bigcup_{h \in H} h G_{(\bar{a})} = \bigcup_{h \in H} \langle \bar{a} \mapsto h\bar{a} \rangle$$

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 $R_\Delta^{\mathfrak{M}} := \Delta$ 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

$$p = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 ,$$

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

$$F(h)\big(\textstyle\sum_n a_n x^n\big) := \textstyle\sum_n h(a_n) x^n .$$

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 $p = c_n x^n + \dots + c_1 x + c_0$, we define

$$h(p) := h_0(c_n) a^n + \dots + h_0(c_1) a + h_0(c_0) .$$

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

$$\begin{aligned} r(sa) &= (rs)a , \\ r(a+b) &= ra + rb , \quad \text{for all } r, s \in R \text{ and } a, b \in M . \\ (r+s)a &= ra + sa , \end{aligned}$$

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

$$\begin{aligned} 0a &:= 0 , \\ (n+1)a &:= na + a , \quad \text{for } n \in \mathbb{N} \text{ and } a \in A . \\ (-n)a &:= -(na) , \end{aligned}$$

(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,

$$a = -r^{-1} s_0 b_0 - \dots - r^{-1} s_{n-1} b_{n-1} \in \langle\langle I \rangle\rangle_{\mathfrak{V}} . \qquad \square$$

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

$$a = s_0 b_0 + \dots + s_{n-1} b_{n-1} .$$

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 $h \upharpoonright B = h_0$.*

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 $h \upharpoonright B = h_0$ 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 $a \in \langle\!\langle I \rangle\!\rangle_{\mathfrak{V}} \smallsetminus \langle\!\langle I_0 \rangle\!\rangle_{\mathfrak{V}}$, there exists ane element $b \in I \smallsetminus I_0$ such that $(I \smallsetminus \{b\}) \cup \{a\}$ is linearly independent and $b \in \langle\!\langle (I \smallsetminus \{b\}) \cup \{a\} \rangle\!\rangle_{\mathfrak{V}}$.*

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 $b_i \in I \smallsetminus I_0$, for some i. By renumbering the elements we may assume that $b_0 \in I \smallsetminus I_0$. 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 $b_0 \in \langle\!\langle (I \smallsetminus b_0) \cup \{a\} \rangle\!\rangle_{\mathfrak{V}}$. Hence, it remains to prove that $(I \smallsetminus b_0) \cup \{a\}$ 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

$$p = a_m x^m + \dots + a_0 \quad \text{and} \quad q = b_n x^n + \dots + b_0\,,$$

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 $q \in \mathfrak{a}$. 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 $h_a \restriction R = \mathrm{id}$. 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, $s = a_{n-1}x^{n-1} + \dots + a_0$, for some $a_0, \dots, a_{n-1} \in K$, and

$$\pi(q) = \pi(s) = a_{n-1}\pi(x^{n-1}) + \dots + a_1\pi(x) + a_0 .$$

It remains to prove that $1, \pi(x), \dots, \pi(x^{n-1})$ are linearly independent. For a contradiction, suppose that there are nonzero coefficients $a_0, \dots, a_{n-1} \in K$ such that

$$a_0 + a_1\pi(x) + \dots + a_{n-1}\pi(x^{n-1}) = 0 .$$

Then there is some $b \in K[x]$ such that

$$a_0 + a_1 x + \dots + a_{n-1}x^{n-1} = bp .$$

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 $a_0, \dots, a_{n-1} \in X$ and a polynomial $p \in U[x_0, \dots, x_{n-1}]$ such that $p[a_0, \dots, a_{n-1}] = 0$. 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 $a \in L \smallsetminus \operatorname{rng} h$ is algebraic over $\operatorname{rng} h$. Similarly, we call h *transcendental* if every $a \in L \smallsetminus \operatorname{rng} h$ 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 $K \cup \{a\} \subseteq \operatorname{rng} h$. The kernel $\operatorname{Ker} h$ is an ideal of $\mathfrak{K}[x]$. By Lemma 5.14, there exists a polynomial $p \in K[x]$ such that $\operatorname{Ker} h = (p)$. Let $\pi : \mathfrak{K}[x] \to \mathfrak{K}[x]/(p)$ be the canonical projection. By Theorem B1.4.12, there exists an isomorphism $g : \mathfrak{K}[x]/(p) \to \operatorname{rng} h = \mathfrak{K}(a)$ 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*

$$p = (x-a)^n q , \quad \text{for } a \in L,\ q \in L[x],\ n < \omega .$$

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 $a \in K \smallsetminus I$, 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

$$f : K \smallsetminus I \to \bigcup_{n<\omega} (\mathbb{Q}[x, \bar{y}] \times I^n) .$$

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 $x = x_0^2 + \dots + x_m^2$ and $y = y_0^2 + \dots + y_n^2$. Then

$$-x_0^2 = x_1^2 + \dots + x_m^2 + y_0^2 + \dots + y_n^2$$

implies

$$-1 = (x_1/x_0)^2 + \dots + (x_m/x_0)^2 + (y_0/x_0)^2 + \dots + (y_n/x_0)^2\,.$$

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

$$-1 = \sum_{i<n} (b_i + c_i\sqrt{a})^2 = \sum_{i<n} (b_i^2 + 2b_ic_i\sqrt{a} + ac_i^2) .$$

Since $\mathfrak{K}(\sqrt{a})$ is a $\mathfrak{K}$-vector space with basis $\{1, \sqrt{a}\}$ it follows that

$$-1 = \sum_{i<n} (b_i^2 + ac_i^2) \quad \text{and} \quad 0 = \sum_{i<n} 2b_ic_i\sqrt{a} .$$

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

$$-1 = b_0^2 + \cdots + b_k^2 .$$

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

$$-1 \equiv q_0^2 + \cdots + q_k^2 \pmod{p} .$$

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,

$$\begin{aligned} -1 &= \sum_{i<n} a_i (u_i + v_i \sqrt{b_{k-1}})^2 \\ &= \sum_{i<n} \bigl(a_i u_i^2 + a_i b_{k-1} v_i^2 + 2 a_i u_i v_i \sqrt{b_{k-1}} \bigr) . \end{aligned}$$

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

$$-1 = \sum_{i<n} (a_i u_i^2 + a_i b_{k-1} v_i^2) \quad \text{and} \quad 0 = \sum_{i<n} 2 a_i u_i v_i \sqrt{b_{k-1}} .$$

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 $-\sqrt{a^2+b^2} \le a \le \sqrt{a^2+b^2}$ 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 $a_i + a_k + c a_i a_k$. By inductive hypothesis, one of them is in $\mathfrak{K}(\sqrt{-1})$. Suppose that $a_i + a_k + c a_i a_k \in K(\sqrt{-1})$.

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

$$a \le b \quad :\text{iff} \quad b - a \text{ is a sum of squares}$$

on $\mathfrak{K}$. We claim that $\le$ is linear. This implies that $\mathfrak{K}$ is real.

It is sufficient to show that every element $a \in K$ satisfies $a \ge 0$ or $-a \ge 0$. 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 $p = x^{2n+1} + a_{2n}x^{2n} + \dots + a_1 x + a_0$ be a polynomial of odd degree. Choose $b \in K$ such that $b > 1 + |a_0| + \dots + |a_{2n}|$. 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

$$q = c_n x^n + \dots + c_1 x + c_0.$$

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

$$\begin{aligned} q[z] &= c_0 + c_1 z + \dots + c_n z^n \\ &\ge c_0 - \varepsilon|c_1| - \dots - \varepsilon^n |c_n| \\ &\ge c_0 - \varepsilon k - \dots - \varepsilon k \\ &= c_0 - \varepsilon kn \\ &\ge \frac{c_0}{2} = \frac{p[a]}{2} > 0 . \end{aligned}$$

□

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*

$$\begin{aligned} p[z] &> p[a] , &&\text{for } a < z < a + \varepsilon , \\ p[z] &< p[a] , &&\text{for } a - \varepsilon < z < a . \end{aligned}$$

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

$$\begin{aligned} & q[z] > 0 , && \text{for } 0 < z < \varepsilon , \\ \text{and} \quad & q[z] < 0 , && \text{for } -\varepsilon < z < 0 . \end{aligned}$$

□

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*

$$p'[z] \ge 0 , \quad \text{for all } a \le z \le b ,$$

then $p[a] < p[b]$.

Proof. First, suppose that $p'[z] > 0$, for all $a \le z \le b$. If $p[a] \ge p[b]$ then applying Lemma 6.19 to a and b, respectively, we obtain elements $a < c < d < b$ with $p[d] < p[b] \le p[a] < p[c]$. 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

$$\begin{aligned} a < \frac{a + c_0}{2} < c_0 < \frac{c_0 + c_1}{2} < c_2 < \dots \\ < \frac{c_{k-2} + c_{k-1}}{2} < c_{k-1} < \frac{c_{k-1} + b}{2} < b . \end{aligned}$$

It is sufficient to prove that $p[d_i] < p[d_{i+1}]$, for all $i \le 2k$. 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 $p[a] \le p[b]$. By the same argument it follows that $p[a] \le p[(a + b)/2]$. Hence, the first part of the proof implies that $p[a] \le p[(a + b)/2] < p[b]$, 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 $a \in L_0 \smallsetminus K$ 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 $a_0 < \dots < a_{l-1}$ be an increasing enumeration of all roots of p in L_0 and let $b_0 < \dots < b_{m-1}$ 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 $a \in L_0 \smallsetminus L_0^\alpha$ 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 $d \in L_1 \smallsetminus L_1^\alpha(b)$ and extend the isomorphism to $\pi_{\alpha+1} : \mathfrak{L}_0^\alpha(a, c) \to \mathfrak{L}_1^\alpha(b, d)$, for some $c \in L_0$. □

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 E6

Chapter E7

Chapter F1

Chapter F2

Chapter F3

Chapter F4

Chapter F5

Chapter G1

Chapter G2

Index

The Roman and Fraktur alphabets							
A	*a*	𝔄	𝔞	*N*	*n*	𝔑	𝔫
B	*b*	𝔅	𝔟	*O*	*o*	𝔒	𝔬
C	*c*	ℭ	𝔠	*P*	*p*	𝔓	𝔭
D	*d*	𝔇	𝔡	*Q*	*q*	𝔔	𝔮
E	*e*	𝔈	𝔢	*R*	*r*	ℜ	𝔯
F	*f*	𝔉	𝔣	*S*	*s*	𝔖	ſ 𝔰
G	*g*	𝔊	𝔤	*T*	*t*	𝔗	𝔱
H	*h*	ℌ	𝔥	*U*	*u*	𝔘	𝔲
I	*i*	ℑ	𝔦	*V*	*v*	𝔙	𝔳
J	*j*	𝔍	𝔧	*W*	*w*	𝔚	𝔴
K	*k*	𝔎	𝔨	*X*	*x*	𝔛	𝔵
L	*l*	𝔏	𝔩	*Y*	*y*	𝔜	𝔶
M	*m*	𝔐	𝔪	*Z*	*z*	ℨ	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