Ultraproducts

David Pierce

September 15, 2014
reformated July 23, 2015

Mimar Sinan Güzel Sanatlar Üniversitesi
Matematik Bölümü
http://mat.msgsu.edu.tr/~dpierce/

Ultraproducts

Mathematics Department
Mimar Sinan Fine Arts University
Istanbul, Turkey
`http://mat.msgsu.edu.tr/~dpierce/`
`dpierce@msgsu.edu.tr`

Preface

In preparing the first edition of this text, I tried to write down everything that I might talk about in an upcoming course on ultraproducts. I had no clear plan for a coherent whole. From my records, here is a summary of the six days of the course (August 13–19, 2012, Monday to Sunday, with Thursday off, 8–10 o'clock in the morning):

1. $\mathbb{R}^\omega/M$.
2. The ordering of $\mathbb{R}^\omega/M$; bad statement of Łoś's Theorem.
3. Better statement of Łoś's Theorem; proof.
4. Compactness.
5. Voting (Arrow's Theorem).
6. The ultraproduct scheme.

Later I edited the text for my own use in a two-week course, in July, 2014. For the sake of completeness, at least, I incorporated more background. In the fall of 2013, I had taught a graduate course on groups and rings, and I thoroughly edited my notes for *that* course; then I took sections from those notes to add to the present ones.

I added and rearranged a lot. I worked out quite generally the notion of a Galois correspondence and its relation to topology. I also investigated the Axiom of Choice and distinguished the results that need it from those that need only the Prime Ideal Theorem. Some of this work would be relevant to a talk on the Compactness Theorem of logic given at the Caucasian Mathematics Conference, Tbilisi, September 5–6, 2014, and then again at a tutorial on the Compactness Theorem given June 20–1, 2015, at the 5th World Congress and School on Universal Logic, Istanbul.

I have not properly revisited the text since 2014. It is still quite rough. It uses more field theory than it actually develops. It grew so long that to read it straight through, checking for coherence, would be difficult. I have not done this. I *did* try to add many cross-references.

Preface to the first edition

These notes are for a course called Ultraproducts and Their Consequences, to be given at the Nesin Mathematics Village in Şirince, Selçuk, İzmir, Turkey, in August, 2012. The notes are mainly for my use; they do not constitute a textbook, although parts of them may have been written in textbook style. The notes have not been thoroughly checked for correctness; writing the notes has been my own way of learning some topics.

The notes have grown like a balloon, at all points: I have added things here and there as I have seen that they are needed or useful. I have also rearranged sections. There is too much material here for a week-long course. Some of the material is background necessary for thorough consideration of some topics; this background may be covered in a simultaneous course in Şirince.

The catalogue listing for the course[1](with abstract as submitted by me on January 27, 2012) is as follows.

Title of course: Ultraproducts and their consequences

Instructor: Assoc. Prof. David Pierce

Institution: Mimar Sinan GSÜ

Dates: 13–19 Ağustos 2012

Prerequisites: Some knowledge of algebra, including the theorem that a quotient of a ring by an ideal is a field if and only if the ideal is maximal.

Level: Advanced undergraduate and graduate

Abstract: An ultraproduct is a kind of average of infinitely many structures. The construction is usually traced to a 1955 paper

[1] From `http://matematikkoyu.org/etkinlikler/2012-tmd-lisans-lisansustu/ultra_pierce.pdf`, to which there is a link on `http://matematikkoyu.org/etkinlikler/2012-tmd-lisans-lisansustu/` as of August 6, 2012.

> of Jerzy Los; however, the idea of an ultraproduct can be found in Kurt Goedel's 1930 proof (from his doctoral dissertation) of the Completeness Theorem for first-order logic. Non-standard analysis, developed in the 1960s by Abraham Robinson, can be seen as taking place in an ultraproduct of the ordered field of real numbers: more precisely, in an ultrapower. Indeed, for each integer, the 'average' real number is greater than that integer; therefore an ultrapower of the ordered field of real numbers is an ordered field with infinite elements and therefore infinitesimal elements. Perhaps the first textbook of model theory is Bell and Slomson's *Models and Ultraproducts* of 1969: the title suggests the usefulness of ultraproducts in the development various model-theoretic ideas. Our course will investigate ultraproducts, starting from one of the simplest interesting examples: the quotient of the cartesian product of an infinite collection of fields by a maximal ideal that has nontrivial projection onto each coordinate. No particular knowledge of logic is assumed.

Such was the abstract that I submitted in January. I have written the following notes since then, by way of working out for myself some of the ideas that might be presented in the course. I have tried to emphasize examples. In some cases, I may have sacrificed generality for concreteness. A theorem that I might have covered, but have not, is the theorem of Keisler and Shelah that elementary equivalence is the same thing as isomorphism of ultrapowers.

Contents

List of Figures

1. Introduction

In this text, the **natural numbers** begin with 0 and compose the set $\upomega$. Thus,[1]

$$\upomega = \{0, 1, 2, \dots\}.$$

We shall use this set in two ways:

1) as an index-set for countably infinite sequences $(a_k \colon k \in \upomega)$;
2) as the cardinal number of each countably infinite set.

We shall also make use of the following feature of the elements of $\upomega$: each of them is a set whose cardinal number is itself. That is, each n in $\upomega$ is an n-element set. More precisely,

$$n = \{0, \dots, n-1\},$$

so that

$$0 = \varnothing, \qquad 1 = \{0\}, \qquad 2 = \{0, 1\}, \qquad 3 = \{0, 1, 2\},$$

and so on. If k and n are in $\upomega$, then

$$k \in n \iff k \subset n;$$

in this case we may write simply

$$k < n.$$

[1] The letter $\upomega$ is not the minuscule English letter called *double u,* but the minuscule Greek *omega,* which is probably in origin a double o. Obtained with the control sequence `\upomega` from the `upgreek` package for LaTeX, the $\upomega$ used here is upright, unlike the standard slanted ω (obtained with `\omega`). The latter ω might be used as a variable. We shall similarly distinguish between the constant $\uppi$ (used for the ratio of the circumference to the diameter of a circle, as well as for the *coordinate projections* defined on page 70) and the variable π.

If A and B are sets, then a **function** from B to A is just a subset f of $B \times A$ such that, for every x in B, there is exactly one y in A such that $(x, y) \in f$. In this case we write

$$y = f(x).$$

Then the function f is the set

$$\{(x, f(x)) \colon x \in B\}.$$

We may abbreviate this as

$$x \mapsto f(x);$$

this notation is useful when we do not actually have a single letter for f itself, but have an expression for $f(x)$. When we do have a letter like f, then, in place of $f(x)$, we may use one of the notations

$$f_x, \qquad\qquad f^x$$

(see below). The set of all functions from B to A will be denoted by

$$A^B.$$

If $f \in A^B$, then B is the **domain** of f, while the **range** of f is the subset

$$\{f(x) \colon x \in B\}$$

of A. One may say that A is a **codomain** of f, but in this case, if $A \subseteq C$, then C is also a codomain of f. In the expression for the range of f, if we replace the braces with round brackets (parentheses), we obtain

$$(f(x) \colon x \in B),$$

which we shall understand as yet another notation for the function f itself (strictly, we may understand it as an *indexed set*: see page 69).

As a special case of the foregoing notation, if $n \in \omega$, we have

$$A^n = \{\text{functions from } n \text{ to } A\}.$$

Instead of $(b_k\colon k<n)$ or $(b^k\colon k<n)$, an element of A^n may be written as one of

$$(b_0,\dots,b_{n-1}), \qquad\qquad (b^0,\dots,b^{n-1}).$$

In a slight departure from the foregoing notation, we may abbreviate this element of A^n by

$$\boldsymbol{b},$$

in boldface: it is an n-**tuple** of elements of A. We shall occasionally use both upper and lower indices at the same time, as for example in consideration of sequences $(\boldsymbol{b}_k\colon k\in\omega)$, where $\boldsymbol{b}_k\in A^n$, so that

$$\boldsymbol{b}_k=(b_k^0,\dots,b_k^{n-1}).$$

Note that

$$A^0=\{0\}=1.$$

According to what seems to be all but universal usage today, the ring of (rational) integers is

$$\mathbb{Z};$$

this is a sub-ring of the field

$$\mathbb{Q}$$

of rational numbers, which is in turn is a subfield of the field

$$\mathbb{R}$$

of real numbers.

We shall use $\mathbb{N}$ to denote the set of *positive* integers, so that

$$\mathbb{N}=\{1,2,3,\dots\}.$$

Literally then ω is the set $\{0\}\cup\mathbb{N}$ of *non-negative* integers. However, when we consider an element n of ω as an integer and hence as a rational number, we are not interested in the internal structure of n as a set. This is a reason why it may be useful to introduce the notation $\mathbb{N}$. It is useful *not* to put 0 in $\mathbb{N}$, because then we can describe the set $\mathbb{Q}^+$ of positive rational numbers as $\{x/y\colon(x,y)\in\mathbb{N}\times\mathbb{N}\}$ (see page 49).

2. Mathematical foundations

2.1. Sets as collections

Most objects of mathematical study can be understood as *sets.* A set is a special kind of *collection.* A **collection** is many things, considered as one thing. Those many things are the **members** or **elements** of the collection. The members **compose** the collection, and the collection **comprises** them.[1] Each member **belongs** to the collection and is **in** the collection, and the collection **contains** the member.

We shall designate certain collections as **sets.** We shall not define the collection of all sets; rather, we shall identify some rules for obtaining sets that will allow us to do the mathematics that we want. These rules will be expressed by *axioms.* We shall use versions of the so-called Zermelo–Fraenkel Axioms with the Axiom of Choice. The collection of these axioms is denoted by ZFC. Most of these axioms were described by Zermelo in 1908 [63].

We study study sets axiomatically, because a naïve approach can lead to contradictions. For example, one might think naïvely that there was a collection of all collections. But there can be no such collection, because if there were, then there would be a collection of all collections that did not contain themselves, and *this* collection would contain itself if and only if it did not. This result is the **Russell Paradox,** described in a letter [51] from Russell to Frege in 1902.

The elements of every set will be sets themselves. This is a conceptual and notational convenience that will turn out to be adequate for our purposes, even though, in ordinary life, the members of a collection are not usually collections themselves.

By the definition to be given officially on page 18, two sets will be

[1] Thus the relations named by the verbs "compose" and "comprise" are converses of one another; but native English speakers often confuse these two verbs.

equal if they have the same elements.[2] There will be an *empty set,* denoted by

$$\varnothing;$$

this will have no elements. If a is a set, then there will be a set denoted by

$$\{a\},$$

with the unique element a. If b is also a set, then there will be a set denoted by

$$a \cup b,$$

whose members are precisely the members of a and the members of b. Thus there will be sets $a \cup \{b\}$ and $\{a\} \cup \{b\}$; the latter is usually written as

$$\{a, b\}.$$

If c is another set, we can form the set $\{a, b\} \cup \{c\}$, which we write as

$$\{a, b, c\},$$

and so forth. This will allow us to build up the following infinite sequence:

$$\varnothing, \qquad \{\varnothing\}, \qquad \{\varnothing, \{\varnothing\}\}, \qquad \Big\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\Big\}, \qquad \dots$$

By definition, these sets will be the natural numbers 0, 1, 2, 3, ... To be more precise, they are the **von Neumann natural numbers** [60].

[2] This definition of equality is usually an axiom, rather than a definition. That is because equality is confused with *identity,* and the identity of two objects is considered to be an inherent property of the objects themselves, rather than a property that we assign to them. By this way of thinking, we say $1/2 = 2/4$ because the expressions $1/2$ and $2/4$ are names of the same equivalence-class $\{(x, y) \in \mathbb{N} \times \mathbb{N} \colon 2x = y\}$. But we can just as well say that, if a, b, c, and d are positive integers, then, *by definition,* the expression $a/b = c/d$ means that the products ad and bc are the same.

2.2. Set theory

2.2.1. Notation

Our formal axioms for set theory will be written in a certain *logic,* whose symbols are:

1) **variables,** as x, y, and z;
2) **constants,** as a, b, and c, or A, B, and C;
3) the symbol $\in$ denoting the membership relation;
4) the **Boolean connectives** of propositional logic:
    a) the singulary connective $\neg$ ("not"), and
    b) the binary connectives
        i) $\vee$ ("or"),
        ii) $\wedge$ ("and"),
        iii) $\Rightarrow$ ("implies"), and
        iv) $\Leftrightarrow$ ("if and only if");
5) parentheses;
6) the **quantification symbols**
    a) $\exists$ ("there exists") and
    b) $\forall$ ("for all").

We could do without constants as distinct from variables; but they seem to be useful. The distinction between constants and variables can be traced back at least as far as Descartes's *Geometry* [14] of 1637, where letters like a, b, and c are used for known lengths, and z, y, and x, for unknown lengths.

A variable or a constant is called a **term.** If t and u are terms, then the expression

$$t \in u$$

is called an **atomic formula.** It means t is a member of u. From atomic formulas, other formulas are built up *recursively* by use of the symbols above, according to certain rules, as follows:

1. If φ is a formula, then so is its **negation** $\neg\varphi$.
2. If φ and ψ are formulas, then so are
    a) the **disjunction** $(\varphi \vee \psi)$,
    b) the **conjunction** $(\varphi \wedge \psi)$,

    c) the **implication** $(\varphi \Rightarrow \psi)$, and
    d) the **equivalence** $(\varphi \Leftrightarrow \psi)$.

3. If φ is a formula and x is variable, then
    a) the **instantiation** $\exists x\ \varphi$ and
    b) the **generalization** $\forall x\ \varphi$

   are both formulas.

The expressions $\exists x$ and $\forall x$ are called **quantifiers.** The negation of the formula $t \in u$ is usually written as

$$t \notin u$$

rather than $\neg\, t \in u$; it says t is *not* a member of u. The expression

$$\forall z\ (z \in x \Rightarrow z \in y)$$

is the formula saying that every element of x is an element of y. Another way to say this is that x is a **subset** of y, or x is **included** in y, or y **includes** x. We abbreviate the formula by[3]

$$x \subseteq y.$$

Then the expression

$$(x \subseteq y \wedge y \subseteq x)$$

stands for the formula saying that x and y have the same members, so that they are **equal** by the definition foretold above (page 16); in this case we use the abbreviation

$$x = y.$$

The negation of this is usually written as

$$x \neq y.$$

Another abbreviation that we use is to eliminate the outer parentheses from a formula (when they are present) and to eliminate internal parentheses when they can be resupplied according to the following rules:

[3] The relation $\subseteq$ of *being included in* is completely different from the relation $\in$ of *being contained in.* However, many mathematicians confuse these relations in words, using "contained" to describe both.

1. The binary connectives $\land$ and $\lor$ have priority over $\Rightarrow$ and $\Leftrightarrow$, so that, for example, $\varphi \land \psi \Rightarrow \chi$ means $(\varphi \land \psi) \Rightarrow \chi$.
2. When two connectives $\Rightarrow$ appear without an intervening parenthesis, the arrow on the right has priority, so $\varphi \Rightarrow \psi \Rightarrow \chi$ means $\varphi \Rightarrow (\psi \Rightarrow \chi)$.

2.2.2. Truth and falsity

The same variable may have several **occurrences** in a particular formula. All occurrences of the variable x in the formulas $\exists x\ \varphi$ and $\forall x\ \varphi$ are said to be **bound,**[4] and they remain bound when other formulas are built up from these formulas. Occurrences of a variable that are not bound are **free.** The same variable can have both bound and free occurrences in the same formula, although this can always be avoided. For example, in the formula $x \in y \Rightarrow \forall y\ x \in y$, the first occurrence of y is free, but the other two occurrences are bound; nonetheless, the formula will have the same meaning as $x \in y \Rightarrow \forall z\ x \in z$, in which the only occurrence of y is free.

If a variable has free occurrences in a formula, then the variable is said to be a **free variable** of the formula, even though the variable might also have bound occurrences in the formula. A **sentence** is a formula like $\forall x\ \exists y\ x \in y$ or $\forall x\ x \notin a$, with no free variables. A **singulary**[5] formula is a formula with only one free variable. If φ is a singulary formula, and its free variable is x, then we may write φ as

$$\varphi(x).$$

[4] The word “bound” here is the past participle of the verb “to bind,” meaning *tie up or restrain.* There is another verb, “to bound,” meaning *put a bound or limit on*: this is also used in mathematics, but its past participle is “bounded.” Although they have similar meanings, the two verbs “to bind” and “to bound” have different origins. The verb “to bind” has been part of English for as long as that language is recognized to have existed: since the eighth century. That is, the precursor of “to bind” is found in Old English. The verb “to bound” is based on the noun “bound,” which entered Middle English in the 12th century from the Old French noun that became the modern French *borne.*

[5] In place of “singulary,” the word **unary** is more common, but less etymologically correct.

By replacing every free occurrence of x in φ with a constant a, we obtain the formula

$$\varphi(a),$$

which is a sentence.

An arbitrary sentence has a **truth-value,** which is either **true** or **false,** but not both. However, the truth-value of a sentence in which constants occur may depend on which sets are named by those constants. Like the definition of formulas in the first place, the definition of the truth-value of sentences is *recursive,* as follows.

1. The atomic sentence $a \in b$ is true if and only if the set a is an element of the set b.
2. If σ and τ are sentences, and $*$ is a binary Boolean connective, then the truth-value of the sentence $(\sigma * \tau)$ depends on the truth-value of σ and τ according to the usual rules of propositional logic:
    a) $(\sigma \vee \tau)$ is true in $\mathfrak{A}$ if and only if at least one of σ and τ is true in $\mathfrak{A}$.
    b) $(\sigma \wedge \tau)$ is true in $\mathfrak{A}$ if and only if both σ and τ are true in $\mathfrak{A}$.
    c) $(\sigma \Rightarrow \tau)$ is true in $\mathfrak{A}$ if and only if $(\neg\sigma \vee \tau)$ is true in $\mathfrak{A}$.
    d) $(\sigma \Leftrightarrow \tau)$ is true in $\mathfrak{A}$ if and only if both $(\sigma \Rightarrow \tau)$ and $(\tau \Rightarrow \sigma)$ are true in $\mathfrak{A}$.
3. Suppose $\varphi(x)$ is a singulary formula.
    a) The instantiation $\exists x\ \varphi(x)$ is true if and only if $\varphi(a)$ is true for *some* set a.
    b) The generalization $\forall x\ \varphi(x)$ is true if and only if $\varphi(a)$ is true for *all* sets a.

The validity of this definition relies on:

Theorem 1 (Unique Readability)**.** *A given formula can be built up from atomic formulas in only one way.*

This means two things:

1. Each formula is of exactly one of the eight kinds named in the previous subsection: (i) an atomic formula, (ii) a negation, (iii) a disjunction, (iv) a conjunction, (v) an implication, (vi) an equivalence, (vii) an instantiation, or (viii) a generalization.

2. Each formula is of one of these kinds in only one way.

These two conclusions are obvious for atomic formulas, negations, generalizations, and instantiations. For disjunctions, conjunctions, implications, and equivalences, the theorem is a consequence of the following.

Lemma 1. *No proper initial segment of a formula is a formula.*

Proof. We prove by induction that every formula neither *is* a proper initial segment of another formula, nor *has* a proper initial segment that is a formula. This is obviously true for atomic formulas. Suppose this is true for the formulas φ and ψ. Then it is obviously true for the the three formulas that can be obtained in one step from φ, as well as for the four formulas that can be obtained in one step from φ and ψ. Therefore the claim is true for all formulas. □

Another difficulty with the definition of truth and falsity is as follows. The definition assigns truth-values, not to arbitrary formulas, but to sentences only. However, sentences as such are not defined recursively. Strictly, the recursive definition of truth-value determines, for each formula φ, an assignment of a truth-value to each sentence that results from φ by replacing each free occurrence of a variable with a constant.

Note that $\forall x\ \varphi(x)$ and "For all a, $\varphi(a)$" are two ways of saying the same thing. The former expression is a sentence of our logic; the latter expression is a sentence of English that incorporates the constant a and the sentence $\varphi(a)$ of our logic. In particular, in English, the constant a plays the role of a variable. In place of "For all a, $\varphi(a)$," we may say simply $\varphi(a)$, if it is clear that a is an *arbitrary* set.

2.2.3. Logical truth

The truth-value of a sentence is determined by the truth-values of all atomic sentences. However, some sentences are true, regardless of the truth-values of atomic sentences. Such sentences are **logically true.**

For example, the sentences

$$(\sigma \Rightarrow \tau) \Leftrightarrow \neg\sigma \lor \tau, \qquad\qquad \forall x\ \varphi(x) \Leftrightarrow \neg\exists x\ \neg\varphi(x)$$

are logically true. A *formula* is logically true if every generation of it that is a sentence is logically true. Then two formulas φ and ψ are **logically equivalent** to one another if the equivalence $\varphi \Leftrightarrow \psi$ is logically true. For example, $\varphi \Rightarrow \psi$ and $\neg\varphi \lor \psi$ are logically equivalent to one another. So are the formulas

$$\psi \Rightarrow \forall x\ \varphi(x), \qquad\qquad \forall x\ \bigl(\psi \Rightarrow \varphi(x)\bigr);$$

and so are the formulas

$$\forall x\ \varphi(x) \Rightarrow \psi, \qquad\qquad \exists x\ \bigl(\varphi(x) \Rightarrow \psi\bigr).$$

We shall use these logical equivalences in examining equality below.

2.2.4. Classes and equality

If φ is a singulary formula $\varphi(x)$, and the sentence $\varphi(a)$ is true, then a can be said to **satisfy** φ. There is a collection of all sets that satisfy φ, and we denote this collection by

$$\{x\colon \varphi(x)\}.$$

Such a collection is called a **class.** In particular, it is the class **defined** by the formula φ. We may give this class a name like $\boldsymbol{C}$, written in boldface: in this case the expression

$$x \in \boldsymbol{C}$$

means just $\varphi(x)$.

A formula in which only two variables occur freely is **binary.** If ψ is such a formula, with free variables x and y, then we may write ψ as

$$\psi(x, y).$$

We shall want this notation for proving Theorem 2 below. If needed, we can talk about ternary formulas $\chi(x, y, z)$, and so on.

By definition of equality, the sentences

$$\begin{aligned}&\forall x\ \forall y\ \forall z\ \bigl(x = y \Rightarrow (z \in x \Leftrightarrow z \in y)\bigr),\\&\forall x\ \forall y\ \exists z\ \bigl((z \in x \Leftrightarrow z \in y) \Rightarrow x = y\bigr).\end{aligned} \tag{2.1}$$

are logically true. We can write the former as

$$\forall x\ \forall y\ \bigl(x = y \Rightarrow (a \in x \Leftrightarrow a \in y)\bigr). \tag{2.2}$$

Axiom 1 (Equality). *Equal sets belong to the same sets:*

$$\forall x\ \forall y\ \bigl(x = y \Rightarrow (x \in a \Leftrightarrow y \in a)\bigr). \tag{2.3}$$

Theorem 2. *Equal sets satisfy the same formulas:*

$$\forall x\ \forall y\ \Bigl(x = y \Rightarrow \bigl(\varphi(x) \Leftrightarrow \varphi(y)\bigr)\Bigr). \tag{2.4}$$

Proof. Suppose $a = b$. By symmetry, it is enough to show

$$\varphi(a) \Rightarrow \varphi(b) \tag{2.5}$$

for all singulary formulas $\varphi(x)$. We use **induction;** this is possible because formulas are defined recursively. See §2.4 below (page 35).

By (2.2) and (2.3), (2.5) holds when $\varphi(x)$ is an atomic formula $x \in c$ or $c \in x$. There is another form of singulary atomic formula, namely $x \in x$. If $a \in a$, then a satisfies $x \in a$, and therefore so does b; thus $b \in a$, so a satisfies $b \in x$, and therefore so does b. Thus $a \in a \Rightarrow b \in b$. So we have (2.5) when φ is any singulary atomic formula.

If we have (2.5) when φ is ψ, then we have it when φ is $\neg\psi$. If we have (2.5) when φ is ψ or χ, then we have it when φ is $(\psi * \chi)$, where $*$ is one of the binary connectives. If, for some binary formula $\psi(x, y)$, we have (2.5) whenever $\varphi(x)$ is $\psi(x, c)$ for some set c, then we have (2.5) when $\varphi(x)$ is $\forall y\ \psi(x, y)$ or $\exists y\ \psi(x, y)$. Therefore we do have (2.5) in all cases. □

For many writers, equality is a logical concept, and the sentence (2.4) is taken as logically true. Then (2.2) and (2.3) are special cases of this, but (2.1) is not logically true. In this case, (2.1) must also be taken as an axiom, which is called the **Extension Axiom.** No matter which approach one takes, all of the sentences (2.1), (2.2), (2.3), and (2.4) end up being true. They tell us that equal sets are precisely those sets that are logically indistinguishable.

As with sets, so with classes, one of them **includes** another if every element of the latter belongs to the former. Hence if formulas $\varphi(x)$ and $\psi(y)$ define classes $\boldsymbol{C}$ and $\boldsymbol{D}$ respectively, and if

$$\forall x\ \bigl(\varphi(x) \Rightarrow \psi(x)\bigr),$$

this means $\boldsymbol{D}$ includes $\boldsymbol{C}$, and we write

$$\boldsymbol{C} \subseteq \boldsymbol{D}.$$

If also $\boldsymbol{C}$ includes $\boldsymbol{D}$, then the two classes are **equal,** and we write

$$\boldsymbol{C} = \boldsymbol{D};$$

this means $\forall x\ \bigl(\varphi(x) \Leftrightarrow \psi(x)\bigr)$. Likewise set and a class can be considered as **equal** if they have the same members. Thus if again $\boldsymbol{C}$ is defined by $\varphi(x)$, then the expression

$$a = \boldsymbol{C}$$

means $\forall x\ \bigl(x \in a \Leftrightarrow \varphi(x)\bigr)$.

Theorem 3. *Every set is equal to a class.*

Proof. $a = \{x \colon x \in a\}$. □

However, there is no reason to expect the converse to be true.

Theorem 4. *Not every class is equal to a set.*

Proof. There are formulas $\varphi(x)$ such that

$$\forall y\ \neg\forall x\ \bigl(x \in y \Leftrightarrow \varphi(x)\bigr); \tag{2.6}$$

for example, $\varphi(x)$ could be $x \notin x$, so that $\forall y\ \neg\bigl(y \in y \Leftrightarrow \varphi(y)\bigr)$. In any case, if (2.6) holds, then no set can be equal to the class $\{x\colon \varphi(x)\}$. □

More informally, the argument is that the class $\{x\colon x \notin x\}$ is not a set, because if it were a set a, then $a \in a \Leftrightarrow a \notin a$, which is a contradiction. This is what was given above as the Russell Paradox (page 15). Another example of a class that is not a set is given by the *Burali-Forti Paradox* on page 59 below.

2.2.5. Construction of sets

We have established what it means for sets to be equal. We have established that sets are examples, but not the only examples, of the collections called classes. However, we have not officially exhibited any sets. We do this now.

Axiom 2 (Empty Set). *The empty class is a set:*

$$\exists x\ \forall y\ y \notin x.$$

As noted above (page 16), the set whose existence is asserted by this axiom is denoted by $\varnothing$. This set is the class $\{x\colon x \neq x\}$.

We now obtain the sequence 0, 1, 2, ..., described above (page 16). We use the Empty Set Axiom to start the sequence. We continue by means of:

Axiom 3 (Adjunction). *If a and b are sets, then there is a set denoted by $a \cup \{b\}$:*

$$\forall x\ \forall y\ \exists z\ \forall w\ (w \in z \Leftrightarrow w \in x \lor w = y).$$

In writing the axiom formally, we have followed the abbreviative conventions on page 18. We can understand the Adjunction Axiom as saying that, for all sets a and b, the class $\{x\colon x \in a \lor x = b\}$ is actually a set. Adjunction is not one of Zermelo's original axioms of 1908; but the following is Zermelo's **Pairing Axiom:**

Theorem 5. *For any two sets a and b, the set $\{a, b\}$ exists:*

$$\forall x\ \forall y\ \exists z\ \forall w\ (w \in z \Leftrightarrow w = x \lor w = y).$$

Proof. By Empty Set and Adjunction, $\varnothing \cup \{a\}$ exists, but this is just $\{a\}$. Then $\{a\} \cup \{b\}$ exists by Adjunction again. □

The theorem is that the class $\{x\colon x = a \lor x = b\}$ is always a set. Actually Zermelo does not have a Pairing Axiom as such, but he has an **Elementary Sets Axiom,** which consists of what we have called the Empty Set Axiom and the Pairing Axiom.[6]

Every class $\boldsymbol{C}$ has a **union,** which is the class

$$\{x\colon \exists y\ (x \in y \land y \in \boldsymbol{C})\}.$$

This class is denoted by

$$\bigcup \boldsymbol{C}.$$

This notation is related as follows with the notation for the classes involved in the Adjunction Axiom:

Theorem 6. *For all sets a and b, $a \cup \{b\} = \bigcup \bigl\{a, \{b\}\bigr\}$.*

We can now use the more general notation

$$a \cup b = \bigcup \{a, b\}.$$

Axiom 4 (Union)**.** *The union of a* set *is always a set:*

$$\forall x\ \exists y\ y = \bigcup x.$$

The Adjunction Axiom is a consequence of the Empty-Set, Pairing, and Union Axioms. This why Zermelo did not need Adjunction as an axiom. We state it as an axiom, because we can do a lot of mathematics with it that does not require the full force of the Union Axiom.

[6] Zermelo also requires that for every set a there be a set $\{a\}$; but this can be understood as a special case of pairing.

Suppose A is a set and $\boldsymbol{C}$ is the class $\{x\colon \varphi(x)\}$. Then we can form the class

$$A \cap \boldsymbol{C},$$

which is defined by the formula $x \in A \wedge \varphi(x)$. Standard notation for this class is[7]

$$\{x \in A\colon \varphi(x)\}. \tag{2.7}$$

Axiom 5 (Separation). *Every class $\{x \in A\colon \varphi(x)\}$ is a set.*

The Separation Axiom is really a *scheme* of axioms, one for each singulary formula φ:

$$\forall x\ \exists y\ \forall z\ \bigl(z \in y \Leftrightarrow z \in x \wedge \varphi(z)\bigr).$$

In most of mathematics, and in particular in the other sections of this text, one need not worry too much about the distinction between sets and classes. But it is logically important. It turns out that the objects of interest in mathematics can be understood as sets. Indeed, we have already defined natural numbers as sets. We can talk about sets by means of formulas. Formulas define classes of sets, as we have said. Some of these classes turn out to be sets themselves; but again, there is no reason to expect all of them to be sets, and indeed by Theorem 4 (page 24) some of them are not sets. *Sub-classes* of sets are sets, by the Separation Axiom; but some classes are too big to be sets. The class $\{x\colon x = x\}$ of all sets is not a set, since if it were, then the sub-class $\{x\colon x \notin x\}$ would be a set, and it is not.

[7] This notation is unfortunate. Normally the formula $x \in A$ is read as a sentence of ordinary language, namely "x belongs to A" or "x is in A." However, the expression in (2.7) is read as "the set of x in A such that φ holds of x"; in particular, $x \in A$ here is read as the noun phrase "x in A" (or "x belonging to A," or "x that are in A," or something like that). Thus a more precise way to write the expression in (2.7) would be something like $\{x \text{ in } A\colon \varphi(x)\}$. Ambiguity of expressions like $x \in A$ (is it a noun or a sentence?) is common in mathematical writing, as for example in the abbreviation of $\forall\varepsilon\ (\varepsilon > 0 \Rightarrow \varphi)$ as $(\forall\varepsilon > 0)\ \varphi$. Nonetheless, such ambiguity is avoided in this text.

Every set a has a *power class,* namely the class $\{x\colon x \subseteq a\}$ of all subsets of a. This class is denoted by

$$\mathscr{P}(a).$$

Axiom 6 (Power Set)**.** *Every power class is a set:*

$$\forall x\ \exists y\ y = \mathscr{P}(x).$$

Then $\mathscr{P}(a)$ can be called the **power set** of a. The Power Set Axiom is of fundamental importance for allowing us to prove Theorem 10 on page 31 below.

We want the collection $\{0, 1, 2, \dots\}$ of natural numbers as defined on page 16 to be a set. Now, it is not obvious how to formulate this as a sentence of our logic. However, the indicated collection contains 0, which by definition is the empty set; also, for each of its elements n, the collection contains also $n \cup \{n\}$. Let $\mathbf{I}$ be the class of all *sets* with these properties: thus

$$\mathbf{I} = \bigl\{x\colon 0 \in x \land \forall y\ (y \in x \Rightarrow y \cup \{y\} \in x)\bigr\}.$$

If it exists, the set of natural numbers will belong to $\mathbf{I}$. Furthermore, the set of natural numbers will be the *smallest* element of $\mathbf{I}$. But we still must make this precise. For an arbitrary class $\boldsymbol{C}$, we define

$$\bigcap \boldsymbol{C} = \{x\colon \forall y\ (y \in \boldsymbol{C} \Rightarrow x \in y)\}.$$

This class is the **intersection** of $\boldsymbol{C}$.

Theorem 7. *If a and b are two sets, then*

$$a \cap b = \bigcap \{a, b\}.$$

If $a \in \boldsymbol{C}$, then

$$\bigcap \boldsymbol{C} \subseteq a,$$

so in particular $\bigcap \boldsymbol{C}$ is a set. However, $\bigcap \varnothing$ is the class of all sets, which is not a set.[8]

[8] Some writers define $\bigcap \boldsymbol{C}$ only when $\boldsymbol{C}$ is a nonempty set.

Axiom 7 (Infinity). $\mathbf{I} \neq \varnothing$*:*

$$\exists x\ \big(0 \in x \land \forall y\ (y \in x \Rightarrow y \cup \{y\} \in x)\big).$$

We can now define[9]

$$\omega = \bigcap \mathbf{I}, \tag{2.8}$$

knowing that this is a set.[10]

Theorem 8. $\omega \in \mathbf{I}$.

We shall establish the additional properties of ω in §2.5 (p. 43).

2.2.6. The Zermelo–Fraenkel Axioms with Choice

We state the following for the record; but we are not going to use it freely, as we shall use the preceding axioms.

Axiom 8 (Choice). *For every set A of nonempty sets, any two of which are disjoint from one another, there is a set b such that, for each set c in A, the intersection $b \cap c$ has a unique element.*

We have now named all of the axioms given by Zermelo in 1908:

(I) Extension,
(II) Elementary Sets,
(III) Separation,
(IV) Power Set,
(V) Union,
(VI) Choice, and
(VII) Infinity.

[9] See note 1 on page 12 about the letter ω.

[10] Every other axiom of this section is of the form, "Such-and-such classes are sets." We can express the Axiom of Infinity in this form, as "$\bigcap \mathbf{I}$ is a set." However, it would be preferable to define ω as a class, without using the Axiom of Infinity; then this Axiom could be simply, "ω is a set." We can do this. In the terminology of §2.9 (page 58), we can define ω as the class of all ordinals that neither *contain* limits nor *are* limits themselves.

Zermelo assumes that equality is identity: but his assumption is our Theorem 2. In fact Zermelo does not use logical formalism as we have. We prefer to define equality with (2.1) and (2.2) and then use the Axioms of

(i) Equality,
(ii) the Empty Set,
(iii) Adjunction,
(iv) Union,
(v) Separation,
(vi) Power Set,
(vii) Infinity, and
(viii) Choice.

But these two collections of definitions and axioms are logically equivalent: using either collection, we can prove the axioms in the other collection as theorems.

Apparently Zermelo overlooked an axiom, the **Replacement Axiom,** which was supplied in 1922 by Skolem [55] and by Fraenkel.[11] We shall give this axiom on page 34 in the next section.

An axiom never needed in ordinary mathematics is the **Foundation Axiom.** Stated originally by von Neumann [59], it ensures that certain pathological situations, like a set's containing itself, are impossible. It does this by declaring that every nonempty set has an element that is disjoint from it:

$$\forall x\, \exists y\, (x \neq \varnothing \Rightarrow y \in x \land x \cap y = \varnothing).$$

We shall never use this axiom.

Zermelo's axioms, along with Replacement and Foundation, compose the collection called

$$\text{ZFC}.$$

[11] I have not been able to consult Fraenkel's original papers. However, according to van Heijenoort [58, p. 291], Lennes also suggested something like the Replacement Axiom at around the same time (1922) as Skolem and Fraenkel; but Cantor had suggested such an axiom in 1899.

If we leave out Choice, we have what is called

$$\mathrm{ZF}.$$

We shall tacitly assume ZF *throughout this text.* When we want to use the Axiom of Choice, we shall be explicit about it.

2.3. Functions and relations

2.3.1. Cartesian products

Given two sets a and b, we define

$$(a, b) = \big\{\{a\}, \{a, b\}\big\}.$$

This set is the **ordered pair** whose first entry is a and whose second entry is b. The purpose of the definition is to make the following theorem true.

Theorem 9. *Two ordered pairs are equal if and only if their first entries are equal and their second entries are equal:*

$$(a, b) = (x, y) \Leftrightarrow a = x \wedge b = y.$$

If A and B are sets, then we define

$$A \times B = \big\{z \colon \exists x \; \exists y \; \big(z = (x, y) \wedge x \in A \wedge y \in B\big)\big\}.$$

This is the **cartesian product** of A and B.

Theorem 10. *The cartesian product of two sets is a set.*

Proof. If $a \in A$ and $b \in B$, then $\{a\}$ and $\{a, b\}$ are elements of $\mathscr{P}(A \cup B)$, so $(a, b) \in \mathscr{P}(\mathscr{P}(A \cup B))$, and therefore

$$A \times B \subseteq \mathscr{P}(\mathscr{P}(A \cup B)). \qquad \square$$

An **ordered triple** (x, y, z) can be defined as $\big((x, y), z\big)$, and so forth.

2.3.2. Functions

A **function** or **map** from A to B is a subset f of $A \times B$ such that, for each a in A, there is exactly one b in B such that $(a, b) \in f$. Then instead of $(a, b) \in f$, we write

$$f(a) = b. \tag{2.9}$$

We have then

$$A = \{x\colon \exists y\ f(x) = y\},$$

that is, $A = \{x\colon \exists y\ (x, y) \in f\}$. The set A is called the **domain** of f. A function is sometimes said to be a function **on** its domain. For example, the function f here is a function on A. The **range** of f is the subset

$$\{y\colon \exists x\ f(x) = y\}$$

of B. If this range is actually equal to B, then we say that f is **surjective onto** B, or simply that f is **onto** B. Strictly speaking, it would not make sense to say f was surjective or onto, simply.

A function f is **injective** or **one-to-one** if

$$\forall x\, \forall z\ (f(x) = f(z) \Rightarrow x = z).$$

The expression $f(x) = f(z)$ is an abbreviation of $\exists y\ (f(x) = y \wedge f(z) = y)$, which is another way of writing $\exists y\ \big((x, y) \in f \wedge (z, y) \in f\big)$. An injective function from A *onto* B is a **bijection** from A to B.

If it is not convenient to name a function with a single letter like f, we may write the function as

$$x \mapsto f(x),$$

where the expression $f(x)$ would be replaced by some particular expression involving x. As an abbreviation of the statement that f is a function from A to B, we may write

$$f\colon A \to B. \tag{2.10}$$

Thus, while the symbol f can be understood as a *noun,* the expression $f\colon A \to B$ is a complete *sentence.* If we say, "Let $f\colon A \to B$," we mean let f be a function from A to B.

If $f\colon A \to B$ and $D \subseteq A$, then the subset $\{y\colon \exists x\ (x \in D \wedge y = f(x)\}$ of B can be written as one of[12]

$$\{f(x)\colon x \in D\}, \qquad\qquad f[D].$$

This set is the **image** of D under f. Similarly, for the Cartesian product $A \times B$, instead of $\bigl\{z\colon \exists x\ \exists y\ \bigl(z = (x,y) \wedge x \in A \wedge y \in B\bigr)\bigr\}$, we can write

$$\{(x,y)\colon x \in A \wedge y \in B\}.$$

Variations on this notation are possible. For example, if $f\colon A \to B$ and $D \subseteq A$, then the **restriction** of f to D is the set

$$\{(x,y) \in f\colon x \in D\},$$

which we may denote by

$$f \restriction D.$$

Then the following is just an exercise in notation.

Theorem 11. *If $f\colon A \to B$ and $D \subseteq A$, then*

$$f \restriction D\colon D \to B$$

and, for all x in D, $(f \restriction D)(x) = f(x)$.

If $f\colon A \to B$ and $g\colon B \to C$, then we can define

$$g \circ f = \{(x,z)\colon \exists y\ (f(x) = y \wedge g(y) = z)\};$$

this is called the **composite** of (g,f).

[12] The notation $f(D)$ is also used, but the ambiguity is dangerous, at least in set theory as such.

Theorem 12. *If $f\colon A \to B$ and $g\colon B \to C$, then*

$$g \circ f\colon A \to C.$$

If also $h\colon C \to D$, then

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

We define

$$\mathrm{id}_A = \{(x,x)\colon x \in A\};$$

this is the **identity** on A.

Theorem 13. id_A *is a bijection from A to itself. If $f\colon A \to B$, then*

$$f \circ \mathrm{id}_A = f, \qquad \mathrm{id}_B \circ f = f.$$

If f is a bijection from A to B, we define

$$f^{-1} = \{(y,x)\colon f(x) = y\};$$

this is the **inverse** of f.

Theorem 14.

1. *The inverse of a bijection from A to B is a bijection from B to A.*
2. *Suppose $f : A \to B$ and $g : B \to A$. Then f is a bijection from A to B whose inverse is g if and only if*

$$g \circ f = \mathrm{id}_A, \qquad f \circ g = \mathrm{id}_B\,.$$

In the definition of the cartesian product $A \times B$ and of functions from A to B, we may replace the sets A and B with classes. For example, we may speak of the function $x \mapsto \{x\}$ on the class of all sets.

Axiom 9 (Replacement). *If $\boldsymbol{F}$ is a function on some class $\boldsymbol{C}$, and A is a* subset *of $\boldsymbol{C}$, then the image $\boldsymbol{F}[A]$ is also a set.*

For example, if we are given a function $n \mapsto G_n$ on ω, then by Replacement the class $\{G_n \colon n \in \omega\}$ is a set. Then the union of this class is a set, which we denote by

$$\bigcup_{n\in\omega} G_n.$$

A **singulary operation** on A is a function from A to itself; a **binary** on A is a function from $A \times A$ to A.

2.3.3. Relations

A **binary relation** on A is a subset of $A \times A$; if R is such, and $(a, b) \in R$, we often write

$$a \mathrel{R} b.$$

A singulary operation on A is a particular kind of binary relation on A; for such a relation, we already have the special notation in (2.9). The reader will be familiar with other kinds of binary relations, such as *equivalence relations* and *orderings.* Equality of sets is an equivalence relation; see also pages 49 and 76. We are going to define a particular binary relation on page 41 below and prove that it is a linear ordering.

Meanwhile, if $R \subseteq A \times B$, then R is a binary relation on $A \cup B$; but we may say more precisely that R is a relation **from** A **to** B, or a relation **between** A and B (in that order). We consider this situation in the proof of Theorem 15 (page 37), and then again in §4.3 (page 105). The **domain** of a relation R from A to B is the subset $\{x \in A \colon \exists y\ (x \mathrel{R} y)\}$ of A.

2.4. An axiomatic development of the natural numbers

In the preceding sections, we sketched an axiomatic approach to set theory. Now we start over with an axiomatic approach to the natural numbers alone. In the section after this, we shall show that the set ω does actually provide a *model* of the axioms for natural numbers developed in the present section.

For the moment though, we forget the definition of ω. We forget about starting the natural numbers with 0. Children learn to count starting with 1, not 0. Let us understand the natural numbers to compose *some* set called $\mathbb{N}$. This set has a distinguished **initial element,** which we call **one** and denote by

$$1.$$

On the set $\mathbb{N}$ there is also a distinguished singulary operation of **succession,** namely the operation

$$n \mapsto n+1,$$

where $n+1$ is called the **successor** of n. Note that some other expression like $S(n)$ might be used for this successor. For the moment, we have no binary operation called $+$ on $\mathbb{N}$.

I propose to refer to the ordered triple $(\mathbb{N}, 1, n \mapsto n+1)$ as an *iterative structure.* In general, by an **iterative structure,** I mean any set that has a distinuished element and a distinguished singulary operation. Here the underlying set can be called the **universe** of the structure.[13] The iterative structure $(\mathbb{N}, 1, n \mapsto n+1)$ is distinguished among all iterative structures by satisfying the following axioms.

I. 1 is not a successor: $1 \neq n+1$.

II. Succession is injective: if $m+1 = n+1$, then $m = n$.

III. The structure admits **proof by induction,** in the following sense. Every subset A of the universe must be the whole universe, provided A has the following two closure properties.

   A. $1 \in A$.

   B. For all n, if $n \in A$, then $n+1 \in A$.

[13] For a simple notational distinction between a structure and its universe, if the universe is A, the structure itself can be denoted by a fancier version of this letter, such as the Fraktur version $\mathfrak{A}$. See Appendix A (p. 255) for Fraktur versions, and their handwritten forms, for all of the Latin letters. However, we shall not make use of Fraktur letters until defining structures in general on page 46.

These axioms were published first by Dedekind [13, II, VI (71), p. 67]; but they were written down also by Peano [48], and they are often known as the **Peano axioms.**

Suppose (A, b, f) is an iterative structure. If we successively compute b, $f(b)$, $f(f(b))$, $f(f(f(b)))$, and so on, either we always get a new element of A, or we reach an element that we have already seen. In the latter case, if the first repeated element is b, then the first Peano axiom fails. If it is not b, then the second Peano axiom fails. The last Peano axiom, the Induction Axiom, would ensure that every element of A was reached by our computations. None of the three axioms implies the others, although the Induction Axiom implies that exactly one of the other two axioms holds [29].

2.4.1. Recursion

The following theorem will allow us to define all of the usual operations on $\mathbb{N}$. The theorem is difficult to prove. Not the least difficulty is seeing that the theorem *needs* to be proved.[14]

Homomorphisms will be defined generally on page 47, but meanwhile we need a special case. A **homomorphism** from the iterative structure $(\mathbb{N}, 1, n \mapsto n + 1)$ to an arbitrary iterative structure (A, b, f) is a function h from $\mathbb{N}$ to A such that

1) $h(1) = b$, and
2) $h(n + 1) = f(h(n))$ for all n in $\mathbb{N}$,

that is, the diagram in Figure 2.1 **commutes** (any two paths from one point to another represent the same function).

Theorem 15 (Recursion). *For every iterative structure, there is exactly one homomorphism from* $(\mathbb{N}, 1, n \mapsto n + 1)$ *to this structure.*

Proof. Given an iterative structure (A, b, f), we seek a homomorphism h from $(\mathbb{N}, 1, x \mapsto n+1)$ to (A, b, f). Then h will be a particular subset of $\mathbb{N} \times A$. Let $\mathscr{B}$ be the set whose elements are the subsets C of $\mathbb{N} \times A$ such that, if $(n, y) \in C$, then either

[14] Peano did not see this need, but Dedekind did. Landau discusses the matter [40, pp. ix–x].

$$\begin{tikzcd}
\{1\} \arrow[r, "\mathrm{id}_{\{1\}}"] \arrow[d, "h\restriction\{1\}"'] & \mathbb{N} \arrow[r, "n\mapsto n+1"] \arrow[d, "h"'] & \mathbb{N} \arrow[d, "h"] \\
\{b\} \arrow[r, "\mathrm{id}_{\{b\}}"'] & A \arrow[r, "f"'] & A
\end{tikzcd}$$

Figure 2.1.: A homomorphism of iterative structures

1) $(n, y) = (1, b)$ or else

2) C has an element (m, x) such that $(n, y) = (m + 1, f(x))$.

In particular, $\{(1, b)\} \in \mathscr{B}$. Also, if $C \in \mathscr{B}$ and $(m, x) \in C$, then

$$C \cup \{(m + 1, f(x))\} \in \mathscr{B}.$$

Let $R = \bigcup \mathscr{B}$; so R is a subset of $\mathbb{N} \times A$, that is, a relation from $\mathbb{N}$ to A in the sense of page 35. If $(n, y) \in R$, then (on page 35) we may write also

$$n\ R\ y.$$

Since $\{(1, b)\} \in \mathscr{B}$, we have $1\ R\ b$. Also, if $m\ R\ x$, then $(m, x) \in C$ for some C in $\mathscr{B}$, so $C \cup \{(m+1, f(x))\} \in \mathscr{B}$, and therefore $(m+1)\ R\ f(x)$. Thus R is the desired function h, provided R is actually a *function* from $\mathbb{N}$ to A. Proving that R is a function from $\mathbb{N}$ to R has two stages.

1. Let D be the set of all n in $\mathbb{N}$ for which there is y in A such that $n\ R\ y$. Then we have just seen that $1 \in D$, and if $n \in D$, then $n + 1 \in D$. By induction, $D = \mathbb{N}$. Thus if R is a function, its domain is $\mathbb{N}$.

2. Let E be the set of all n in $\mathbb{N}$ such that, for all y in A, if $n\ R\ y$ and $n\ R\ z$, then $y = z$. Suppose $1\ R\ y$. Then $(1, y) \in C$ for some C in $\mathscr{B}$. Since 1 is not a successor, we must have $y = b$, by definition of $\mathscr{B}$. Therefore $1 \in E$. Suppose $n \in E$, and $(n+1)\ R\ y$. Then $(n+1, y) \in C$ for some C in $\mathscr{B}$. Again since 1 is not a successor, we must have

$$(n + 1, y) = (m + 1, f(x))$$

for some (m, x) in C. Since succession is injective, we must have $m = n$. Thus, $y = f(x)$ for some x in A such that $n \mathrel{R} x$. Since $n \in E$, we know x is *unique* such that $n \mathrel{R} x$. Therefore y is unique such that $(n+1) \mathrel{R} y$. Thus $n + 1 \in E$. By induction, $E = \mathbb{N}$.

So R is the desired function h. Finally, h is unique by induction. □

Note well that the proof uses all three of the Peano Axioms. The Recursion Theorem is often used in the following form.

Corollary 15.1. *For every set A with a distinguished element b, and for every function F from $\mathbb{N} \times B$ to B, there is a unique function H from $\mathbb{N}$ to A such that*

1) $H(1) = b$, and

2) $H(n+1) = F(n, H(n))$ for all n in $\mathbb{N}$.

Proof. Let h be the unique homomorphism from $(\mathbb{N}, 1, n \mapsto n+1)$ to $(\mathbb{N} \times A, (1, b), f)$, where f is the operation $(n, x) \mapsto (n+1, F(n, x)))$. In particular, $h(n)$ is always an ordered pair. By induction, the first entry of $h(n)$ is always n; so there is a function H from $\mathbb{N}$ to A such that $h(n) = (n, H(n))$. Then H is as desired. By induction, H is unique. □

2.4.2. Arithmetic operations

We can now use recursion to define, on $\mathbb{N}$, the binary operation

$$(x, y) \mapsto x + y$$

of **addition,** and the binary operation

$$(x, y) \mapsto x \cdot y$$

of **multiplication.** More precisely, for each n in $\mathbb{N}$, we recursively define the operations $x \mapsto n + x$ and $x \mapsto n \cdot x$. The definitions are:

$$\begin{aligned} n + 1 &= n + 1, & n + (m+1) &= (n+m) + 1, \\ n \cdot 1 &= n, & n \cdot (m+1) &= n \cdot m + n. \end{aligned} \tag{2.11}$$

The definition of addition might also be written as $n + 1 = S(n)$ and $n + S(m) = S(n+m)$. In place of $x \cdot y$, we often write xy.

Lemma 2. *For all n and m in $\mathbb{N}$,*

$$1 + n = n + 1, \qquad (m + 1) + n = (m + n) + 1.$$

Proof. Induction. □

Theorem 16. *Addition on $\mathbb{N}$ is*

*1) **commutative:** $n + m = m + n$; and*
*2) **associative:** $n + (m + k) = (n + m) + k$.*

Proof. Induction and the lemma. □

Theorem 17. *Addition on $\mathbb{N}$ allows **cancellation:** if $n + x = n + y$, then $x = y$.*

Proof. Induction, and injectivity of succession. □

The analogous proposition for multiplication is Corollary 23.1 below.

Lemma 3. *For all n and m in $\mathbb{N}$,*

$$1 \cdot n = n, \qquad (m + 1) \cdot n = m \cdot n + n.$$

Proof. Induction. □

Theorem 18. *Multiplication on $\mathbb{N}$ is*

1) commutative: $nm = mn$;
*2) **distributive** over addition: $n(m + k) = nm + nk$; and*
3) associative: $n(mk) = (nm)k$.

Proof. Induction and the lemma. □

Landau [40] proves *using induction alone* that $+$ and $\cdot$ exist as given by the recursive definitions above. However, Theorem 17 needs more than induction. So does the existence of the **factorial** function defined by

$$1! = 1, \qquad (n + 1)! = n! \cdot (n + 1).$$

So does **exponentiation,** defined by

$$n^1 = n, \qquad n^{m+1} = n^m \cdot n.$$

2.4.3. The linear ordering

The usual ordering $<$ of $\mathbb{N}$ is defined recursively as follows. First note that $m \leqslant n$ means simply $m < n$ or $m = n$. Then the definition of $<$ is:

1) $m \nless 1$ (that is, $\neg\, m < 1$);
2) $m < n + 1$ if and only if $m \leqslant n$.

In particular, $n < n + 1$. Really, it is the sets $\{x \in \mathbb{N} \colon x < n\}$ that are defined by recursion:

$$\{x \in \mathbb{N} \colon x < 1\} = \varnothing,$$
$$\{x \in \mathbb{N} \colon x < n + 1\} = \{x \in \mathbb{N} \colon x < n\} \cup \{n\} = \{x \in \mathbb{N} \colon x \leqslant n\}.$$

We now have $<$ as a binary relation on $\mathbb{N}$; we must *prove* that it is an ordering.

Theorem 19. *The relation $<$ is* ***transitive*** *on $\mathbb{N}$, that is, if $k < m$ and $m < n$, then $k < n$.*

Proof. Induction on n. □

Theorem 20. *The relation $<$ is* ***irreflexive*** *on $\mathbb{N}$: $m \nless m$.*

Proof. Since every element k of $\mathbb{N}$ is less than some other element (namely $k + 1$), it is enough to prove

$$k < n \Rightarrow k \nless k.$$

We do this by induction on n. The claim is vacuously true when $n = 1$. Suppose it is true when $n = m$. If $k < m + 1$, then $k < m$ or $k = m$. If $k < m$, then by inductive hypothesis $k \nless k$. If $k = m$, but $k < k$, then $k < m$, so again $k \nless k$. Thus the claim holds when $n = m + 1$. By induction, it holds for all n. □

Because the relation $<$ is transitive and irreflexive on $\mathbb{N}$, the relation is called an **ordering**[15] of $\mathbb{N}$, and $\mathbb{N}$ is **ordered** by $<$.

[15] In some sources, what we are calling an *ordering* is called merely a *partial ordering*.

Lemma 4. $1 \leqslant m$.

Proof. Induction. □

Lemma 5. *If $k < m$, then $k + 1 \leqslant m$.*

Proof. The claim is vacuously true when $m = 1$. Suppose it is true when $m = n$. Say $k < n+1$. Then $k \leqslant n$. If $k = n$, then $k+1 = n+1$, so $k + 1 \leqslant n + 1$. If $k < n$, then $k + 1 \leqslant n$ by inductive hypothesis, so $k+1 < n+1$ by transitivity (Theorem 19), and therefore $k+1 \leqslant n+1$. Thus the claim holds when $m = n + 1$. By induction, the claim holds for all m. □

Theorem 21. *The relation $<$ is* **total** *on $\mathbb{N}$: either $k \leqslant m$ or $m < k$.*

Proof. By the last lemma but one, the claim is true when $k = 1$. Suppose it is true when $k = \ell$. If $m \not< \ell + 1$, then $m \not\leqslant \ell$. In this case, we have both $m \neq \ell$ and $m \not< \ell$. Also, by the inductive hypothesis, $\ell \leqslant m$, so $\ell < m$, and hence $\ell + 1 \leqslant m$ by the last lemma. □

Being a total ordering of $\mathbb{N}$, the relation $<$ is also called a **linear ordering** of $\mathbb{N}$, and $\mathbb{N}$ is **linearly ordered** by $<$.

Theorem 22. *For all m and n in $\mathbb{N}$, we have $m < n$ if and only if the equation*

$$m + x = n \tag{2.12}$$

is soluble in $\mathbb{N}$.

Proof. By induction on k, if $m + k = n$, then $m < n$. We prove the converse by induction on n. We never have $m < 1$. Suppose for some r that, for all m, if $m < r$, then the equation $m + x = r$ is soluble. Suppose also $m < r + 1$. Then $m < r$ or $m = r$. In the former case, by inductive hypothesis, the equation $m + x = r$ has a solution k, and therefore $m + (k + 1) = r + 1$. If $m = r$, then $m + 1 = r + 1$. Thus the equation $m + x = r + 1$ is soluble whenever $m < r + 1$. By induction, for all n in $\mathbb{N}$, if $m < n$, then (2.12) is soluble in $\mathbb{N}$. □

Theorem 23. *If $k < \ell$, then*

$$k + m < \ell + m, \qquad km < \ell m.$$

Here the first conclusion is a refinement of Theorem 17; the second yields the following analogue of Theorem 17 for multiplication.

Corollary 23.1. *If $km = \ell m$, then $k = \ell$.*

Theorem 24. $\mathbb{N}$ *is well-ordered by $<$: every nonempty set of natural numbers has a least element.*

Proof. Suppose A is a set of natural numbers with no least element. Let B be the set of natural numbers n such that, if $m \leqslant n$, then $m \notin A$. Then $1 \in B$, since otherwise 1 would be the least element of A. Suppose $m \in B$. Then $m + 1 \in B$, since otherwise $m + 1$ would be the least element of A. By induction, $B = \mathbb{N}$, so $A = \varnothing$. □

The members of $\mathbb{N}$ are the **positive integers;** the full set $\mathbb{Z}$ of *integers* will be defined formally in §2.7 below, on page 54. As presented in Books VII–IX of Euclid's *Elements* [18, 17], number theory is a study of the positive integers; but a consideration of all integers is useful in this study, and the integers that will constitute a motivating example, first of a group (page 62), and then of a ring (page 73).

2.5. A construction of the natural numbers

For an arbitrary set a, let

$$a' = a \cup \{a\}.$$

If A belongs to the class $\mathbf{I}$ defined in (2.8) on page 29, then $0 \in A$, and A is closed under the operation $x \mapsto x'$, and so $(A, 0, ')$ is an iterative structure. In particular, $(\omega, 0, ')$ is an iterative structure by Theorem 8 (page 29).

Theorem 25. *The structure $(\omega, 0, ')$ satisfies the Peano Axioms.*

Proof. There are three things to prove.

1. In $(\omega, 0, {}')$, the initial element 0 is not a successor, because for all sets a, the set a' contains a, so it is nonempty.

2. $(\omega, 0, {}')$ admits induction, because, if $A \subseteq \omega$, and A contains 0 and is closed under $x \mapsto x'$, then $A \in \mathbf{I}$, so $\bigcap \mathbf{I} \subseteq A$ by Theorem 7 (page 28), that is, $\omega \subseteq A$.

3. It remains to establish that $x \mapsto x'$ is injective on ω. On page 41, we used recursion to define a relation $<$ on $\mathbb{N}$ so that

$$m \not< 1, \qquad m < n + 1 \Leftrightarrow m < n \lor m = n. \tag{2.13}$$

Everything that we proved about this relation required only these properties, and induction. On ω, we do not know whether we have recursion, but we have (2.13) when $<$ is $\in$ and 1 is 0: that is, we have

$$m \notin 0, \qquad m \in n' \Leftrightarrow m \in n \lor m = n.$$

Therefore $\in$ must be a linear ordering of ω, by the proofs in the previous section. Thus, if $m \neq n$, then either $m \in n$ or $n \in m$. We also have the last lemma in that section for $\in$, that is, if $m \in n$, then either $m' = n$ or $m' \in n$; and in either case, $m' \in n'$, so $m' \neq n'$. Thus, assuming $m \neq n$, we have $m' \neq n'$. □

Given sets A and B, we define

$$A \smallsetminus B = \{x \in A \colon x \notin B\}.$$

As a corollary of the foregoing theorem, we have that the iterative structure $(\omega \smallsetminus \{0\}, 1, {}')$ also satisfies the Peano Axioms. We may henceforth assume that $(\mathbb{N}, 1, x \mapsto x + 1)$ is this structure. In particular,

$$\mathbb{N} = \omega \smallsetminus \{0\}.$$

Thus we no longer need the Peano Axioms as axioms; they are theorems about $(\mathbb{N}, 1, x \mapsto x + 1)$ and $(\omega, 0, {}')$.

We extend the definitions of addition and multiplication on $\mathbb{N}$ to allow their arguments to be 0:

$$n + 0 = n = 0 + n, \qquad n \cdot 0 = 0 = 0 \cdot n.$$

Theorem 26. *Addition and multiplication are commutative and associative on ω, and multiplication distributes over addition.*

In particular, the equations (2.11) (page 39) making up the recursive definitions of addition and multiplication on $\mathbb{N}$ are still valid on ω. The same goes for factorials and exponentiation when we define

$$0! = 1, \qquad\qquad n^0 = 1.$$

2.6. Structures

For us, the point of using the von-Neumann definition of the natural numbers is that, under this definition, a natural number n is a particular set, namely $\{0,\dots,n-1\}$, with n elements. We denote the set of functions from a set B to a set A by

$$A^B.$$

In particular then, A^n is the set of functions from $\{0,\dots,n-1\}$ into A. We can denote such a function by one of

$$(x_0,\dots,x_{n-1}), \qquad\qquad (x_i \colon i < n),$$

so that

$$A^n = \{(x_0,\dots,x_{n-1}) \colon x_i \in A\}.$$

Thus, A^2 can be identified with $A \times A$, and A^1 with A itself. There is exactly one function from 0 to A, namely 0; so

$$A^0 = \{0\} = 1.$$

An n-ary **relation** on A is a subset of A^n; an **n-ary operation** on A is a function from A^n to A. Relations and operations that are 2-ary, 1-ary, or 0-ary can be called **binary, singulary,** or **nullary,** respectively; after the appropriate identifications, this agrees with the terminology used in §2.3. A nullary operation on A can be identified with an element of A.

Generalizing the terminology used at the beginning of §2.4 (page 35), we define a **structure** as a set together with some distinguished relations and operations on the set; as before, the set is the **universe** of the structure. If the underlying set of a structure is denoted by a Latin letter, as A or B, then the structure itself may be denoted by the corresponding Fraktur letter, as $\mathfrak{A}$ or $\mathfrak{B}$. See Appendix A, page 255.

The **signature** of a structure comprises a symbol for each distinguished relation and operation of the structure. For example, we have so far obtained $\mathbb{N}$ as a structure in the signature $\{1, +, \cdot, <\}$. We may then write out this structure as

$$(\mathbb{N}, 1, +, \cdot, <).$$

In this way of writing the structure, an expression like $+$ stands not for the *symbol* of addition, but for the actual operation on $\mathbb{N}$. In general, if s is a symbol of the signature of $\mathfrak{A}$, then the corresponding relation or operation on A can, for precision, be denoted by

$$s^{\mathfrak{A}}.$$

Then $s^{\mathfrak{A}}$ is the **interpretation** of s in $\mathfrak{A}$.

The reason why we might distinguish $s^{\mathfrak{A}}$ from s is that two structures can have the same signature. We must be clear what this means. Each symbol of a signature carries with it two pieces of information:

1) whether it symbolizes a relation or an operation, and
2) for which n in ω the relation or operation is n-ary.

A relation symbol can be called a **predicate;** a nullary operation symbol can be called a **constant.** More than one symbol in a signature can symbolize an n-ary relation or an n-ary operation. But we normally do not consider the sign $=$ of equality to belong to a signature.

If $\mathscr{S}$ is a signature, we denote the class of all structures with this signature by

$$\mathbf{Str}_{\mathscr{S}}.$$

Suppose $\mathfrak{A}$ and $\mathfrak{B}$ belong to $\mathbf{Str}_{\mathscr{S}}$. If $s \in \mathscr{S}$, then $s^{\mathfrak{A}}$ is an n-ary operation or relation on A if and only if $s^{\mathfrak{B}}$ is an n-ary operation or relation on B, respectively. A **homomorphism** from $\mathfrak{A}$ to $\mathfrak{B}$ is a function h from A to B that *preserves* the relations and operations symbolized in $\mathscr{S}$: this means

$$\begin{gathered} h(f^{\mathfrak{A}}(x_0,\dots,x_{n-1})) = f^{\mathfrak{B}}(h(x_0),\dots,h(x_{n-1})), \\ (x_0,\dots,x_{n-1}) \in R^{\mathfrak{A}} \Rightarrow (h(x_0),\dots,h(x_{n-1})) \in R^{\mathfrak{B}}, \end{gathered} \tag{2.14}$$

for all n-ary operation symbols f of $\mathscr{S}$ and all n-ary predicates R of $\mathscr{S}$, for all n in ω. To indicate that h is a homomorphism from $\mathfrak{A}$ to $\mathfrak{B}$, we may write

$$h\colon \mathfrak{A} \to \mathfrak{B}$$

(rather than simply $h\colon A \to B$). We have already seen a special case of a homomorphism in the Recursion Theorem (Theorem 15, page 37). The following is easily proved.

Theorem 27. *If $h\colon \mathfrak{A} \to \mathfrak{B}$ and $g\colon \mathfrak{B} \to \mathfrak{C}$, then*

$$g \circ h\colon \mathfrak{A} \to \mathfrak{C}.$$

A homomorphism is an **embedding** if it is injective and if the converse of (2.14) also holds. A surjective embedding is an **isomorphism.**

Theorem 28. *The function id_A is an isomorphism from $\mathfrak{A}$ to itself. The following are equivalent conditions on a bijective homomorphism h from $\mathfrak{A}$ to $\mathfrak{B}$:*

1) $\mathfrak{B}$ is an isomorphism from $\mathfrak{A}$ to $\mathfrak{B}$,
2) h^{-1} is a homomorphism from $\mathfrak{B}$ to $\mathfrak{A}$,
3) h^{-1} is an isomorphism from $\mathfrak{B}$ to $\mathfrak{A}$.

If there is an isomorphism from a structure $\mathfrak{A}$ to a structure $\mathfrak{B}$, then these two structures are said to be **isomorphic** to one another, and we may write

$$\mathfrak{A} \cong \mathfrak{B}.$$

In this case $\mathfrak{A}$ and $\mathfrak{B}$ are indistinguishable as structures, and so (out of laziness perhaps) we may *identify* them, treating them as the *same* structure. We have already done this, in a sense, with $(\mathbb{N}, 1, x \mapsto x+1)$

and $(\omega\smallsetminus\{0\},1,{}')$. However, we never actually had a set called $\mathbb{N}$, until we identified it with $\omega\smallsetminus\{0\}$.

A **substructure** of a structure $\mathfrak{B}$ is a structure $\mathfrak{A}$ of the same signature such that $A\subseteq B$ and the **inclusion** $x\mapsto x$ of A in B is an embedding of $\mathfrak{A}$ in $\mathfrak{B}$. To indicate that $\mathfrak{A}$ is a substructure of $\mathfrak{B}$, we may write

$$\mathfrak{A}\subseteq\mathfrak{B}.$$

Model theory studies structures as such. **Universal algebra** studies **algebras,** which are sets with distinguished operations, but no distinguished relations. In other words, an algebra is a structure in a signature with no symbols for relations.

We shall study mainly the algebras called *groups* and the algebras called *rings.* Meanwhile, we have the algebra $(\mathbb{N},1,+,\cdot)$, and we shall have more examples in the next section.

A **reduct** of a structure is obtained by ignoring some of its operations and relations, while the universe remains the same. The original structure is then an **expansion** of the reduct. For example, $(\mathbb{N},+)$ is a reduct of $(\mathbb{N},+,\cdot,<)$, and the latter is an expansion of the former.

Let us finally note that the universe of a structure is normally considered to be a set, and not just a class. Thus the *universal class* $\{x\colon x=x\}$ is not the universe of a structure with signature $\{\in\}$. Set theory does study structures in this signature that have some of the properties of the universal class. We shall not do this. However, in order to talk precisely about structures as such, in Chapter 5 (page 132) we shall adapt the logic that we developed in §2.2 (page 17) for talking about sets.

2.7. Constructions of the integers and rationals

The next theorem is an example of something like *localization,* which will be the topic of §4.5 (p. 121). One learns the theorem implicitly in school, when one learns about fractions. Perhaps fractions are our first encounter with nontrivial *equivalence classes.*

On page 41, we defined an *ordering* as an irreflexive, transitive relation on a set. A relation R on a set A is **reflexive** and **symmetric** if, respectively,

$$a\ R\ a, \qquad\qquad a\ R\ b \Leftrightarrow b\ R\ a,$$

for all a and b in A. A reflexive, symmetric, transitive relation on a set is an **equivalence relation** on that set. If E is an equivalence relation on A, and $a \in A$, then the set

$$\{x \in A \colon a\ E\ x\}$$

is the **equivalence class** of a in A with respect to R. We may denote by

$$A/E$$

the set of equivalence classes of elements of A with respect to E; this is the **quotient** of A by E. Since the equivalence class of an element of A contains that element and is included in A, we have

$$A = \bigcup A/E.$$

Moreover, two distinct equivalence classes are disjoint, and so A is the **disjoint union** of A/E.

Now let $\approx$ be the binary relation on $\mathbb{N} \times \mathbb{N}$ given by[16]

$$(a, b) \approx (x, y) \Leftrightarrow ay = bx. \tag{2.15}$$

Lemma 6. *The relation $\approx$ is an equivalence relation on $\mathbb{N} \times \mathbb{N}$.*

If $(a, b) \in \mathbb{N} \times \mathbb{N}$, let its equivalence class with respect to $\approx$ be denoted by either of

$$a/b, \qquad\qquad \frac{a}{b}.$$

Let the set $(\mathbb{N} \times \mathbb{N})/{\approx}$ of all such equivalence classes be denoted by

$$\mathbb{Q}^{+}.$$

This set comprises the **positive rational numbers.**

[16] As a binary relation on $\mathbb{N} \times \mathbb{N}$, the relation $\approx$ is a subset of $(\mathbb{N} \times \mathbb{N})^2$, which we identify with $\mathbb{N}^4$.

Theorem 29. *There are* well-defined *binary operations* $+$ *and* $\cdot$ *on* $\mathbb{Q}^+$ *given by the rules*

$$\frac{a}{b}+\frac{x}{y}=\frac{ay+bx}{by}, \qquad\qquad \frac{a}{b}\cdot\frac{x}{y}=\frac{ax}{by}. \tag{2.16}$$

There is a well-defined singulary operation ${}^{-1}$ *on* $\mathbb{Q}^+$ *given by*

$$\left(\frac{x}{y}\right)^{-1}=\frac{y}{x}. \tag{2.17}$$

There is a well-defined linear ordering $<$ *of* $\mathbb{Q}^+$ *given by*

$$\frac{a}{b}<\frac{x}{y}\Leftrightarrow ay<bx. \tag{2.18}$$

The structure $(\mathbb{N},+,\cdot,<)$ *embeds in* $(\mathbb{Q}^+,+,\cdot,<)$ *under the map*

$$x\mapsto\frac{x}{1}.$$

Addition and multiplication are commutative and associative on $\mathbb{Q}^+$*, and multiplication distributes over addition. Moreover,*

$$\frac{1}{1}\cdot\frac{x}{y}=\frac{x}{y}, \qquad\qquad \left(\frac{x}{y}\right)^{-1}\cdot\frac{x}{y}=\frac{1}{1}, \tag{2.19}$$

Finally,

$$\frac{1}{1}<\frac{a}{b}\wedge\frac{1}{1}<\frac{x}{y}\Rightarrow\frac{1}{1}<\frac{a}{b}\cdot\frac{x}{y}. \tag{2.20}$$

The operations on $\mathbb{Q}^+$ in the theorem are said to be **well defined** because it is not immediately obvious that they exist at all. It is possible that $a/b=c/d$ although $(a,b)\neq(c,d)$. In this case one must check that (for example) $(ay+bx)/(by)=(cy+dx)/(dy)$. See page 76.

Because multiplication is commutative and associative on $\mathbb{Q}^+$, and (2.19) holds, the structure $(\mathbb{Q}^+,1/1,{}^{-1},\cdot)$ is an **abelian group.** Because in addition $\mathbb{Q}^+$ is linearly ordered by $<$, and (2.20) holds, the structure $(\mathbb{Q}^+, 1/1, {}^{-1}, \cdot, <)$ is an **ordered group.**[17] The **positive** elements of this group are those elements a such that $a > 1/1$, although we do not usually use this terminology when, as at present, the ordered group is written *multiplicatively.*

For the moment, a natural number is *not* a positive rational number. Therefore, even though we already have a function $(x, y) \mapsto x/y$ from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{Q}^+$, we are free to use the same notation to define a binary operation on $\mathbb{Q}^+$. This operation will be given by

$$\frac{x}{y} = x \cdot y^{-1}. \tag{2.21}$$

We easily have the following.

Theorem 30. *For all m and n in $\mathbb{N}$,*

$$\frac{m/1}{n/1} = \frac{m}{n}. \tag{2.22}$$

The rules (2.16), (2.17), *and* (2.18) *are correct when the letters range over $\mathbb{Q}^+$.*

The meaning of (2.22) is that the diagram in Figure 2.2 commutes. We may now *identify* n and $n/1$, treating them as the same thing. Then $\mathbb{N} \subseteq \mathbb{Q}^+$, and the function $(x, y) \mapsto x/y$ from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{Q}^+$ is just the restriction of the binary operation on $\mathbb{Q}^+$.

In the definition (2.15) of $\approx$, if we replace multiplication with addition, then instead of the positive rational numbers, we obtain the *integers.* Probably this construction of the integers is not learned in school. If it were, possibly students would never think that $-x$ is automatically a negative number. In any case, by applying this construction of the integers to the positive rational numbers, we obtain all of the rational numbers as follows. Let $\sim$ be the binary relation on $\mathbb{Q}^+ \times \mathbb{Q}^+$ given by

$$(a, b) \sim (x, y) \Leftrightarrow a + y = b + x. \tag{2.23}$$

[17] In particular, all of our ordered groups will be abelian.

$$\begin{tikzcd}
\mathbb{N}\times\mathbb{N} \arrow[rr, "(x,y)\mapsto \frac{x}{y}"] \arrow[dr, "(x,y)\mapsto\left(\frac{x}{1},\frac{y}{1}\right)"'] & & \mathbb{Q}^+ \\
& \mathbb{Q}^+\times\mathbb{Q}^+ \arrow[ur, "(x,y)\mapsto x\cdot y^{-1}"'] &
\end{tikzcd}$$

Figure 2.2.: Division of positive rationals

Lemma 7. *The relation $\sim$ on $\mathbb{Q}^+ \times \mathbb{Q}^+$ is an equivalence relation.*

If $(a, b) \in \mathbb{Q}^+ \times \mathbb{Q}^+$, let its equivalence class with respect to $\sim$ be denoted by

$$a - b.$$

Let the set of such equivalence classes be denoted by

$$\mathbb{Q}.$$

This set comprises the **rational numbers.** However, for the moment, a positive rational number is not a rational number. We denote by

$$0$$

the rational number $1 - 1$.

Theorem 31. *There are well-defined operations $-$, $+$, and $\cdot$ on $\mathbb{Q}$ given by the rules*

$$\left.\begin{aligned} -(x-y) &= y-x, \\ (a-b)+(x-y) &= (a+x)-(b+y), \\ (a-b)\cdot(x-y) &= (ax+by)-(ay+bx). \end{aligned}\right\} \tag{2.24}$$

There is a dense linear ordering $<$ of $\mathbb{Q}$ given by

$$a - b < x - y \Leftrightarrow a + y < b + x.$$

The structure $(\mathbb{Q}^+, +, \cdot, <)$ embeds in $(\mathbb{Q}, +, \cdot, <)$ under the map

$$x \mapsto (x + 1) - 1.$$

The structure $(\mathbb{Q}, 0, -, +, <)$ is an ordered group, and its positive elements are just those in the image of $\mathbb{Q}^+$. Multiplication is also commutative and associative on $\mathbb{Q}$, and it distributes over addition.

As before, although we already have a function $(x, y) \mapsto x - y$ from $\mathbb{Q}^+ \times \mathbb{Q}^+$ to $\mathbb{Q}$, we are free to use the same notation for a binary operation on $\mathbb{Q}$ given by

$$x - y = x + (-y).$$

We easily have the following.

Theorem 32. *For all a and b in $\mathbb{Q}^+$,*

$$\big((a + 1) - 1\big) - \big((b + 1) - 1\big) = a - b. \qquad (2.25)$$

The rules (2.24) are correct when the letters range over $\mathbb{Q}$.

The meaning of (2.25) is that the diagram in Figure 2.3 commutes. We now identify $\mathbb{Q}^+$ with its image in $\mathbb{Q}$, so that a positive rational number is indeed just a rational number that is positive.

Theorem 33. $\mathbb{Q}^+ = \{x \in \mathbb{Q} \colon 0 < x\}$. *The singulary operation $^{-1}$ on $\mathbb{Q}^+$ extends to an operation on $\mathbb{Q} \smallsetminus \{0\}$ when*

$$x^{-1} = -(-x)^{-1}$$

on $\{-x \colon x \in \mathbb{Q}^+\}$. Then $(\mathbb{Q} \smallsetminus \{0\}, 1, {}^{-1}, \cdot)$ is an abelian group. The binary operation $/$ on $\mathbb{Q}^+$ is the restriction of the function $/$ from $\mathbb{Q} \times (\mathbb{Q} \smallsetminus \{0\})$ to $\mathbb{Q}$ given by (2.21), and (2.16), (2.17), and (2.18) hold when the letters range over $\mathbb{Q}$ (and the expressions are defined).

$$\begin{tikzcd}
\mathbb{Q}^+ \times \mathbb{Q}^+ \arrow[rr, "(x{,}y)\mapsto x-y"] \arrow[dr, "(x{,}y)\mapsto\big((x+1)-1{,}(y+1)-1\big)"'] & & \mathbb{Q} \\
& \mathbb{Q} \times \mathbb{Q} \arrow[ur, "(x{,}y)\mapsto x+(-y)"'] &
\end{tikzcd}$$

Figure 2.3.: Subtraction of rationals

Because $(\mathbb{Q}, 0, -, +, <)$ and $(\mathbb{Q}^+, 1, {}^{-1}, \,\cdot\,, <)$ are ordered groups, where $\mathbb{Q}^+ = \{x \in \mathbb{Q} \colon x > 0\}$, and multiplication distributes over addition in $\mathbb{Q}$, the structure $(\mathbb{Q}, 0, -, +, 1, \cdot, <)$ is an **ordered field.** However, the ordering of $\mathbb{Q}$ is not **complete,** that is, there are subsets with upper bounds, but no *suprema* (least upper bounds). An example is the set $\{x \in \mathbb{Q} \colon 0 < x \land x^2 < 2\}$.

We can now define

$$\mathbb{Z} = \{x - y \colon (x, y) \in \mathbb{N} \times \mathbb{N}\};$$

this is the subset of $\mathbb{Q}$ comprising the **integers.**

Theorem 34.

1. $(\mathbb{Z}, 0, -, +, 1, \cdot, <) \subseteq (\mathbb{Q}, 0, -, +, 1, \cdot, <)$.
2. *In particular,* $(\mathbb{Z}, 0, -, +, 1, \cdot, <)$ *is well defined.*
3. $(\mathbb{Z}, 0, -, +, <)$ *is an ordered group.*
4. $\mathbb{Q} = \{x/y \colon x \in \mathbb{Z} \land y \in \mathbb{Z} \smallsetminus \{0\}\}$.

Because of the theorem, we can also think of $\mathbb{Q}$ as arising from $\mathbb{Z}$ by the same construction that gives us $\mathbb{Q}^+$ from $\mathbb{N}$. We shall generalize this construction of $\mathbb{Q}$ in §4.5 (page 121).

2.8. A construction of the reals

There is a technique due to Dedekind for completing $(\mathbb{Q}, <)$ to obtain the completely ordered set $(\mathbb{R}, <)$. As Dedekind says explicitly [13, pp. 39–40], the original inspiration for the technique is the definition of *proportion* found in Book V of Euclid's *Elements* [18, 17].

In the geometry of Euclid, a *straight line* is what we now call a *line segment* or just *segment,* and two segments are *equal* to one another if they are congruent to one another. Congruence of segments is an equivalence relation. Let us refer to a congruence class of segments as the *length* of any one of its members. Two lengths can be *added* together by taking two particular segments with those lengths and setting them end to end. Then lengths of segments compose the set of positive elements of an ordered group. In particular, individual lengths can be *multiplied,* in the original sense of being taken several times. Indeed, if A is a length, and $n \in \mathbb{N}$, we can define the multiple nA of x recursively:

$$1A = A, \qquad\qquad (n+1)A = nA + A.$$

It is assumed that, for any two lengths A and B, some multiple of A is greater than B: this is the **archimedean property.** If C and D are two more lengths, then A has to B the *same ratio* that C has to D, provided that, for all k and m in $\mathbb{N}$,

$$kA > mB \Leftrightarrow kC > mD.$$

In this case, the four lengths A, B, C, and D are *proportional,* and we may write

$$A : B :: C : D.$$

We can write the condition for this proportionality as

$$\left\{\frac{x}{y} \in \mathbb{Q}^+ : xB < yA\right\} = \left\{\frac{x}{y} \in \mathbb{Q}^+ : xD < yC\right\}$$

Dedekind's observation is that such sets can be defined independently of all geometrical considerations. Indeed, we may define a **positive real number** as a nonempty, proper subset C of $\mathbb{Q}^+$ such that

1) if $a \in C$ and $b \in \mathbb{Q}^+$ and $b < a$, then $b \in C$, and
2) if C has a supremum in $\mathbb{Q}^+$, this supremum does not belong to C.

Let the set of all positive real numbers be denoted by

$$\mathbb{R}^+.$$

Theorem 35. *The set $\mathbb{R}^+$ is completely ordered by proper inclusion. There are well-defined operations $+$, ${}^{-1}$, and $\cdot$ on $\mathbb{Q}^+$ given by the rules*

$$\begin{gathered} C + D = \{x + y \colon x \in C \wedge y \in D\}, \\ C^{-1} = \{x^{-1} \colon x \in \mathbb{Q}^+ \wedge \exists y\ (y \in \mathbb{Q}^+ \smallsetminus C \wedge y < x)\}, \\ C \cdot D = \{x \cdot y \colon x \in C \wedge y \in D\}. \end{gathered}$$

Then $(\mathbb{Q}^+, +, {}^{-1}, \cdot)$ embeds in $(\mathbb{R}^+, +, {}^{-1}, \cdot)$ under $y \mapsto \{x \in \mathbb{Q}^+ \colon x < y\}$.

Let us identify $\mathbb{Q}^+$ with its image in $\mathbb{R}^+$. We may also write $\subset$ on $\mathbb{R}^+$ as $<$.

For every n in ω, an n-ary operation f on $\mathbb{R}^+$ is **continuous** if, for every $(A_i \colon i < n)$ in $(\mathbb{R}^+)^n$, for every ε in $\mathbb{Q}^+$, there is $(\delta_i \colon i < n)$ in $(\mathbb{Q}^+)^n$ such that, for all $(X_i \colon i < n)$ in $(\mathbb{R}^+)^n$, if

$$\bigwedge_{i<n} A_i - \delta_i < X_i < A_i + \delta_i,$$

then

$$f(A_i \colon i < n) - \varepsilon < f(X_i \colon i < n) < f(A_i \colon i < n) + \varepsilon.$$

Theorem 36. *The operations $+$, ${}^{-1}$, and $\cdot$ on $\mathbb{R}^+$ are continuous. Every composite of continuous functions on $\mathbb{R}^+$ is continuous.*

Lemma 8. *The only continuous singulary operation on $\mathbb{R}^+$ that is 1 on $\mathbb{Q}$ is 1 everywhere.*

Theorem 37. *The structure $(\mathbb{R}^+, 1, {}^{-1}, \cdot, <)$ is an ordered group, and addition is commutative and associative on $\mathbb{R}^+$, and multiplication distributes over addition on $\mathbb{R}^+$.*

Now define $\sim$ on $\mathbb{R}^+ \times \mathbb{R}^+$ as in (2.23). Just as before, this is an equivalence relation. The set of its equivalence classes is denoted by

$$\mathbb{R}.$$

Just as before, we obtain the ordered field $(\mathbb{R}, 0, -, +, {}^{-1}, \cdot, <)$. But now, the ordering is complete. We identify $\mathbb{R}^+$ with its image in $\mathbb{R}$. The elements of $\mathbb{R}$ are the **real numbers.**

Lemma 9. *For every n in $\mathbb{N}$, for every element A of a completely and densely ordered group, the equation*

$$nX = A$$

is soluble in the group.

Theorem 38. *Suppose $(G, 0, -, +, <)$ is a completely and densely ordered group, and u is a positive element of G, and b is an element of $\mathbb{R}^+$ such that $1 < b$. Then there is an isomorphism from $(G, 0, -, +, <)$ to $(\mathbb{R}^+, 1, {}^{-1}, \cdot, <)$ taking u to b.*

By the theorem, the completely ordered groups $(\mathbb{R}, 0, -, +, <)$ and $(\mathbb{R}^+, 1, {}^{-1}, \cdot, <)$ are isomorphic, and indeed for every b in $\mathbb{R}^+$ such that $b > 1$, there is an isomorphism taking 1 to b. This isomorphism is denoted by

$$x \mapsto b^x,$$

and its inverse is

$$x \mapsto \log_b x.$$

Theorem 39 (Intermediate Value Theorem)**.** *If f is a continuous singulary operation on $\mathbb{R}$, and $f(a) \cdot f(b) < 0$, then f has a zero between a and b.*

Hence for example the function $x \mapsto x^2 - 2$ must have a zero in $\mathbb{R}$ between 1 and 2. More generally, if $A \subseteq \mathbb{R}$, then the set of *polynomial functions over A* is obtained from the set of constant functions taking values in A, along with $-$, $+$, $\cdot$, and the projections $(x_0, \dots, x_{n-1}) \mapsto$

x_i, by closing under taking composites. Then all polynomial functions over $\mathbb{R}$ are continuous, and so the Intermediate Value Theorem applies to the singulary polynomial functions. Therefore the ordered field $\mathbb{R}$ is said to be **real-closed.** However, there are smaller real-closed ordered fields: we establish this in the next section.

2.9. Countability

A set is **countable** if it embeds in ω; otherwise the set is **uncountable.**

Theorem 40. *The sets $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$ are all countable.*

Theorem 41. *$\mathscr{P}(\omega)$ is uncountable.*

Proof. Suppose f is an injection from ω to $\mathscr{P}(\omega)$. Then the subset $\{x\colon x \notin f(x)\}$ of ω is not in the range of f, by a variant of the Russell Paradox: if $\{x\colon x \notin f(x)\} = f(a)$, then $a \in f(a) \Leftrightarrow a \notin f(a)$. □

Theorem 42. *The set $\mathbb{R}$ is uncountable.*

Proof. For every subset A of ω, let $g(A)$ be the set of rational numbers x such that, for some n in ω,

$$x < \sum_{k \in A \cap n} \frac{2}{3^k}.$$

Then $g(A)$ is a real number by the original definition. The function $A \mapsto g(A)$ from $\mathscr{P}(\omega)$ to $\mathbb{R}$ is injective. □

In the theorem, the image of the function g is the *Cantor Set*; see page 106.

If $A \subseteq \mathbb{R}$, suppose we let A^{rc} be the smallest field that contains all zeros from $\mathbb{R}$ of singulary polynomial functions over A. If we define $A_0 = \mathbb{Q}$ and $A_{n+1} = {A_n}^{\mathrm{rc}}$, then $\bigcup_{n\in\omega} A_n$ will contain all zeros from $\mathbb{R}$ of singulary polynomial functions over itself. Thus it will be real-closed. In fact it will be $\mathbb{Q}^{\mathrm{rc}}$. But this field is countable.

We can say more about a set than whether it is countable or uncountable. A class is **transitive** if it properly includes all of its elements. A transitive *set* is an **ordinal** if it is well-ordered by the relation of membership. Then all of the elements of ω are ordinals, and so is ω itself. The class of all ordinals can be denoted by

$$\mathbf{ON}.$$

Theorem 43. *The class* **ON** *is transitive and well-ordered by membership.*

In particular, **ON** cannot contain itself; so it is not a set. This result is the **Burali-Forti Paradox** [7].

Theorem 44. *Every well-ordered set* $(A, <)$ *is isomorphic to a unique ordinal. The isomorphism is a certain function* f *on* A*, and this function is determined by the rule*

$$f(b) = \{f(x) \colon x < b\}.$$

There are three classes of ordinals.

1. A **successor** is an ordinal α' for some ordinal α.
2. The least ordinal, 0, is in a class by itself.
3. A **limit** is an ordinal that is neither 0 nor a successor.

Then ω is the least limit ordinal.

Two sets are **equipollent** if there is a bijection between them. An ordinal is a **cardinal** if it is the least ordinal that is equipollent with it.

Theorem 45. *Every element of* ω *is a cardinal. So is* ω *itself.*

The class of cardinals can be denoted by

$$\mathbf{CN}.$$

Every set is equipollent with at most one cardinal, which is called the **cardinality** or **size** of that set. The cardinality of an arbitrary set A is denoted by

$$|A|.$$

A countable set has cardinality ω or less; uncountable sets have cardinality greater than ω. The **finite** sets are those whose cardinalities are less then ω; other sets are **infinite.**

3. Groups and Rings

3.1. Groups and rings

3.1.1. Groups

Given a set A, we may refer to a bijection from A to itself as a **symmetry** or **permutation** of A. Let us denote the set of these symmetries by

$$\mathrm{Sym}(A).$$

This set can be equipped with:

1) the element id_A, which is the **identity** on A;
2) the singulary operation $f \mapsto f^{-1}$, which is **inversion;**
3) the binary operation $(f, g) \mapsto f \circ g$, which is **composition.**

The 4-tuple

$$(\mathrm{Sym}(A), \mathrm{id}_A, {}^{-1}, \circ)$$

is the **complete group of symmetries** of A. We may speak of the set $\mathrm{Sym}(A)$ as the *underlying set* of the group. We may also use $\mathrm{Sym}(A)$ to denote the group $(\mathrm{Sym}(A), \mathrm{id}_A, {}^{-1}, \circ)$. A **subgroup** of this group is a subset of $\mathrm{Sym}(A)$ that contains e and is closed under inversion and composition. Such a subgroup can be called simply a **group of symmetries** of A. The following is easily verified.

Theorem 46. *For all sets A, for all elements f, g, and h of a group of symmetries of A,*

$$\begin{gathered} f \circ \mathrm{id}_A = f, \quad f \circ f^{-1} = \mathrm{id}_A, \\ \mathrm{id}_A \circ f = f, \quad f^{-1} \circ f = \mathrm{id}_A, \\ (f \circ g) \circ h = f \circ (g \circ h). \end{gathered}$$

A group of symmetries is an example of an **algebra,** that is, a set equipped with some operations. More generally, a **structure** is a set equipped with operations *and* relations. A **group** is an algebra with the properties of a group of symmetries given by the last theorem (Theorem 46). That is, a group is an algebra $(G, \mathrm{e}, {}^{-1}, \cdot)$ in which the following equations are *identities* (are true for all values of the variables):

$$\begin{gathered} x \cdot \mathrm{e} = x, \quad x \cdot x^{-1} = \mathrm{e}, \\ \mathrm{e} \cdot x = x, \quad x^{-1} \cdot x = \mathrm{e}, \\ (x \cdot y) \cdot z = x \cdot (y \cdot z). \end{gathered}$$

We may say also that these equations are the *axioms* of groups: this means that their *generalizations* ($\forall x\ x \cdot \mathrm{e} = x$ and so forth) are true in every group, by definition. According to these axioms, in every group $(G, \mathrm{e}, {}^{-1}, \cdot)$,

1) the binary operation $\cdot$ is **associative,**
2) the element e is an **identity** with respect to $\cdot$,
3) the singulary operation ${}^{-1}$ is **inversion** with respect to $\cdot$ and e.

The identity and the inversion will turn out to be uniquely determined by the binary operation, by Theorem 51 on page 66.

A group is called **abelian** if its binary operation is commutative. If A has at least three elements, then $\mathrm{Sym}(A)$ is not abelian. However, every one-element set $\{a\}$ becomes an abelian group when we define

$$\mathrm{e} = a, \qquad a^{-1} = a, \qquad a \cdot a = a.$$

This group is a **trivial group.** For example, both $\mathrm{Sym}(0)$ and $\mathrm{Sym}(1)$ are trivial groups. All trivial groups are isomorphic to one another. Therefore we tend to identify them with one another, referring to each of them as *the* trivial group, which we shall denote by

$$\{\mathrm{e}\}.$$

Besides this and the symmetry groups, we have the following seven examples of groups, namely

$$(\mathbb{Z}, 0, -, +), \qquad (\mathbb{Q}, 0, -, +), \qquad (\mathbb{R}, 0, -, +)$$

along with

$$\begin{aligned}
&(\mathbb{Q}^+, 1, {}^{-1}, \cdot), &\qquad &(\mathbb{Q} \smallsetminus \{0\}, 1, {}^{-1}, \cdot),\\
&(\mathbb{R}^+, 1, {}^{-1}, \cdot), &\qquad &(\mathbb{R} \smallsetminus \{0\}, 1, {}^{-1}, \cdot),
\end{aligned}$$

In the first three examples, the symbols $-$ and $+$ mean something different in each case, although we understand 0 to be the same in each case. In the last four examples, the symbols ${}^{-1}$ and $\cdot$ mean something different in each case, although we understand 1 to be the same in each case. All seven examples are abelian. The last four of them are the origin of a terminological convention. In an arbitrary group $(G, \mathrm{e}, {}^{-1}, \cdot)$, the operation $\cdot$ is usually called **multiplication.** We usually write $g \cdot h$ as gh. The element g^{-1} is the **inverse** of g. The element e is the **identity,** and it is sometimes denoted by 1 rather than e.

Evidently the first three examples use different notation. These groups are said to be written **additively.** Additive notation is often used for abelian groups, but almost never for other groups. It will be useful to have one more example of an abelian group. Actually there will be one example for each positive integer. If a and b are arbitrary integers for which the equation

$$ax = b$$

has a solution in $\mathbb{Z}$, then we say that a **divides** b, or a is a **divisor** or **factor** of b, or b is a **multiple** of a, and we may write

$$a \mid b.$$

Using the notation due to Gauss [22, p. 1], for a positive integer n and arbitrary integers a and b we write

$$a \equiv b \pmod{n}$$

if $n \mid a - b$. In this case we say a and b are **congruent** with respect to the **modulus** n. This manner of speaking is abbreviated by putting the Latin word *modulus* into the ablative case: a and b are congruent ***modulo*** n.[1] Still following Gauss, we may say too that a is a **residue** of b with respect to the modulus n.

Theorem 47. *Let $n \in \mathbb{N}$.*

1. *Congruence* modulo n *is an equivalence relation on* $\mathbb{Z}$.
2. *If $a \equiv x$ and $b \equiv y \pmod{n}$, then*

$$-a \equiv -x \;\;\&\;\; a+b \equiv x+y \;\;\&\;\; ab \equiv xy \pmod{n}.$$

The set of congruence-classes of integers *modulo* n can be denoted by

$$\mathbb{Z}_n.$$

If a is some integer, we can denote its congruence-class *modulo* n by something like $[a]$ or $\bar{a}$, or more precisely by

$$a + n\mathbb{Z}.$$

Theorem 48. *For every positive integer n, the function*

$$x \mapsto x + n\mathbb{Z}$$

from $\{0, \dots, n-1\}$ to $\mathbb{Z}_n$ is a bijection.

Again given a positive integer n, we may treat an arbitary integer as a name for its own congruence-class *modulo* n. In particular, by the last theorem, we may consider $\mathbb{Z}_n$ as being the set $\{0, \dots, n-1\}$, where these n elements are understood to be distinct. By Theorem 47, we have a well-defined algebra $(\mathbb{Z}_n, 0, -, +, 1, \cdot)$, where 0 and 1 stand for their respective congruence-classes $n\mathbb{Z}$ and $1 + n\mathbb{Z}$. The following theorem is now an easy consequence of Theorem 47.

Theorem 49. *For each n in $\mathbb{N}$, the algebra $(\mathbb{Z}_n, 0, -, +)$ is an abelian group.*

[1] The ablative case of Latin corresponds roughly to the *-den hali* of Turkish. Gauss writes in Latin; however, instead of *modulo n*, he says *secundum modulum n*, "according to the modulus *n*" [23, p. 2].

A **homomorphism** from a group $(G, \mathrm{e}, {}^{-1}, \cdot)$ to a group $(H, \mathrm{e}, {}^{-1}, \cdot)$ is a function h from G to H that "preserves structure" in the sense that

$$h(\mathrm{e}) = \mathrm{e}, \qquad h(x^{-1}) = h(x)^{-1}, \qquad h(x \cdot y) = h(x) \cdot h(y).$$

An **embedding** of groups is a homomorphism that is injective as a function. The (multiplicative) groups of positive rational numbers, of nonzero rational numbers, of positive real numbers, and of nonzero real numbers, and the (additive) groups of integers, rational numbers, real numbers, and integers with respect to some modulus, are not obviously symmetry groups. But they can be *embedded* in symmetry groups. Indeed, every element g of a group G (written multiplicatively) determines a singulary operation λ_g on G, given by

$$\lambda_g(x) = gx.$$

(Here λ stands for "left" as in "multiplication from the left.") Then we have the following.

Theorem 50 (Cayley). *For every group $(G, \mathrm{e}, {}^{-1}, \cdot)$, the function*

$$x \mapsto \lambda_x$$

embeds $(G, \mathrm{e}, {}^{-1}, \cdot)$ in the group $(\mathrm{Sym}(G), \mathrm{id}_G, {}^{-1}, \circ)$ of symmetries.

Proof. We first observe that

$$\lambda_{\mathrm{e}} = \mathrm{id}_G, \qquad \lambda_{g \cdot h} = \lambda_g \circ \lambda_h,$$

because

$$\begin{gathered} \lambda_{\mathrm{e}}(x) = \mathrm{e} \cdot x = x = \mathrm{id}_G(x), \\ \lambda_{g \cdot h}(x) = (g \cdot h) \cdot x = g \cdot (h \cdot x) = \lambda_g(\lambda_h(x)) = (\lambda_g \circ \lambda_h)(x). \end{gathered}$$

Consequently each λ_g has an inverse, and

$$(\lambda_g)^{-1} = \lambda_{g^{-1}}.$$

This establishes that $x \mapsto \lambda_x$ is a homomorphism from $(G, \mathrm{e}, {}^{-1}, \cdot)$ to $(\mathrm{Sym}(G), \mathrm{id}_G, {}^{-1}, \circ)$. It is an embedding, since if $\lambda_g = \lambda_h$, then in particular

$$g = g\,\mathrm{e} = \lambda_g(\mathrm{e}) = \lambda_h(\mathrm{e}) = h\,\mathrm{e} = h. \qquad \square$$

By Cayley's Theorem, every group can be considered as a symmetry group.

3.1.2. Simplifications

A **reduct** of a structure is a structure with the same underlying set, but equipped with fewer operations and relations. The original structure is then called an **expansion** of the reduct. We shall establish that a group $(G, \mathrm{e}, {}^{-1}, \cdot)$ is determined by the reduct $(G, \cdot)$ and that a homomorphism of such reducts is a homomorphism of the whole groups.

A **semigroup** is an algebra $(S, \cdot)$, where $\cdot$ is an associative operation on S. If $(G, \mathrm{e}, {}^{-1}, \cdot)$ is a group, the the reduct $(G, \cdot)$ is a semigroup. Often the semigroup $(G, \cdot)$ itself is called a group. But this usage must be justified.

Theorem 51. *A semigroup can expand to a group in at most one way.*

Proof. Let $(G, \mathrm{e}, ^{-1}, \cdot)$ be a group. If e' were a second identity, then

$$\mathrm{e}'\,x = \mathrm{e}\,x, \qquad \mathrm{e}'\,xx^{-1} = \mathrm{e}\,xx^{-1}, \qquad \mathrm{e}' = \mathrm{e}\,.$$

If a' were a second inverse of a, then

$$a'a = a^{-1}a, \qquad a'aa^{-1} = a^{-1}aa^{-1}, \qquad a' = a^{-1}. \qquad \square$$

Establishing that a particular algebra is a group is made easier by the following.

Theorem 52. *Any algebra satisfying the identities*

$$\mathrm{e}x = x,$$
$$x^{-1}x = \mathrm{e},$$

$$x(yz) = (xy)z$$

is a group. In other words, any semigroup with a left-identity and with left-inverses is a group.

Proof. We need to show $x\,\mathrm{e} = x$ and $xx^{-1} = \mathrm{e}$. To establish the latter, using the given identies we have

$$(xx^{-1})(xx^{-1}) = x(x^{-1}x)x^{-1} = x\mathrm{e}x^{-1} = xx^{-1},$$

and so

$$xx^{-1} = \mathrm{e}xx^{-1} = (xx^{-1})^{-1}(xx^{-1})(xx^{-1}) = (xx^{-1})^{-1}(xx^{-1}) = \mathrm{e}.$$

Hence also

$$x\mathrm{e} = x(x^{-1}x) = (xx^{-1})x = \mathrm{e}x = x. \qquad \square$$

The theorem has an obvious "dual" involving right-identities and right-inverses. By the theorem, the semigroups that expand to groups are precisely the semigroups that satisfy the axiom

$$\exists z\ (\forall x\ zx = x \land \forall x\ \exists y\ yx = z),$$

which is logically equivalent to

$$\exists z\ \forall x\ \forall y\ \exists u\ (zx = x \land uy = z). \tag{3.1}$$

We shall show that this sentence is more complex than need be.

Thanks to Theorem 51, if a semigroup $(G, \cdot)$ does expand to a group, then we may unambiguously refer to $(G, \cdot)$ itself as a group. We may even refer to G as a group, although, theoretically, it may lead to ambiguity.

Theorem 53. *Let G be a nonempty semigroup. The following are equivalent.*

1. *G expands to a group.*
2. *Each equation $ax = b$ and $ya = b$ with parameters from G has a solution in G.*

3. *Each equation $ax = b$ and $ya = b$ with parameters from G has a unique solution in G.*

Proof. Immediately (3)⇒(2). Almost as easily, (1)⇒(3). For, if a and b belong to some semigroup that expands to a group, we have $ax = b \Leftrightarrow x = a^{-1}b$; and we know by Theorem 51 that a^{-1} is uniquely determined. Likewise for $ya = b$.

Finally we show (2)⇒(1). Suppose G is a nonempty semigroup in which all equations $ax = b$ and $ya = b$ have solutions. If $c \in G$, let e be a solution to $yc = c$. If $b \in G$, let d be a solution to $cx = b$. Then

$$\mathrm{e}b = \mathrm{e}(cd) = (\mathrm{e}c)d = cd = b.$$

Since b was chosen arbitrarily, e is a left identity. Since the equation $yc = \mathrm{e}$ has a solution, c has a left inverse. But c is an arbitrary element of G. By Theorem 52, we are done. □

Now we know that the semigroups that expand to groups are just the semigroups that satisfy the axiom

$$\forall x\, \forall y\, (\exists z\ xz = y \wedge \exists w\ wx = y).$$

This may not look simpler than (3.1), but it is. It should be understood as

$$\forall x\, \forall y\, \exists z\, \exists w\, (xz = y \wedge wx = y),$$

which is a sentence of the general form ∀∃; whereas (3.1) is of the form ∃∀∃).

Theorem 54. *A map f from one group to another is a homomorphism, provided it is a homomorphism of semigroups, that is,*

$$f(xy) = f(x)f(y).$$

Proof. In a group, if a is an element, then the identity is the unique solution of $xa = a$, and a^{-1} is the unique solution of $yaa = a$. A semigroup homomorphism f takes solutions of these equations to solutions of $xb = b$ and $ybb = b$, where $b = f(a)$. □

Inclusion of a substructure in a larger structure is a homomorphism. Therefore we have, as a special case of Theorem 54, that if $(G, \mathrm{e}, {}^{-1}, \cdot)$ and $(H, \mathrm{e}, {}^{-1}, \cdot)$ are groups, then

$$(G, \cdot) \subseteq (H, \cdot) \implies (G, \mathrm{e}, {}^{-1}, \cdot) \subseteq (H, \mathrm{e}, {}^{-1}, \cdot).$$

3.1.3. Direct products of groups

As on page 45, if Ω and A are sets, then A^{Ω} is the set of functions from Ω to A. If A is the underlying set of a group, then a multiplication can be defined on A^{Ω} so that this power is also a group. The following will be used on page 74 in case Ω is ω.

Theorem 55. *If Ω is a set and $(G, \cdot)$ is a group, then $(G^{\Omega}, \cdot)$ is a group, where for all f and g in G^{Ω} and all x in Ω,*

$$(f \cdot g)(x) = f(x) \cdot g(x).$$

The foregoing theorem can be generalized as follows. We can think of the power A^{Ω} as the product of copies of A, each copy being indexed by an element of Ω. Then we can replace some of these copies with different sets. To be precise, we define an **indexed set** as a set together with a function whose range is that set. If that function is f in A^B, then the corresponding indexed set can be denoted by

$$(f(x) \colon x \in B)$$

(this notation was introduced on page 13). We may identify this indexed set with the function f itself. Note however that the same set can be the range of many functions with many domains (unless the set is empty; then it is the range of only one function). That is, we may know $\{f(x) \colon x \in B\}$ without knowing what f and B are. However, knowing $(f(x) \colon x \in B)$ means knowing f and hence knowing B and $\{f(x) \colon x \in B\}$.

An indexed set $(a_n \colon n \in \omega)$ is also called a **sequence** and can be written also as

$$(a_0, a_1, a_2, \dots).$$

The word *family* is a synonym for *set;* it is often used for sets whose elements are themselves sets whose elements will be of interest.[2]

Suppose $\mathscr{A}$ is an indexed family $(A_i\colon i \in \Omega)$, where each A_i is a group. We can form the **direct product** of the family $\mathscr{A}$. This direct product is denoted by one of the expressions

$$\prod_{i\in\Omega} A_i, \qquad\qquad \prod \mathscr{A}.$$

If a belongs to this direct product, this means

$$a = (a_i\colon i \in \Omega),$$

where $a_i \in A_i$ in each case. Thus a is simply a function on Ω that, at every element i of this domain, takes a value in A_i; we write this value as a_i, though as in Chapter 1 it could be written also[3] as $a(i)$ or a^i. For each j in Ω, there is a function π_j from $\prod \mathscr{A}$ to A_j given by

$$\pi_j(x) = x_j,$$

so that, for each a in $\prod \mathscr{A}$,

$$a = \bigl(\pi_i(a)\colon i \in \Omega\bigr). \tag{3.2}$$

The function π_j is the **coordinate projection** onto A_j.

Theorem 56. *If $\mathscr{G}$ is an indexed family $\bigl((G_i, \cdot)\colon i \in \Omega\bigr)$ of groups, then $(\prod \mathscr{G}, \cdot)$ is a group, where*

$$(x_i\colon i \in \Omega) \cdot (y_i\colon i \in \Omega) = (x_i \cdot y_i\colon i \in \Omega).$$

[2] In the usual formulation of set theory, every element of every set is itself a set. Since for example a *group* is a set equipped with a certain operation of multiplication, the elements of a group must themselves be sets; but in ordinary mathematics these elements are not thought of as sets, and so one does not refer to the underlying set of a group as a family. One may however speak of a set of groups as a family.

[3] Even the notation i^a might be used. Indeed, x^σ is used below (page 110) for the image under an automorphism σ of an element x of a given field.

Each of the coordinate projections π_j on $\prod \mathscr{G}$ is a homomorphism of groups. If H is a group, and f_j is a homomorphism from $(H, \cdot)$ to $(G_j, \cdot)$ for each j in Ω, then the map

$$x \mapsto \big(f_i(x) \colon i \in \Omega\big)$$

is the unique homomorphism f from H to $\prod \mathscr{G}$ such that, for each j in Ω,

$$\pi_j \circ f = f_j.$$

In the indexed set $(a_i \colon i \in \Omega)$, each element a_i can be called a **term.** Then the multiplication on $\prod_{i \in \Omega} G_i$ defined in the theorem can be described as **termwise** multiplication. The theorem is easily generalized to cover arbitrary algebras and even structures. This will lead to the definition of *ultraproducts.* See for example Theorem 73 on page 83 below.

3.1.4. Rings

A homomorphism from a structure to itself is an **endomorphism.** Recall from page 62 that a group in which the multiplication is commutative is said to be an **abelian group,** and (page 63) its operation is usually written additively. The set of endomorphisms of an abelian group can be made into an abelian group in which:

1) the identity is the constant function $x \mapsto \mathrm{e}$;
2) additive inversion converts f to $x \mapsto -f(x)$;
3) addition converts (f, g) to $x \mapsto f(x) + g(x)$.

If E is an abelian group, let the abelian group of its endomorphisms be denoted by

$$\mathrm{End}(E).$$

A **monoid** is an algebra $(M, \mathrm{e}, \cdot)$, where $\cdot$ is an associative operation, and e is an identity with respect to this operation. The set of endomorphisms of the abelian group E is the underlying set of a monoid in which the identity is the identity function id_E, and multiplication is functional composition. This multiplication distributes in both senses over addition:

$$f \circ (g + h) = f \circ g + f \circ h, \qquad (f + g) \circ h = f \circ h + g \circ h.$$

We may denote the two combined structures—abelian group and monoid together—by

$$(\operatorname{End}(E), \mathrm{id}_E, \circ);$$

this is the **complete ring of endomorphisms of** E. A substructure of $(\operatorname{End}(E), \mathrm{id}_E, \circ)$ can be called simply a **ring of endomorphisms** E.

A **ring** is a structure $(R, 0, -, +, 1, \cdot)$ such that

1) $(R, 0, -, +)$ is an abelian group,
2) $(R, 1, \cdot)$ is a monoid,
3) the multiplication distributes in both senses over addition.

Then rings of endomorphisms are indeed rings. It may be convenient to write a ring as $(R, 1, \cdot)$, where R is implicitly an abelian group. We might even say simply that R is a ring. Let us note the trivial example:

Theorem 57.

1. *In every ring,* $0 \cdot x = 0$.
2. *In a ring,* $1 = 0$ *if and only if there are no other elements.*

A one-element ring is **trivial.**

Some authors might not require a ring to have a multiplicative identity.[4] We require it, so that the next theorem holds. As with a group, so with a ring, an element a determines a singulary operation λ_a on the structure, the operation being given by

$$\lambda_a(x) = ax.$$

Then we have an analogue of Cayley's Theorem (page 65):

[4] For Lang [41, ch. II, §1, p. 83], a ring is as we have defined it. For Hungerford [34, ch. III, §1, p. 115], what we call a ring is a *ring with identity.*

Theorem 58. *For every ring* $(R, 1, \cdot)$, *the function*

$$x \mapsto \lambda_x$$

embeds $(R, 1, \cdot)$ *in* $(\mathrm{End}(R), \mathrm{id}_R, \circ)$.

In a ring, if the multiplication commutes, then the ring is a **commutative ring.** For example, the algebras

$$(\mathbb{Z}, 0, -, +, 1, \cdot), \qquad (\mathbb{Q}, 0, -, +, 1, \cdot), \qquad (\mathbb{R}, 0, -, +, 1, \cdot)$$

are commutative rings. The following is easy to check.

Theorem 59. $(\mathbb{Z}_n, 0, -, +, 1, \cdot)$ *is a commutative ring.*

If R is a sub-ring of a commutative ring S, and $a \in S$, then we denote by

$$R[a]$$

the smallest sub-ring of S that includes R and contains a. Then every nonzero element of $R[a]$ can be written in the form

$$b_0 + b_1 a + \cdots + b_n a^n$$

for some b_i in R, for some n in ω. We may replace a with X, this being, not an element of a particular ring, but an **indeterminate.** Then we obtain the **polynomial ring**

$$R[X],$$

whose elements are **formal sums**

$$b_0 + b_1 X + \cdots + b_n X^n.$$

We can continue this construction, getting rings

$$R[X_0, \ldots, X_{m-1}].$$

In an arbitrary ring, an element with both a left and a right multiplicative inverse can be called simply **invertible;** it is also called a **unit.**

Theorem 60. *In a ring, the units compose a group with respect to multiplication. In particular, a unit has a unique left inverse, which is also a right inverse.*

The group of units of a ring R is denoted by

$$R^{\times}.$$

For example, $\mathbb{Z}^{\times} = \{1, -1\}$. Evidently all two-element groups are isomorphic to this one.

By the theorem, if an element of a ring has both a left inverse and a right inverse, then they are equal. However, possibly an element can have a right inverse, but not a left inverse. We can construct an example by means of Theorem 55. Let G be any nontrivial group. An arbitrary element $(x_n\colon n \in \omega)$ of G^{ω} can be written also as

$$(x_0, x_1, \dots).$$

Then $\operatorname{End}(G^{\omega})$ contains elements f and g given by

$$\begin{aligned} f(x_0, x_1, \dots) &= (x_1, x_2, x_3, x_4, \dots), \\ g(x_0, x_1, \dots) &= (x_0, x_0, x_1, x_2, \dots), \end{aligned}$$

so that

$$\begin{aligned} fg(x_0, x_1, \dots) &= (x_0, x_1, x_2, \dots), \\ gf(x_0, x_1, \dots) &= (x_1, x_1, x_2, \dots). \end{aligned}$$

In particular, g is a right inverse of f, but not a left inverse.

3.1.5. Fields

If R is a commutative ring, and $R^{\times} = R \setminus \{0\}$, then R is called a **field.** For example, $\mathbb{Q}$ and $\mathbb{R}$ are fields. The field $\mathbb{C}$ can be defined as $\mathbb{R} \times \mathbb{R}$ with the appropriate operations. Additional examples are given by Theorem 62 below.

A positive integer n is **prime** if $n \neq 1$ and the only divisors of n in $\mathbb{N}$ are 1 and n. The **greatest common divisor** of two positive integers a and b is just that: the largest of the positive integers that divide both a and b. It can be denoted by

$$\gcd(a, b).$$

This can be found by the **Euclidean algorithm,** used in Propositions VII.1 and 2 of Euclid's *Elements* [18, 17]. The algorithm constructs a sequence $(a_0, a_1, \dots)$, where a_0 is the greater of a and b, and a_1 is the lesser, and for each k in ω, if $a_{k+1} \mid a_k$, then a_{k+2} is undefined, but if $a_{k+1} \nmid a_k$, then a_{k+2} is the remainder on dividing a_k by a_{k+1}, that is, a_{k+2} is the least positive integer r such that

$$a_{k+1} \mid a_k - r;$$

equivalently, a_{k+2} is the least positive integer in the set

$$\{a_k - a_{k+1}x \colon x \in \mathbb{N}\}.$$

Then $a_0 > a_1 > a_2 > \cdots$, so the sequence must terminate, since $\mathbb{N}$ is well-ordered (Theorem 24, page 43).

Theorem 61. *For all positive integers a and b, the last entry of the sequence constructed by the Euclidean algorithm is $\gcd(a, b)$. This is the least positive element of the set*

$$\{ax + by \colon (x, y) \in \mathbb{Z} \times \mathbb{Z}\}.$$

Theorem 62. *The ring $\mathbb{Z}_n$ is a field if and only if n is prime.*

A special case of the theorem is that the trivial ring $\mathbb{Z}_1$ is not a field. If p is prime, then, considered as a field, $\mathbb{Z}_p$ will be denoted by

$$\mathbb{F}_p.$$

3.2. Quotients

3.2.1. Congruence relations

The groups $(\mathbb{Z}_n, 0, -, +)$ and the rings $(\mathbb{Z}_n, 0, -, +, 1, \cdot)$ are instances of a general construction. Suppose $\sim$ is an equivalence relation on a set A, so that it partitions A into equivalence classes

$$\{x \in A \colon x \sim a\};$$

each such class can be denoted by an expression like one of the following:

$$a^{\sim}, \qquad [a], \qquad \overline{a}.$$

Each element of an equivalence class is a **representative** of that class. The **quotient** of A by $\sim$ is the set of equivalence classes of A with respect to $\sim$; this set can be denoted by

$$A/{\sim}.$$

Suppose for some n in ω and some set B, we have a function f from A^n to B. Then there may or may not be a function $\tilde{f}$ from $(A/{\sim})^n$ to B such that the equation

$$\tilde{f}([x_0], \dots, [x_{n-1}]) = f(x_0, \dots, x_{n-1}) \tag{3.3}$$

is an identity. If there is such a function $\tilde{f}$, then it is unique. In this case, the function $\tilde{f}$ is said to be **well-defined** by the given identity (3.3). Note however that there are no "ill-defined" functions. An ill-defined function would be a nonexistent function. The point is that choosing a function f and writing down the equation (3.3) does not automatically give us a function $\tilde{f}$. To know that there is such a function, we must check that

$$a_0 \sim x_0 \land \dots \land a_{n-1} \sim x_{n-1} \Rightarrow f(a_0, \dots, a_{n-1}) = f(x_0, \dots, x_{n-1}).$$

When this does hold (for all a_i), so that $\tilde{f}$ exists as in (3.3), then

$$\tilde{f} \circ \mathrm{p} = f, \tag{3.4}$$

$$\begin{tikzcd}
A^n \arrow[r, "f"] \arrow[d, "\mathrm{p}"'] & B \\
(A/{\sim})^n \arrow[ur, "\tilde{f}"'] &
\end{tikzcd}$$

Figure 3.1.: A well-defined function

where p is the function $(x_0, \dots x_{n-1}) \mapsto ([x_0], \dots, [x_{n-1}])$ from A^n to $(A/{\sim})^n$. Another way to express the equation (3.4) is to say that the diagram in Figure 3.1 **commutes.**

Suppose now $\mathfrak{A}$ is an algebra with universe A. If for all n in ω, for every distinguished n-ary operation f of $\mathfrak{A}$, there is an n-ary operation $\tilde{f}$ on $(A/{\sim})^n$ as given by (3.3), then $\sim$ is a **congruence-relation** or **congruence** on $\mathfrak{A}$. In this case, the $\tilde{f}$ are the distinguished operations of an algebra with universe $A/{\sim}$. This new algebra is the **quotient** of $\mathfrak{A}$ by $\sim$ and can be denoted by

$$\mathfrak{A}/{\sim}.$$

For example, by Theorem 47 on page 64, for each n in $\mathbb{N}$, congruence *modulo* n is a congruence on $(\mathbb{Z}, 0, -, +, 1, \cdot)$. Then the structure $(\mathbb{Z}_n, 0, -, +)$ can be understood as the quotient $(\mathbb{Z}, 0, -, +)/{\sim}$, and $(\mathbb{Z}_n, 0, -, +, 1, \cdot)$ as $(\mathbb{Z}, 0, -, +, 1, \cdot)/{\sim}$. The former quotient is an abelian group by Theorem 49, and the latter quotient is a commutative ring by Theorem 59 on page 73. These theorems are special cases of the next two theorems. In fact the first of these makes verification of Theorem 49 easier.

Theorem 63. *Suppose $\sim$ is a congruence-relation on a semigroup $(G, \cdot)$.*

1. *$(G, \cdot)/{\sim}$ is a semigroup.*
2. *If $(G, \cdot)$ expands to a group, then $\sim$ is a congruence-relation on this group, and the quotient of the group by $\sim$ is a group. If the original group is abelian, then so is the quotient.*

Theorem 64. *Suppose* $(R, 0, -, +, 1, \cdot)$ *is a ring, and* $\sim$ *is a congruence-relation on the reduct* $(R, +, \cdot)$. *Then* $\sim$ *is a congruence-relation on* $(R, 0, -, +, 1, \cdot)$, *and the quotient* $(R, 0, -, +, 1, \cdot)/{\sim}$ *is also a ring. If the original ring is commutative, so is the quotient.*

3.2.2. Normal subgroups of groups

We defined subgroups of symmetry groups on page 61, and of course subgroups of arbitrary groups are defined the same way. A **subgroup** of a group is a subset containing the identity that is closed under multiplication and inversion.

The subset $\mathbb{N}$ of $\mathbb{Q}^+$ contains the identity and is closed under multiplication, but is not closed under inversion, and so it is not a subgroup of $\mathbb{Q}^+$. The subset ω of $\mathbb{Z}$ contains the additive identity and is closed under addition, but is not closed under additive inversion, and so it is not a subgroup of $\mathbb{Z}$.

Theorem 65. *A subset of a group is a subgroup if and only if it is non-empty and closed under the binary operation* $(x, y) \mapsto xy^{-1}$.

If H is a subgroup of G, we write

$$H < G.$$

One could write $H \leqslant G$ instead, if one wanted to reserve the expression $H < G$ for the case where H is a *proper* subgroup of G. We shall not do this.[5]

Theorem 66. *An arbitrary intersection of subgroups is a subgroup.*

Suppose $H < G$. If $a \in G$, let

$$aH = \{ax \colon x \in H\},$$
$$Ha = \{xa \colon x \in H\}.$$

[5] I do think it is useful to reserve the notation $A \subset B$ for the case where A is a proper subset of B, writing $A \subseteq B$ when A is allowed to be equal to B.

Each of the sets aH is a **left coset** of H, and the set $\{xH\colon x \in G\}$ of left cosets is denoted by

$$G/H.$$

Each of the sets Ha is a **right coset** of H, and the set $\{Hx\colon x \in G\}$ of right cosets is denoted by

$$H\backslash G.$$

Note that H itself is both a left and a right coset of itself.

Theorem 67. *Suppose $H < G$. The left cosets of H in G compose a partition of G. Likewise for the right cosets. For each a in G, the map $x \mapsto ax$ is a bijection from H to aH, and $x \mapsto xa$ is a bijection from H to Ha. Thus all cosets are in bijection with one another. The map $xH \mapsto Hx^{-1}$ is a well-defined bijection from G/H to $H\backslash G$.*

Proof. We have $a \in aH$. Suppose $aH \cap bH \neq \varnothing$. Then $ah = bh_1$ for some h and h_1 in H, so that $a = bh_1h^{-1}$, which is in bH. Thus $a \in bH$, and hence $aH \subseteq bH$. By symmetry of the argument, we have also $bH \subseteq aH$, and therefore $aH = bH$. Hence the left cosets compose a partition of G. By symmetry again, the same is true for the right cosets. □

Corollary 67.1. *If $H < G$, then the relation $\sim$ on G defined by*

$$a \sim x \Leftrightarrow aH = xH$$

is an equivalence relation, and

$$G/H = G/{\sim}.$$

Corollary 67.2. *If $H < G$ and $aH = Hb$, then $aH = Ha$.*

Proof. Under the assumption, $a \in Hb$, so $Ha \subseteq Hb$, and therefore $Ha = Hb$. □

Theorem 68. *Suppose $H < G$. The following are equivalent:*

1. *G/H is a group whose multiplication is given by*

$$(xH)(yH) = xyH.$$

2. *Every left coset of H is a right coset.*
3. *$aH = Ha$ for all a in G.*
4. *$a^{-1}Ha = H$ for all a in G.*

Proof. Immediately the last two conditions are equivalent, and they imply the second. The second implies the third, by Corollary 67.2 (p. 79).

Suppose now the first condition holds. For all h in H, since $hH = H$, we have

$$aH = \mathrm{e}\,aH = \mathrm{e}\,HaH = hHaH = haH,$$

hence $a^{-1}haH = H$, so $a^{-1}ha \in H$. Thus $a^{-1}Ha \subseteq H$, so $a^{-1}Ha = H$.

Conversely, if the third condition holds, then $(xH)(yH) = xHHy = xHy = xyH$. In this case, the equivalence relation $\sim$ on G given as in Corollary 67.1 (p. 79) by

$$a \sim x \Leftrightarrow aH = xH$$

is a congruence-relation, and so, by Theorem 63 (p. 77), G/H is a group with respect to the proposed multiplication. □

A subgroup H of G meeting any of these equivalent conditions is called **normal,** and in this case we write

$$H \lhd G.$$

As trivial examples, we have

$$G \lhd G, \qquad\qquad \{\mathrm{e}\} \lhd G.$$

Only slightly less trivially, all subgroups of abelian groups are normal subgroups. If f is a homomorphism from a group G to a group H, then we define

$$\ker(f) = \{x \in G \colon f(x) = \mathrm{e}\},$$

$$\operatorname{im}(f) = \{f(x) \colon x \in G\};$$

these are, respectively, the **kernel** and **image** of the homomorphism f. A homomorphism whose inverse is a well-defined homomorphism is an **isomorphism.**

Theorem 69. *If f is a homomorphism from a group G to a group H, then*

$$\ker(f) \lhd G$$

and there is a well-defined isomorphism

$$x \ker(f) \mapsto f(x)$$

from $G/\ker(f)$ to $\operatorname{im}(f)$.

3.2.3. Ideals of rings

Theorem 70. *Suppose $(R, 1, \cdot)$ is a ring and $A < R$. The group R/A expands to a ring with multiplication given by*

$$(x + A)(y + A) = xy + A$$

if and only if

$$r \in R \;\&\; a \in A \implies ra \in A \;\&\; ar \in A. \tag{3.5}$$

Proof. If R/A does expand to a ring, and $a \in A$, then $a + A$ is 0 in this ring, and hence so are $ra + A$ and $ar + A$, so that (3.5) holds. Conversely, suppose this holds. If $a + A = x + A$ and $b + A = y + A$, then A contains $a - x$ and $b - y$, so A contains also

$$(a - x) \cdot y + a \cdot (b - y),$$

which is $ab - xy$, so $ab + A = xy + A$. □

Under the equivalent conditions of the theorem, A is called an **ideal** of R. We can express (3.5) as

$$RA \subseteq A, \qquad\qquad AR \subseteq A.$$

A homomorphism of rings has the obvious definition. If $(R, 1, \cdot)$ and $(S, 1, \cdot)$ are rings, then a homomorphism from the former to the latter is a homomorphism f from R to S (these considered as groups) such that

$$f(1) = 1, \qquad\qquad f(x) \cdot f(y) = f(x \cdot y).$$

We define the kernel and image of a homomorphism of rings as we do for a homomorphism of groups. Then we have the following analogue of Theorem 69.

Theorem 71. *If f is a homomorphism from a ring R to a ring S, then* $\ker(f)$ *is an ideal of R, and there is a well-defined isomorphism*

$$x + \ker(f) \mapsto f(x)$$

from $R/\ker(f)$ *to* $\operatorname{im}(f)$.

If R is a ring, and A is a subset of R, then there is at least one ideal of R that includes A, namely the **improper ideal** R itself. There is a *smallest* ideal that includes A, by the following.

Theorem 72. *The intersection of a family of ideals of a ring is an ideal.*

Thus, by the terminology to be developed on page 106, the family of ideals of a ring is a *Moore family.* Given a subset A of a ring R, we define

$$(A) = \bigcap\{I \colon I \text{ is an ideal of } R \text{ and } A \subseteq I\}.$$

This is an ideal that includes A and is included in every ideal that includes A. It is the ideal **generated by** A.

3.3. Direct products and sums of commutative rings

Analogously to Theorem 56 on page 70, we have the following.

Theorem 73. *If $\mathscr{R}$ is an indexed family $(R_i\colon i \in \Omega)$ of rings, then $\prod \mathscr{R}$ is a ring under the termwise operations. Each of the coordinate projections π_j on $\prod \mathscr{R}$ is a homomorphism of rings. If S is a ring, and f_j is a homomorphism from S to R_j for each j in Ω, then the map*

$$x \mapsto \bigl(f_i(x)\colon i \in \Omega\bigr)$$

is the unique homomorphism f from S to $\prod \mathscr{R}$ such that, for each j in Ω,

$$\pi_j \circ f = f_j.$$

Suppose $\mathscr{R}$ is an indexed family $(R_i\colon i \in \Omega)$ of commutative rings. If $a \in \prod \mathscr{R}$, we define

$$\operatorname{supp}(a) = \{i \in \Omega\colon a_i \neq 0\};$$

this is the **support** of a. Then a has **finite support** if (obviously) its support is a finite set, that is,

$$|\operatorname{supp}(a)| < \omega.$$

Theorem 74. *The elements having finite support in the direct product of an indexed family of commutative rings compose an ideal of the product ring. That is, if $\mathscr{R}$ is an indexed family of commutative rings, then the subset*

$$\Bigl\{x \in \prod \mathscr{R}\colon |\operatorname{supp}(x)| < \omega\Bigr\}$$

of $\prod \mathscr{R}$ is an ideal.

The ideal of $\prod \mathscr{R}$ given by the theorem is the **direct sum** of $\mathscr{R}$ and can be denoted by one of

$$\bigoplus \mathscr{R}, \qquad\qquad \bigoplus_{i \in \Omega} R_i,$$

the latter assuming $\mathscr{R}$ is $(R_i\colon i \in \Omega)$. In this case, for each j in Ω, there is a function ι_j from R_j to $\bigoplus \mathscr{R}$ that can be given by

$$\pi_i\bigl(\iota_j(x)\bigr) = \begin{cases} x, & \text{if } i = j,\\ 0, & \text{if } i \neq j.\end{cases}$$

This ι_j is the **coordinate injection** of R_j in $\bigoplus \mathscr{R}$.

Theorem 75. *If $\mathscr{R}$ is an indexed family $(R_i \colon i \in \Omega)$ of commutative rings, then each of the coordinate injections ι_j of R_j in $\bigoplus \mathscr{R}$ is an embeddings of rings, and for each a in $\bigoplus \mathscr{R}$,*

$$a = \sum_{i \in \operatorname{supp}(a)} \iota_i\bigl(\pi_i(a)\bigr). \tag{3.6}$$

Since $\iota_i(\pi_i(a)) = 0$ when $i \notin \operatorname{supp}(a)$, it makes sense to write (3.6) in the form

$$a = \sum_{i \in \Omega} \iota_i(\pi_i(a)). \tag{3.7}$$

This should be compared with (3.2) on page 70, namely

$$a = \bigl(\pi_i(a) \colon i \in \Omega\bigr).$$

The latter holds for all a in $\prod \mathscr{R}$; but in (3.7), the sum is defined only when $a \in \bigoplus \mathscr{R}$, that is, only finitely many of the summands are nonzero.

If a is an element of an arbitrary commutative ring, then the ideal generated by $\{a\}$ is denoted by

$$(a)$$

as well as $(\{a\})$. Such an ideal is called a **principal ideal.**

Theorem 76. *If R is a commutative ring and $a \in R$, then*

$$(a) = \{ax \colon x \in R\}.$$

The principal ideal in the last theorem can be denoted also by one of

$$aR, \qquad\qquad Ra.$$

Thus the ring $\mathbb{Z}_n$ (Theorem 59, page 73) can be written as one of

$$\mathbb{Z}/(n), \qquad\qquad \mathbb{Z}/n\mathbb{Z}.$$

There is a homomorphism $k \mapsto \underbrace{1 + \dots + 1}_{k}$ from $\mathbb{Z}$ to R, whose kernel is (n) for some n in ω; in this case n is called the **characteristic** of R.

An ideal is in particular a ring. Thus, if A is a subset of the commutative ring R, we can form an indexed family $(Ra \colon a \in A)$ of commutative rings. Such a family has a direct sum.

Theorem 77. *If R is a commutative ring and A is a subset, then*

$$\begin{aligned}(A) &= \left\{\sum_{a \in A} x_a \colon x \in \bigoplus_{a \in A} Ra\right\} \\ &= \left\{\sum_{a \in A} x_a a \colon x \in \bigoplus_{a \in A} R\right\}.\end{aligned}$$

That is, the ideal (A) consists of the **R-linear combinations** of elements of A. The ideal can be denoted by one of

$$\sum_{a \in A} Ra, \qquad \sum_{a \in A} (a).$$

If $A = \{a_i \colon i < n\}$, then (A) can be written as one of

$$(a_0, \dots, a_{n-1}), \qquad Ra_0 + \dots + Ra_{n-1}, \qquad (a_0) + \dots + (a_{n-1}).$$

Such an ideal is said to be **finitely generated.**

3.4. Ultraproducts of fields

The improper ideal of a commutative ring R is the principal ideal

$$(1).$$

The subset $\{0\}$ of R is the **zero ideal** and can be considered[6] as the principal ideal

$$(0).$$

A **proper ideal** of a ring is an ideal that is not improper, that is, is not the whole ring. An ideal of a commutative ring is called a **maximal ideal** if it is a proper ideal, but is not properly included in a proper ideal. (Thus the improper ideal is not a maximal ideal.) An ideal I of the commutative ring R is maximal just in case, for every ideal J of R,

$$I \subset J \iff J = R.$$

Theorem 78. *Let R be a commutative ring.*

1. *The ideal (0) of R is maximal if and only if R is a field.*
2. *An ideal I of R is maximal if and only if the quotient R/I is a field.*

Proof. If R is a field and $(0) \subset I$, then $I \smallsetminus (0)$ contains some a, and then $a^{-1} \cdot a \in I$, so $I = R$. Conversely, if (0) is maximal, then for all a in $R \smallsetminus (0)$ we have $(a) = (1)$, so a is invertible.

Every ideal of R/I is J/I for some subgroup J of R. Moreover, this J must be an ideal of R. In this case, J is maximal if and only if J/I is a maximal ideal of R/I. □

Suppose $\mathscr{K}$ is an indexed family $(K_i \colon i \in \Omega)$ of fields. For example, each K_i might be $\mathbb{R}$, or each K_i might be a different finite field. Suppose M is a maximal ideal of the ring $\prod \mathscr{K}$. By the last theorem, the quotient $\prod \mathscr{K}/M$ is a field. Such a field is called an **ultraproduct** of the indexed family $\mathscr{K}$. The ultraproduct is called **principal** or **nonprincipal,** according as M itself is principal or nonprincipal.

[6] Since every ideal contains 0, the zero ideal is also the ideal $(\varnothing)$ generated by the empty set. However, when we write this ideal as (0), we mean by 0 the zero element of the ring, rather than the first von Neumann natural number (page 16), which is the empty set. There is no need to include 0 in the generating set of any ideal. Nonetheless, there is no harm in including it, and we do want to consider the zero ideal as being a principal ideal.

If I is an arbitrary ideal of $\prod\mathscr{K}$, we define

$$\operatorname{supp}[I]=\{\operatorname{supp}(x)\colon x\in I\}.$$

Theorem 79. *Let $\mathscr{K}$ be an indexed family of fields. If I is an ideal of $\prod\mathscr{K}$, then*

$$I=\left\{x\in\prod\mathscr{K}\colon\ \operatorname{supp}(x)\in\operatorname{supp}[I]\right\}.$$

Proof. Obviously $I\subseteq\{x\in\prod\mathscr{K}\colon\ \operatorname{supp}(x)\in\operatorname{supp}[I]\}$. For the reverse inclusion, if $a\in I$ and $\operatorname{supp}(b)=\operatorname{supp}(a)$ or even $\operatorname{supp}(b)\subseteq\operatorname{supp}(a)$, then $b\in I$, since $b=ca$, where

$$c_i=\begin{cases}b_i/a_i, & \text{if } i\in\operatorname{supp}(b),\\ 0, & \text{if } i\notin\operatorname{supp}(b);\end{cases}$$

this shows $\{x\in\prod\mathscr{K}\colon\ \operatorname{supp}(x)\in\operatorname{supp}[I]\}\subseteq I$. □

If $A\subseteq\Omega$, we define the element χ_A of ${\mathbb{F}_2}^{\Omega}$ by

$$\chi_A(i)=\begin{cases}1, & \text{if } i\in A,\\ 0, & \text{if } i\in\Omega\smallsetminus A.\end{cases}$$

If A and B are both subsets of Ω, we define

$$A\mathbin{\triangle}B=(A\smallsetminus B)\cup(B\smallsetminus A);$$

this is the **symmetric difference** of A and B. See Figure 3.2.

Theorem 80. *For every set Ω, the map*

$$A\mapsto\chi_A$$

from $\mathscr{P}(\Omega)$ to ${\mathbb{F}_2}^{\Omega}$ is a bijection, whose inverse is

$$x\mapsto\operatorname{supp}(x).$$

Moreover,

$$\chi_{A\triangle B}=\chi_A+\chi_B,\qquad\qquad \chi_{A\cap B}=\chi_A\cdot\chi_B.$$

In $\mathscr{P}(\Omega)$,

$$A\cup B=A\mathbin{\triangle}B\mathbin{\triangle}(A\cap B).$$

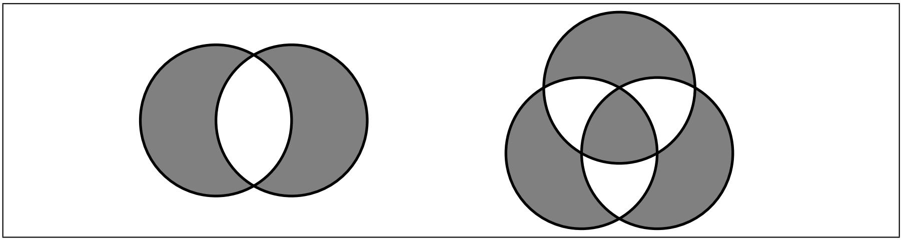

Figure 3.2.: Symmetric differences of two sets and three sets

Corollary 80.1. *$\mathscr{P}(\Omega)$ is a ring in which sums are symmetric differences and products are intersections. Moreover, an ideal of $\mathscr{P}(\Omega)$ is just a subset I such that*

$$\begin{gathered}\varnothing \in I,\\ X \in I \ \&\ Y \in I \implies X \cup Y \in I,\\ Y \in I \ \&\ X \subseteq Y \implies Y \in I.\end{gathered}$$

See Figure 3.3. There are two important examples of ideals of the ring $\mathscr{P}(\Omega)$:

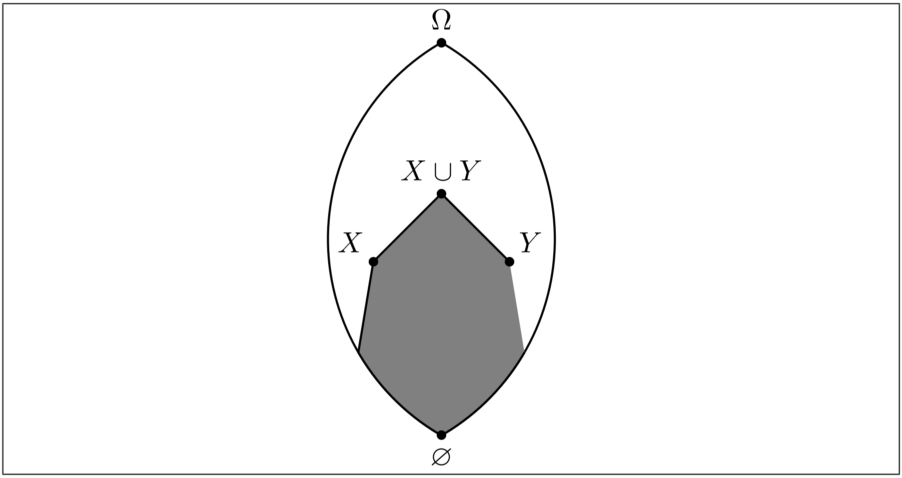

Figure 3.3.: An ideal of $\mathscr{P}(\Omega)$

1. If $A \subseteq \Omega$, then $\mathscr{P}(A)$ is the principal ideal of $\mathscr{P}(\Omega)$ generated by A.
2. The set of finite subsets of Ω is an ideal of $\mathscr{P}(\Omega)$, called the **Fréchet ideal** of $\mathscr{P}(\Omega)$; this ideal can be denoted by

$$\mathscr{P}_\omega(\Omega).$$

If $\mathscr{K}$ is again the indexed family $(K_i\colon i \in \Omega)$ of fields, we want to show that the map $I \mapsto \operatorname{supp}[I]$ is a bijection from the family of ideals of $\prod \mathscr{K}$ to the family of ideals of $\mathscr{P}(\Omega)$.

The underlying set of the field $\mathbb{F}_2$ can be considered as the subset $\{0,1\}$ of each field K_i. The field $\mathbb{F}_2$ is not a *subfield* of K_i unless K has characteristic 2; but it can be understood as is a multiplicative submonoid. Hence ${\mathbb{F}_2}^\Omega$ is a multiplicative submonoid of $\prod \mathscr{K}$. For each subset A of Ω, the function χ_A can be understood as belonging to $\prod \mathscr{K}$.

If x belongs to an arbitrary field, we define

$$x^* = \begin{cases} 1/x, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

If now x belongs to $\prod \mathscr{K}$, we can let

$$x^* = ({x_i}^*\colon i \in \Omega). \tag{3.8}$$

Then easily

$$\operatorname{supp}(x^*) = \operatorname{supp}(x).$$

Theorem 81. *Let $\mathscr{K}$ be again the indexed family $(K_i\colon i \in \Omega)$ of fields. The map $I \mapsto \{x^*x\colon x \in I\}$ is a bijection from the family of ideals of $\prod \mathscr{K}$ to the family of ideals of ${\mathbb{F}_2}^\Omega$.*

Proof. If $x \in \prod \mathscr{K}$, then

$$x^*x = \chi_{\operatorname{supp}(x)}.$$

Hence we have the commutative diagram in Figure 3.4. If I is an ideal

$$\begin{tikzcd}
 & \mathbb{F}_2{}^{\Omega} \arrow[dd, bend left, "x\mapsto\operatorname{supp}(x)"] \\
\prod\mathscr{K} \arrow[ur, "x\mapsto x^*x"] \arrow[dr, "x\mapsto\operatorname{supp}(x)"'] & \\
 & \mathscr{P}(\Omega) \arrow[uu, bend left, "A\mapsto\chi_A"]
\end{tikzcd}$$

Figure 3.4.: Products of fields

of $\prod\mathscr{K}$, let us denote $\{x^*x\colon x\in I\}$ by I^*. Then this is an ideal of $\mathbb{F}_2{}^{\Omega}$ if and only if $\operatorname{supp}[I]$ is an ideal of $\mathscr{P}(\Omega)$. Evidently

$$I^* = \mathbb{F}_2{}^{\Omega}\cap I.$$

Since $\mathbb{F}_2{}^{\Omega}$ is a submonoid of $\prod\mathscr{K}$, we have that I^* is an ideal of $\mathbb{F}_2{}^{\Omega}$ if and only if it is closed under addition, or equivalently $\operatorname{supp}[I]$ is closed under symmetric differences. But $\operatorname{supp}[I]$ is so closed, since in $\prod\mathscr{K}$ we have

$$\operatorname{supp}(x)\mathbin{\triangle}\operatorname{supp}(y)\subseteq\operatorname{supp}(x+y)$$

and so

$$\operatorname{supp}(x)\mathbin{\triangle}\operatorname{supp}(y)=\operatorname{supp}((x+y)\cdot\chi_{\operatorname{supp}(x)\triangle\operatorname{supp}(y)}).$$

So I^* is indeed an ideal of $\mathbb{F}_2{}^{\Omega}$. Since $(I^*)=I$, the map $I\mapsto I^*$ is injective. Suppose J is an arbitrary ideal of $\mathbb{F}_2{}^{\Omega}$, and let $I=\{x\in\prod\mathscr{K}: x^*x\in J\}$. Evidently this is nonempty. If it contains x and y, then it contains $x-y$, since

$$\operatorname{supp}(x-y)\subseteq\operatorname{supp}(x)\cup\operatorname{supp}(y).$$

Also, if $z\in\prod\mathscr{K}$, then I contains zx, since

$$\operatorname{supp}(zx)\subseteq\operatorname{supp}(x).$$

Thus I is an ideal of $\prod \mathscr{K}$, and $I^* = J$. □

Under the one-to-one correspondence of the theorem,

1) a principal ideal $\mathscr{P}(A)$ of $\mathscr{P}(\Omega)$ corresponds to the obvious image of $\prod_{i\in A} K_i$ in $\prod \mathscr{K}$;
2) the Fréchet ideal of $\mathscr{P}(\Omega)$ corresponds to the ideal $\bigoplus \mathscr{K}$ of $\prod \mathscr{K}$.

Theorem 82. *Let $\mathscr{K}$ be an indexed family $(K_i \colon i \in \Omega)$ of fields.*

1. *If $j \in \Omega$, then*
$$\ker(\pi_i) = (\iota_j(1)) = (\chi_{\{j\}}).$$
This is a maximal ideal of $\prod \mathscr{K}$, and every principal maximal ideal of $\prod \mathscr{K}$ is of this form. Thus every principal ultraproduct of $\mathscr{K}$ is isomorphic to one of the K_j.
2. *Every nonprincipal maximal ideal of $\prod \mathscr{K}$ includes the ideal $\bigoplus \mathscr{K}$.*

We are going to be interested in nonprincipal ultraproducts.

4. Products of fields

The main results of this chapter are the following.

1. All maximal ideals of a commutative ring are *prime ideals* (Corollary 85.1, page 94).
2. Every proper ideal of a commutative ring is included in a maximal ideal (Theorem 99, page 102), by *Zorn's Lemma* (page 99).
3. The set $\operatorname{Spec}(R)$ of prime ideals of a commutative ring R is a *compact Kolmogorov topological space* (Theorem 113, page 114) whose closed sets are in one-to-one correspondence with the *radical ideals* of R (Corollary 117.1, page 118).
4. A proper ideal of a commutative ring R is radical if and only if R/I is *reduced* (Theorem 114, page 116).
5. A commutative ring is *regular* if and only if it is reduced and all of its prime ideals are maximal (Theorem 127, page 127).
6. A commutative ring is regular if and only if it embeds, as a regular ring, in a product of fields (Theorem 128, page 128).
7. The Tychonoff Theorem (page 131) is equivalent to the Axiom of Choice (Theorem 132, page 131).

4.1. Prime ideals

4.1.1. Properties

The following is Proposition VII.30 of Euclid's *Elements* [18, 17]. It will motivate the definition of *prime ideal* below.

Theorem 83 (Euclid's Lemma). *If p is a prime number, then for all integers a and b,*

$$p \mid ab \ \& \ p \nmid a \implies p \mid b.$$

Proof. Given that $p \nmid a$, we know that $\gcd(p, a) = 1$, so we can solve

$ax + py = 1$ by Theorem 61 (page 75). We obtain

$$abx + pby = b,$$

so if $p \mid ab$, then, since immediately $p \mid pby$, we must have $p \mid b$. □

Noting that, in $\mathbb{Z}$,

$$a \mid b \iff b \in (a),$$

we refer to an ideal $\mathfrak{p}$ of a commutative ring R as **prime** if $\mathfrak{p}$ is a *proper* ideal of R and, for all a and b in R,

$$ab \in \mathfrak{p} \ \& \ a \notin \mathfrak{p} \implies b \in \mathfrak{p}. \tag{4.1}$$

(See Appendix A, page 255, for Fraktur letters like $\mathfrak{p}$.) Then the prime ideals of $\mathbb{Z}$ are precisely the ideals (0) and (p), where p is prime. A trivial ring has no prime ideal.

We shall establish an analogue of Theorem 78 (page 86), with prime ideals in place of maximal ideals. A **zero-divisor** of the commutative ring R is a nonzero element b such that the equation

$$bx = 0$$

has a nonzero solution in R. So zero-divisors are not units. For example, if $m > 1$ and $n > 1$, then $m + (mn)$ and $n + (mn)$ are zero-divisors in $\mathbb{Z}_{mn}$. The unique element of the trivial ring $\mathbb{Z}_1$ is a unit, but not a zero-divisor.

Theorem 84. *In a non-trivial commutative ring, the zero-divisors, together with* 0 *itself, compose a prime ideal.*

A commutative ring is an **integral domain** if it has no zero-divisors and $1 \neq 0$. If $n \in \mathbb{N}$, the ring $\mathbb{Z}_n$ is an integral domain if and only if n is prime.[1] Hence the characteristic of an integral domain must be prime or 0. Fields are integral domains, but $\mathbb{Z}$ is an integral domain that is not a field. We now establish an analogue of Theorem 78 (page 86).

[1] Lang refers to integral domains as *entire* rings [41, p. 91]. It would appear that integral domains were originally defined as subgroups of $\mathbb{C}$ that are closed under multiplication *and* that include the integers [11, p. 47].

Theorem 85. *Let R be a commutative ring.*

1. *The ideal* (0) *of R is prime if and only if R is an integral domain.*
2. *An ideal I of R is prime if and only if the quotient R/I is an integral domain.*

Proof. 1. This is immediate from the definitions of integral domain and prime ideal, once we note that $x \in (0)$ means $x = 0$.

2. The ideal (0) of R/I is $\{I\}$, and

$$(a + I)(b + I) = I \iff ab \in I. \qquad \square$$

We might summarize Theorems 78 and 85 thus:

prime ideal : integral domain :: maximal ideal : field.

Since fields are integral domains, we have:

Corollary 85.1. *Maximal ideals are prime.*

The converse of the corollary fails easily, since (0) is a prime but non-maximal ideal of $\mathbb{Z}$. However, every prime ideal of $\mathbb{Z}$ other than (0) is maximal. The same is true for $\mathbb{Q}[X]$ (see Theorem 197, page 225), but not for $\mathbb{Q}[X, Y]$, which has the prime but non-maximal ideal (X).

In some commutative rings, *every* prime ideal is maximal. This is so for fields, since their only proper ideals are (0). We are going to show that all prime ideals of direct products of indexed families of fields are maximal. Thus the quotient of such a product by an arbitrary prime ideal will be an ultraproduct.

We first consider a special case: the direct power ${\mathbb{F}_2}^{\Omega}$. By Theorem 80 (page 87), we can consider $\mathscr{P}(\Omega)$ as a ring in which the sum of two sets is their symmetric difference, and the product of two sets is their intersection; and this ring is isomorphic to ${\mathbb{F}_2}^{\Omega}$.

The rings $\mathscr{P}(\Omega)$ and ${\mathbb{F}_2}^{\Omega}$ are examples of *Boolean rings.* An arbitrary nontrivial ring is called **Boolean** if it satisfies the identity

$$x^2 = x.$$

Immediately from the definition, every sub-ring of a Boolean ring is a Boolean ring.

Theorem 86. *Every Boolean ring is commutative and satisfies the equivalent identities*

$$2x = 0, \qquad\qquad -x = x.$$

Proof. In a Boolean ring,

$$\begin{aligned} x + y = (x+y)^2 &= x^2 + xy + yx + y^2 \\ &= x + xy + yx + y, \end{aligned}$$

so $0 = xy + yx$. Replacing y with x gives $0 = 2x^2 = 2x$. Hence generally $yx = -xy = xy$. □

Theorem 87. *Let I be an ideal of a Boolean ring R.*

1. *If I is prime, then I is maximal.*
2. *If I is maximal, then*
$$R/I \cong \mathbb{F}_2.$$
3. *I is maximal if and only if*
$$x \in R \smallsetminus I \iff 1 + x \in I.$$

Proof. In a Boolean ring, by the last theorem,

$$x \cdot (1 + x) = x + x^2 = x + x = 0,$$

and also

$$x \in \{0,1\} \iff 1 + x \in \{0,1\}.$$

Therefore every x is a zero-divisor unless x is 0 or 1. Thus there are no Boolean integral domains besides $\{0,1\}$, which is the field $\mathbb{F}_2$. □

4.1.2. Existence

So far, we do not know whether an arbitrary nontrivial commutative ring has a maximal or even a prime ideal. However, settling the question is easy in one special case.

For an arbitrary set Ω, a subset C of $\mathscr{P}(\Omega)$ is called a **chain** if proper inclusion is also a total relation on C, so that C is linearly ordered by proper inclusion (see Theorem 21, page 42).

Lemma 10.

1. *The union of a chain of ideals of a commutative ring is an ideal.*
2. *The union of a chain of proper ideals of a commutative ring is a proper ideal.*

Theorem 88. *Every countable commutative ring has a maximal ideal.*

Proof. Suppose R is a countable nontrivial commutative ring. This means there is a function $k \mapsto a_k$ from ω onto R. Using the Recursion Theorem (page 37), we define recursively a function $k \mapsto I_k$ from ω into the set of ideals of R. Let $I_0 = (0)$, which is a proper ideal of R. If $(I_k \cup \{a_k\})$ is a proper ideal of R, we let I_{k+1} be this ideal; otherwise $I_{k+1} = I_k$. By induction, each I_k is a proper ideal of R. Let

$$J = \bigcup_{k \in \omega} I_k.$$

By the lemma, J is a proper ideal of R Moreover, every element of $R \smallsetminus J$ is a_k for some k, and then $a_k \notin I_{k+1}$, so $(I_k \cup \{a_k\})$ must be the improper ideal. Therefore $(J \cup \{a_k\})$ is improper. Thus J is a maximal ideal of R. □

One way that countable rings arise is as follows. Let S be a commutative ring. Then the additive subgroup of S generated by 1 is actually a sub-ring of S. This sub-ring is the **prime ring** of S. It is the image in S of $\mathbb{Z}$ under the homomorphism $k \mapsto \underbrace{1 + \cdots + 1}_{k}$ mentioned also on page 85, and so it is isomorphic either to $\mathbb{Z}$ itself or to a quotient $\mathbb{Z}_n$.

Suppose R is the prime ring of S. If $a \in S$, we defined the notation $R[a]$ on page 73: it stands for the smallest sub-ring of S that includes R and contains a. If $(a_k \colon k \in \omega)$ is an indexed family of elements of S, we define the sub-rings

$$R[a_0, \dots, a_{n-1}]$$

of S recursively, in the obvious way: The ring is R if $n = 0$, and also

$$R[a_0, \dots, a_k] = \bigl(R[a_0, \dots, a_{k-1}]\bigr)[a_k].$$

The rings that can be written in this form are called **finitely generated.**

Note that being finitely generated has different meanings for commutative rings and ideals. (See page 85.) As an improper ideal, every commutative ring can be written as (1) and is thus finitely generated as an ideal. But a commutative ring as such need not be finitely generated: an example is $\mathbb{Q}$.

Theorem 89. *Every finitely generated nontrivial commutative ring is countable and therefore has a maximal ideal.*

We shall adapt the proof of Theorem 88 to rings whose underlying sets are well-ordered. We need a generalization of the Recursion Theorem.

Theorem 90 (Transfinite Recursion). *For all sets A, for all ordinals α, for all functions f from $\bigcup\{A^\beta \colon \beta < \alpha\}$ to A, there is a unique element*

$$(a_\beta \colon \beta < \alpha)$$

of A^α such that, for all β in α,

$$f(a_\gamma \colon \gamma < \beta) = a_\beta.$$

Proof. We first prove uniqueness. Suppose, if possible, $(a'_\beta \colon \beta < \alpha)$ is another element of A^α as desired, and let β be minimal such that $a_\beta \neq a'_\beta$. Then

$$(a_\gamma \colon \gamma < \beta) = (a'_\gamma \colon \gamma < \beta),$$

so by definition $a_\beta = a'_\beta$. We now prove existence. If the theorem fails for some α, let α be minimal such that it fails. Say $f \colon \bigcup\{A^\beta \colon \beta < \alpha\} \to A$. By hypothesis, for each β in α, there is a unique element $(a_\gamma \colon \gamma < \beta)$ of A^β such that, for all γ in β,

$$f(a_\delta \colon \delta < \gamma) = a_\gamma.$$

As before, a_γ is independent of the choice of β such that $\gamma < \beta < \alpha$. Then for all β in α we are free to define

$$a_\beta = f(a_\gamma \colon \gamma < \beta).$$

Then the element $(a_\beta\colon \beta<\alpha)$ of A^α shows that the theorem does not fail for α. □

Our proof used the method of the **minimal counterexample:** we showed that there could not be such a counterexample. The Transfinite Recursion Theorem is used for example to show that there is a bijection, denoted by

$$\alpha\mapsto\aleph_\alpha,$$

from the class **ON** of ordinals to the class $\mathbf{CN}\smallsetminus\omega$ of infinite cardinals: $\aleph_\alpha$ is the least infinite cardinal that is greater than all of the cardinals in $\{\aleph_\beta\colon \beta<\alpha\}$. (One must show that such cardinals exist.) The *Continuum Hypothesis* is that $|\mathbb{R}|=\aleph_1$, but we shall make no use of this.

Theorem 91. *Every nontrivial commutative ring with a cardinality has a maximal ideal.*

Proof. Let R be a nontrivial commutative ring, and suppose $\alpha\mapsto a_\alpha$ is a surjection from a cardinal κ onto R. If $\alpha<\kappa$, and a function $\beta\mapsto I_\beta$ on α has been defined whose range is a chain of proper ideals of R, we define I_α to be $\left(\bigcup_{\beta<\alpha}I_\beta\cup\{a_\alpha\}\right)$, if this is a proper ideal of R, and otherwise $I_\alpha=\bigcup_{\beta<\alpha}I_\beta$. Then $\bigcup_{\alpha<\kappa}I_\kappa$ is a maximal ideal of R. □

4.1.3. Zorn's Lemma

We want to show that the last theorem applies to every ring, so that every nontrivial ring has a maximal ideal. Doing this will be our first use of the Axiom of Choice; and here as always, we shall make this use explicit.

AC **Theorem 92** (Well Ordering)**.** *By the Axiom of Choice, every set has a cardinality.*

Proof. Given a set A, we define

$$A^*=\bigl\{\{X\}\times X\colon X\in\mathscr{P}(A)\smallsetminus\{\varnothing\}\bigr\}.$$

By the Axiom of Choice, there is a set that contains exactly one element of each element of A^*. Such a set is a function g from $\mathscr{P}(A)\smallsetminus\{\varnothing\}$ to A such that $f(X) \in X$ for each nonempty subset of X. Now say $c \notin A$. Given an ordinal α, we define a function from α to $A \cup \{c\}$ by letting

$$f_\alpha(\beta) = g(A \smallsetminus \{f_\alpha(\gamma) \colon \gamma < \beta\}),$$

if possible; otherwise, $f_\alpha(\beta) = c$. If $\beta < \alpha$, then $f_\beta \subset f_\alpha$. Now let β be the least α such that c is in the range of f_α. (Such α must exist; otherwise **ON** embeds in A.) Then β must be γ' for some γ, and then f_γ is a bijection from γ to A. □

Not only *can* we use the Axiom of Choice to prove the foregoing theorem, but we *must* use it, or something equivalent to it:

Theorem 93. *The Well Ordering Theorem implies the Axiom of Choice.*

Proof. Suppose every set has a cardinality, and A is a set of nonempty, pairwise-disjoint sets. Let $\alpha \mapsto a_\alpha$ be a bijection from some cardinal κ to $\bigcup A$, and let B contain those a_α such that, for some X in A, α is the least β such that $a_\beta \in X$. □

For algebraic results that logically require the Axiom of Choice, it may be more convenient to use this in the form of *Zorn's Lemma.* Suppose Ω is a set and $A \subseteq \mathscr{P}(\Omega)$. Then proper inclusion ($\subset$) is a transitive irreflexive relation on A and on each of its subsets (see Theorems 19 and 20, page 41). Suppose $C \subseteq A$. An **upper bound** of C is a set that includes each element of C. In particular, $\bigcup C$ is an upper bound, and every upper bound includes this union. A **maximal element** of A is an element that is not properly included in any other element.

Theorem 94 (Zorn's Lemma). *By the Axiom of Choice, for all sets Ω, for all subsets A of $\mathscr{P}(\Omega)$, if A contains an upper bound for each of its chains, then A has a maximal element.*[2] **AC**

[2] In 1935, Zorn [64] presented this result for the case where the upper bounds

Proof. By the Axiom of Choice, there is a bijection $\alpha \mapsto B_\alpha$ from some cardinal κ to A. By the Recursion Theorem, there is a function $\alpha \mapsto C_\alpha$ from κ to A such that, for all α in κ, if $\{C_\beta \colon \beta < \alpha\}$ is a chain, and if γ is minimal such that B_γ is an upper bound of this chain, then

$$C_\alpha = \begin{cases} B_\gamma, & \text{if } B_\gamma \nsubseteq B_\alpha, \\ B_\alpha, & \text{if } B_\gamma \subseteq B_\alpha; \end{cases}$$

in particular, $\{C_\beta \colon \beta \leqslant \alpha\}$ is a chain. If $\{C_\beta \colon \beta < \alpha\}$ is *not* a chain, then we can define $C_\alpha = B_0$. But we never have to do this: for all α in κ, the set $\{C_\beta \colon \beta < \alpha\}$ *is* a chain, since there can be no minimal counterexample to this assertion. Indeed, if α is minimal such that $\{C_\beta \colon \beta < \alpha\}$ is not a chain, there must be β and γ in α such that $\gamma < \beta$ and neither of C_β and C_γ includes the other. But by assumption $\{C_\delta \colon \delta < \beta\}$ is a chain, and so by definition $\{C_\delta \colon \delta \leqslant \beta\}$ is a chain, and in particular one of C_β and C_γ must include the other.

By a similar argument, $\{C_\alpha \colon \alpha < \kappa\}$ is a chain, so it has an upper bound D in A. Suppose for some α we have $D \subseteq B_\alpha$. Then $C_\alpha = B_\alpha$, since otherwise, by definition, $C_\alpha = B_\gamma$ for some γ such that $B_\gamma \nsubseteq B_\alpha$: in this case $C_\alpha \nsubseteq B_\alpha$, so $C_\alpha \nsubseteq D$, which is absurd. Thus $C_\alpha = B_\alpha$, and hence $B_\alpha \subseteq D$, so $D = B_\alpha$. Therefore D is a maximal element of A. □

We sometimes want to use Zorn's Lemma in a more general form. If $<$ is an arbitrary ordering of a set A, a **chain** of $(A, <)$ is a subset of A that is linearly ordered by $<$. If $C \subseteq A$, an **upper bound** of C (with respect to $<$) in A is an element a of A such that, for all x in C, $x \leqslant a$. A **maximal element** of A (with respect to $<$) is an element b such that, for all x in A, if $b \leqslant x$, then $b = x$.

of the chains are actually the unions of the chains. He called the conclusion the "maximum principle" and suggested that using it would make proofs more algebraic than when the Well-Ordering Theorem is used. Zorn promised to prove the converse in a later paper, which would imply the full equivalence of the maximum principle and the Axiom of Choice; but it seems such a paper never appeared. Earlier, in 1922, Kuratowski [39, (42), p. 89] proved "Zorn's Lemma" for the case where the chains in question are well-ordered.

Corollary 94.1. *By the Axiom of Choice, an order whose every chain has an upper bound has a maximal element.* **AC**

Proof. Given an order $(A,<)$, for each b in A we let

$$(b)=\{x\in A\colon x\leqslant b\}.$$

Now let

$$\mathscr{A}=\{(x)\colon x\in A\}.$$

Then $x\mapsto(x)$ is an isomorphism from $(A,<)$ to $(\mathscr{A},\subset)$; so since the claim holds for the latter structure, it holds for the former. □

We now have easily:

Theorem 95 (Maximal Ideal). *By Zorn's Lemma, every nontrivial commutative ring has a maximal ideal.* **AC**

Proof. The family of proper ideals of a nontrivial commutative ring has an upper bound (namely the union) for each of its chains. □

Theorem 96. *The Maximal Ideal Theorem implies the Axiom of Choice.*

Proof. The proof is given in Rubin and Rubin [50, p. 113], where it is attributed to Hodges, "Krull implies Zorn" (J. London Math. Soc. **19** (1979), 285–287). □

Then the following statements are equivalent:

- the Axiom of Choice;
- the Well Ordering Theorem;
- Zorn's Lemma;
- the Maximal Ideal Theorem.

By Corollary 85.1 (page 94), we obtain the following.

Theorem 97 (Prime Ideal). *By the Maximal Ideal Theorem, every nontrivial commutative ring has a prime ideal.* **AC**

Recall from Theorem 87 that all prime ideals of Boolean rings are maximal.

Theorem 98 (Boolean Prime Ideal). *By the Prime Ideal Theorem, every Boolean ring has a maximal ideal.* **PI**

We shall show later that the Boolean Prime Ideal Theorem implies the Prime Ideal Theorem. However, these theorems do *not* imply the Maximal Ideal Theorem.[3] So we are going to be careful about which theorems need the full Axiom of Choice (or one of its equivalent forms) and which need only the Prime Ideal Theorem. For example, we have the following.

Theorem 99. *Suppose I is a proper ideal of a commutative ring R.*

1. *By the Maximal Ideal Theorem, I is included in a maximal ideal of R.*
2. *By the Prime Ideal Theorem, I is included in a prime ideal of R.*

Proof. By the Maximal Ideal Theorem, R/I has a maximal ideal M. Then $\{x \in R\colon x + I \in M\}$ is a maximal ideal of R. Similarly in the prime case. □

4.2. Determinacy

This section is about why the Axiom of Choice is not "obviously" or "intuitively" correct. The axiom contradicts another set-theoretic axiom that might be considered "obviously" or "intuitively" correct. That axiom is the Axiom of Determinacy, according to which, in certain *games* of infinite length, one of the players always has a winning strategy.

We consider games with two players. Hodges [32] calls these players $\forall$ and $\exists$, after Abelard and Eloise; but I propose to call them simply 0 and 1, for notational purposes. A **game** that 0 and 1 can play is determined by a partition $A_0 \amalg A_1$ of the set ${}^{\omega}2$ of binary sequences on ω. A particular **play** of the game can be analyzed as a sequence of **rounds,** indexed by ω. In round m, player 0 chooses an element

[3] See the discussion in Hodges [31, pp. 272f.] or Rubin and Rubin [50, p. 99]. The latter comprehensive reference does not however mention that the Prime Ideal Theorem is implied by the Boolean Prime Ideal Theorem.

a_{2m} of 2; this is the **move** of 0 in this round. Then player 1 moves by choosing an element a_{2m+1} of 2. The play itself is then the sequence $(a_n\colon n \in \omega)$ or a, which is an element of ${}^{\omega}2$. The play is **won** by that player e such that $a \in A_e$; and then player $1 - e$ has **lost.**

Each player e may use a **strategy,** namely a function f_e from $\bigcup_{m\in\omega} {}^{m+e}2$ to 2. (So f_0 assigns an element of 2 to each finite binary sequence; f_1 does this to every *nonempty* finite binary sequence.) If both f_0 and f_1 are chosen, then a play is determined, namely the sequence $(a_n\colon n \in \omega)$ given by

$$a_{2m} = f_0(a_1, a_3, \dots, a_{2m-1}), \qquad a_{2m+1} = f_1(a_0, a_2, \dots, a_{2m}),$$

or simply by

$$a_{2m+e} = f_e(a_{1-e}, a_{3-e}, \dots, a_{2m-1+e}).$$

That is, f_e determines the move of player e from the previous moves by the *other* player. The player's own previous moves need not be formally considered, since they themselves were already determined by the player's strategy and the other player's previous moves.

Suppose player $1 - e$ has chosen strategy f_{1-e}. For every b in ${}^{\omega}2$, player e might choose a strategy f_e that is constant on each set ${}^{m+e}2$, having the value b_m there. The resulting play will be a, where

$$a_{2m+1-e} = f_{1-e}(b_0, b_1, \dots, b_{m-e}), \qquad a_{2m+e} = b_m.$$

This shows that, for every choice of f_{1-e}, there are continuum-many plays that can result if player $1 - e$ uses this strategy.

If, using a strategy f_e, player e wins all plays of a game, then f_e is a **winning** strategy for that game. The game is **determined** if one of the players has a winning strategy. The **Axiom of Determinacy** is that in every game, one of the players has a winning strategy: in other words, for every choice of the A_e, one of the following sentences of infinitary logic is true:

$$\exists x_0\ \forall x_1\ \exists x_2\ \cdots\ (x_0, x_1, x_2, \dots) \in A_0,$$

$$\forall x_0\ \exists x_1\ \forall x_2\ \cdots\ (x_0, x_1, x_2, \dots) \in A_1.$$

However, this Axiom is false under the assumption of the Axiom of Choice, or more precisely under the assumption that the Continuum can be well-ordered, so that there is a least ordinal, called 2^ω, whose cardinality is that of ${}^{\omega}2$.

Indeed, every ordinal is $\alpha + n$ for some unique limit ordinal α and finite ordinal n. Then $\alpha + n$ is even or odd, according as n is even or odd. Assuming the Axiom of Choice, we can list all possible strategies as $(f^\alpha \colon \alpha < 2^\omega)$, where f^α will be a strategy for e if and only if $\alpha + e$ is even.

We shall now define a list $(a^\alpha \colon \alpha < 2^\omega)$ of possible plays (that is, elements of ${}^{\omega}2$) so that,

- for all α, if $\alpha + e$ is even, then e can use strategy f^α for the play a^α; that is, for all m in ω,

$$a^\alpha_{2m+e} = f^\alpha(a^\alpha_{1-e}, a^\alpha_{3-e}, \dots, a^\alpha_{2m-1+e});$$

- $a^\alpha \neq a^\beta$ for all distinct α and β such that $\alpha + \beta$ is odd.

We do this recursively. If $(a^\beta \colon \beta < \alpha)$ has been defined, and $\alpha < 2^\omega$, then since there are continuum-many plays in which the strategy f^α is used, one of them, to be called a^α, is not among those a^β such that $\beta < \alpha$ and $\beta + \alpha$ is odd.

Since, if $\alpha + e$ is even, player e can use strategy f^α for the play a^α, this means player $1-e$ has *some* strategy that, with f^α, determines a^α. That is, player $1 - e$ can win against strategy f^α, provided $a^\alpha \in A_{1-e}$. We now choose the partition of ${}^{\omega}2$ so that

$$\{a^\alpha \colon \alpha \text{ even}\} \subseteq A_1, \qquad \{a^\alpha \colon \alpha \text{ odd}\} \subseteq A_0.$$

Then neither player has a winning strategy for the game: the game is not determined.

4.3. Spectra

The **spectrum** of a commutative ring is the set of its prime ideals. The spectrum of a commutative ring R can be denoted by

$$\operatorname{Spec}(R).$$

We are going to define a *topology* on $\operatorname{Spec}(R)$. Let us recall what this means.

4.3.1. Topologies

Topologies can be defined in terms of *open sets* or *closed sets.* We shall use *closed sets.* Given an arbitrary set A, let us understand a **topology** on A to be a family τ of subsets of A such that

1) if X and Y are in τ, then $X \cup Y \in \tau$;
2) if $\mathscr{X} \subseteq \tau$, then $\bigcap \mathscr{X} \in \tau$;
3) $\varnothing \in \tau$.

In words, (1) τ is closed under finite unions and (2) arbitrary intersections, and (3) τ contains the empty set. The pair (A, τ) is called a **topological space.**

In condition (2) of the definition, we allow $\mathscr{X}$ to be $\varnothing$, and then we understand $\bigcap \varnothing$ to be A itself; thus we have (4) $A \in \tau$. Perhaps most writers will give this fourth condition as part of the *definition* of a topology as a fourth condition, without noting that it can be derived from condition (2).

Conditions (1) and (3) together are that τ is the universe of a substructure of the monoid $(\mathscr{P}(A), \varnothing, \cup)$: in short, τ is a submonoid of $(\mathscr{P}(A), \varnothing, \cup)$. (It would be ambiguous to say τ is a submonoid of $\mathscr{P}(A)$ simply, because $(\mathscr{P}(A), A, \cap)$ is also a monoid.)

The elements of the topology τ on A are the **closed** subsets of A with respect to the topology. The complement in A of a closed subset is an **open** subset. For example, in the *Euclidean topology* on $\mathbb{R}$, the open subsets are the unions of open intervals. Hence the closed subsets of $\mathbb{R}$ in this topology are intersections of closed intervals. In particular, finite unions of closed intervals are closed sets. However, some closed subsets of $\mathbb{R}$ are not unions of closed intervals. The **Cantor Set** is an example: this is the complement of the union of $(-\infty, 0)$ and $(1, \infty)$ and all of the intervals

$$\left(\sum_{k<n} \frac{2e_k}{3^{k+1}} + 1, \sum_{k<n} \frac{2e_k}{3^{k+1}} + 2\right),$$

where $n \in \mathbb{N}$ and $(e_k \colon k < n) \in 2^n$. The Cantor set is the set to which a bijection from $\mathscr{P}(\omega)$ is defined in the proof of the uncountability of $\mathbb{R}$ (Theorem 42, page 58).

Given a topology τ on A and an arbitrary subset X of A, we define

$$\bar{X} = \bigcap \{Y \in \tau \colon X \subseteq Y\}.$$

Theorem 100. *In an arbitrary topological space,*

1) $X \subseteq \bar{X}$ *and*

2) $\bar{X}$ *is closed,*

so $\bar{X}$ *is the smallest closed subset that includes* X*. Moreover,*

$$X \subseteq \bar{X}, \qquad X \subseteq \bar{Y} \implies \bar{X} \subseteq \bar{Y}. \tag{4.2}$$

The set $\bar{X}$ is called the **closure** of X with respect to the topology.

4.3.2. Closure operations and Moore families

An arbitrary operation $X \mapsto \bar{X}$ on $\mathscr{P}(A)$ with the properties in (4.2) is called a **closure operation** on A. We easily have the following.

Theorem 101. *An operation* $X \mapsto \bar{X}$ *on* $\mathscr{P}(A)$ *is a closure operation on* A *if and only if*

$$X \subseteq \bar{X}, \qquad X \subseteq Y \implies \bar{X} \subseteq \bar{Y}, \qquad \bar{\bar{X}} = \bar{X}.$$

To obtain a closure operation from a topology does not actually require every part of the definition of a topology. Weakening the definition, we shall say that a subset $\mathscr{F}$ of $\mathscr{P}(A)$ is a **Moore family** on A if

$$\mathscr{X} \subseteq \mathscr{F} \implies \bigcap \mathscr{X} \in \mathscr{F}.$$

Again we understand $\bigcap\varnothing$ to be A; so this is in $\mathscr{F}$. A topology on A is then just a Moore family on A that is also a submonoid of $(\mathscr{P}(A),\varnothing,\cup)$.

We have already encountered Moore families. By Theorem 72 (page 82), the family of ideals of a commutative ring R is a Moore family on R. But the family of *prime* ideals of R is *not* always a Moore family on R. For example, in $\mathbb{Z}$, $(2)\cap(3)=(6)$, which is not prime. Birkhoff [5, p. 111] attributes to Moore the following theorem.[4]

Theorem 102. *Let A be a set.*

1. *If $X \mapsto \bar{X}$ is a closure operation on A, then $\{\bar{X}\colon X \subseteq A\}$ is a Moore family on A.*
2. *If $\mathscr{F}$ is a Moore family on A, then the operation*

$$X \mapsto \bigcap\{Y \in \mathscr{F}\colon X \subseteq Y\}$$

   on $\mathscr{P}(A)$ is a closure operation on A.

3. *The given conversions between closure operations and Moore families are inverses of one another.*

Proof. Suppose $X \mapsto \bar{X}$ is a closure operation on A, and $\mathscr{F}=\{\bar{X}\colon X \subseteq A\}$. Let $\mathscr{X} \subseteq \mathscr{F}$. If $Y \in \mathscr{X}$, then

$$\bigcap\mathscr{X} \subseteq Y, \qquad\qquad \overline{\bigcap\mathscr{X}} \subseteq Y.$$

Therefore

$$\bigcap\mathscr{X} \subseteq \overline{\bigcap\mathscr{X}} \subseteq \bigcap\mathscr{X},$$

so these last inclusions must be equations, and $\bigcap\mathscr{X} \in \mathscr{F}$. The rest is easy. □

For example, since the family of ideals of a commutative ring is a Moore family, the operation $X \mapsto (X)$ on the ring is a closure operation.

[4] The precise reference is to E. H. Moore's *Introduction to a form of general analysis,* 1910.

In general, if $X \mapsto \bar{X}$ is a closure operation on A, it is reasonable to say that each subset $\bar{X}$ of A is **closed** and is the **closure** of X, with respect to the given closure operation. However, the resulting Moore family of closed subsets of A need not be a topology, because it need not be closed under finite unions and it need not contain $\varnothing$. For example, the ideals of a commutative ring do not compose a topology on the ring.

4.3.3. Galois correspondences

Closure operations arise in the following setting. Let A and B be two arbitrary sets, and suppose there are functions $X \mapsto X^*$ from $\mathscr{P}(A)$ to $\mathscr{P}(B)$ and $Y \mapsto Y^\dagger$ from $\mathscr{P}(B)$ to $\mathscr{P}(A)$ such that

$$X \subseteq X_1 \implies {X_1}^* \subseteq X^*, \qquad Y \subseteq Y_1 \implies {Y_1}^\dagger \subseteq Y^\dagger$$

(that is, the two functions are inclusion-reversing), and also

$$X \subseteq (X^*)^\dagger, \qquad Y \subseteq (Y^\dagger)^*.$$

Then the two functions constitute a **Galois correspondence** between $\mathscr{P}(A)$ and $\mathscr{P}(B)$. We shall show on page 110 how the original Galois correspondence in field theory is a special case. The general definition is apparently due to Øystein Ore,[5] who proves the following [47, Theorem 2, §2, p. 496]:

Theorem 103. *Suppose $X \mapsto X^*$ and $Y \mapsto Y^\dagger$ constitute a Galois correspondence between $\mathscr{P}(A)$ and $\mathscr{P}(B)$. Then the operations*

$$X \mapsto (X^*)^\dagger, \qquad Y \mapsto (Y^\dagger)^*$$

are closure operations on A and B respectively. The closed subsets of A and the closed subsets of B are in one-to-one, inclusion-reversing correspondence under the Galois correspondence.

[5] Ore's situation is even more general, with arbitrary (partially) ordered sets in place of $\mathscr{P}(A)$ and $\mathscr{P}(B)$.

Proof. The defining properties of a Galois correspondence give

$$X^* \subseteq ((X^*)^\dagger)^*, \qquad\qquad ((X^*)^\dagger)^* \subseteq X^*,$$

and therefore

$$X^* = ((X^*)^\dagger)^*.$$

By symmetry

$$Y^\dagger = ((Y^\dagger)^*)^\dagger.$$

Then we have, as special cases,

$$(Y^\dagger)^* = (((Y^\dagger)^*)^\dagger)^*, \qquad\qquad (X^*)^\dagger = (((X^*)^\dagger)^*)^\dagger.$$

All claims now follow. □

It will be useful to note the following.

Theorem 104. *Suppose $X \mapsto X^*$ and $Y \mapsto Y^\dagger$ constitute a Galois correspondence between $\mathscr{P}(A)$ and $\mathscr{P}(B)$. Then*

$$X^* = \bigcap_{a \in X} \{a\}^*, \qquad\qquad Y^\dagger = \bigcap_{b \in Y} \{b\}^\dagger.$$

Proof. Let $b \in X$. Then

$$X^* \subseteq \bigcap_{a \in X} \{a\}^* \subseteq \{b\}^*,$$

$$(\{b\}^*)^\dagger \subseteq \left(\bigcap_{a \in X} \{a\}^* \right)^\dagger \subseteq (X^*)^\dagger,$$

$$X \subseteq \bigcup_{a \in X} (\{a\}^*)^\dagger \subseteq \left(\bigcap_{a \in X} \{a\}^* \right)^\dagger \subseteq (X^*)^\dagger.$$

Since $(X^*)^\dagger$ is the closure of X, while $\left(\bigcap_{a \in X} \{a\}^*\right)^\dagger$ is closed, we have

$$\left(\bigcap_{a \in X} \{a\}^* \right)^\dagger = (X^*)^\dagger.$$

Since both $\bigcap_{a \in X} \{a\}^*$ and X^* are closed, we are done. □

In particular, the subsets $\{a\}^*$ of B compose a *basis* of the induced Moore family of closed subsets of B, in the sense of the next subsection (page 112).

The notion of a Galois correspondence is a generalization from the following special case. Let A and B be two arbitrary sets, and let R be a relation from A to B, so that, formally,

$$R \subseteq A \times B.$$

Given subsets X of A and Y of B, we define

$$X^* = \bigcap_{a \in X} \{y \in B\colon a \mathrel{R} y\}, \qquad Y^\dagger = \bigcap_{b \in Y} \{x \in A\colon x \mathrel{R} b\}.$$

These definitions are due to Birkhoff,[6] who refers to the functions $X \mapsto X^*$ and $Y \mapsto Y^\dagger$ as **polarities.** Then he easily observes the following.

Theorem 105. *The polarities induced by a relation constitute a Galois correspondence.*

For example, suppose L is a field with subfield K. Then a Galois correspondence—the original Galois correspondence—is induced by the relation R between L and $\mathrm{Aut}(L/K)$ given by

$$x \mathrel{R} \sigma \iff x^\sigma = x.$$

The existence of a one-to-one correspondence between the closed subsets of L and the closed subsets of $\mathrm{Aut}(L/K)$ is now easy: it follows from Theorem 103 (page 108). The hard part is identifying what those closed subsets are. Easily they are subfields of L that include K, and subgroups of $\mathrm{Aut}(L/K)$, respectively. If F is such a subfield, and G is such a subgroup, then the Galois correspondence is given by

$$F^* = \mathrm{Aut}(L/F), \qquad G^\dagger = \mathrm{Fix}(F).$$

[6] In the third edition of his *Lattice Theory* [5, ch. V, §7, p. 122], Birkhoff cites the first edition of his book, from 1940, as being the origin.

But it is not always the case that F and G are closed. It *is* the case if $\mathrm{Aut}(L/K)$ is finite and K is closed: this is the great theorem of the original Galois theory.[7]

Øystein Ore shows that *every* Galois correspondence arises from a relation [47, Theorem 10, §5, p. 503]:

Theorem 106. *For every Galois correspondence between power sets $\mathscr{P}(A)$ and $\mathscr{P}(B)$, there is a relation between A and B whose induced polarities constitute the Galois correspondence.*

Proof. Let the Galois correspondence be constituted by $X \mapsto X^*$ and $Y \mapsto Y^\dagger$. By Theorem 104, if we define the relation R between A and B by

$$x \mathrel{R} y \iff y \in \{x\}^*,$$

then $X \mapsto X^*$ is the induced polarity. The same is true for $Y \mapsto Y^\dagger$ by symmetry, since

$$y \in \{x\}^* \implies \{y\} \subseteq \{x\}^* \implies (\{x\}^*)^\dagger \subseteq \{y\}^\dagger \implies x \in \{y\}^\dagger. \quad \square$$

Finally we observe that every Moore family arises from a Galois correspondence:

Theorem 107. *A Moore family $\mathscr{F}$ on a set A consists of the closed subsets of A determined by the Galois correspondence induced by the relation $\in$ between A and $\mathscr{F}$.*

Proof. The given Galois correspondence is

$$X \mapsto \{Y \in \mathscr{F} : X \subseteq Y\}, \qquad \mathscr{Y} \mapsto \bigcap \mathscr{Y}. \qquad \square$$

[7] This follows from the theorem that Hungerford [34, Ch. V, Theorem 2.15, p. 252] names for Artin. In his *Galois Theory* [1, Theorem 13, p. 36], Artin first shows $|\mathrm{Aut}(L/K)| \leqslant [L : K]$. "Artin's Theorem" [1, Theorem 14, p. 42] is that, if G is a finite subgroup of $\mathrm{Aut}(L)$ and $K = G^\dagger$, then $[L : K] = |G|$. In this case, we must also have $[L : K] = |(G^\dagger)^*|$; so $G = (G^\dagger)^*$ and thus G is closed. Also L/K must be separable, and from this it follows that, if $K \subseteq F \subseteq L$, then F is closed.

4.3.4. Bases

If $\mathscr{F}$ is a Moore family on A, and $\mathscr{B}$ is a subset of $\mathscr{F}$ such that

$$F \in \mathscr{F} \implies F = \bigcap\{X \in \mathscr{B} \colon F \subseteq X\},$$

then $\mathscr{B}$ is a **basis** for $\mathscr{F}$.

Theorem 108. *Let A be a set.*

1. *A Moore family on A is a basis of itself.*
2. *The family of Moore families on A is a Moore family on $\mathscr{P}(A)$.*
3. *Every subset of $\mathscr{P}(A)$ is a basis of its closure with respect to the Moore family of Moore families on A.*

If $\mathscr{B}$ is a basis of the Moore family $\mathscr{F}$ on A, then $\mathscr{B}$ may be said to **generate** $\mathscr{F}$. As a corollary of Theorem 104, we have:

Theorem 109. *Suppose $X \mapsto X^*$ and $Y \mapsto Y^\dagger$ constitute a Galois correspondence between $\mathscr{P}(A)$ and $\mathscr{P}(B)$. The sets $\{a\}^*$, where $a \in A$, compose a basis for the Moore family of closed subsets of B.*

We can also generalize Theorem 107:

Theorem 110. *If $\mathscr{B}$ generates the Moore family $\mathscr{F}$ on A, then $\mathscr{F}$ consists of the closed subsets of A determined by the Galois correspondence induced by the relation $\in$ between A and $\mathscr{B}$.*

Proof. The given Galois correspondence is

$$X \mapsto \{Y \in \mathscr{B} \colon X \subseteq Y\}, \qquad\qquad \mathscr{Y} \mapsto \bigcap \mathscr{Y}. \qquad \square$$

A basis for a *topology* τ on A need not contain $\varnothing$ or be closed under $\cup$; that is, the basis need not be a submonoid of τ. However, it may be, since τ is a basis of itself.

Theorem 111. *If $\mathscr{B}$ is a submonoid of $(\mathscr{P}(A), \varnothing, \cup)$, then the Moore family generated by $\mathscr{B}$ is a topology on A.*

If $\mathscr{C}$ is an arbitrary subset of $\mathscr{P}(A)$, then $\mathscr{C}$ generates a submonoid $\mathscr{B}$ of $(\mathscr{P}(A), \varnothing, \cup)$, and $\mathscr{C}$ may be called a **sub-basis** of the topology generated by $\mathscr{B}$.

As a corollary of Theorems 109 and 111, we have:

Theorem 112. *Suppose $X \mapsto X^*$ and $Y \mapsto Y^\dagger$ constitute a Galois correspondence between $\mathscr{P}(A)$ and $\mathscr{P}(B)$. If also A has an element 1 and a binary operation $\cdot$ such that*

$$\varnothing = \{1\}^*, \qquad\qquad \{x\}^* \cup \{y\}^* = \{x \cdot y\}^*,$$

then the closed subsets of B compose a topology on B.

The topology on $\operatorname{Spec}(R)$ (promised on page 105) will arise in this way in Theorem 113. Indeed, *every* topology arises in this way, by Theorem 110.

In the theorem, the structure $(A, 1, \cdot\,)$ need not be a monoid. However, if we define the binary relation $\sim$ on A by

$$x \sim y \iff \{x\}^* = \{y\}^*,$$

then $\sim$ will be a congruence-relation on $(A, 1, \cdot\,)$ in the sense of §3.2.1 (page 76), and the quotient $(A, 1, \cdot\,)/{\sim}$ will be a monoid. We shall establish a variant of this result as Theorem 148 (page 152). Meanwhile, in the situations of interest, $(A, 1, \cdot\,)$ will already be known to be a monoid, and $\sim$ will be equality.

All of the notions of this section can now be defined in terms of a relation between two sets:

1. A *Galois correspondence* between $\mathscr{P}(A)$ and $\mathscr{P}(B)$ consists of the polarities induced by a relation from A to B.
2. A *Moore family* of subsets of A consists of the closed subsets of A with respect to the Galois correspondence between $\mathscr{P}(A)$ and $\mathscr{P}(B)$ induced by some relation R from A to B for some set B.
3. That Moore family is a *topology* on A, if B has an element 1 and a binary operation $\cdot$ such that

$$\neg\, a \mathrel{R} 1, \qquad\qquad a \mathrel{R} (x \cdot y) \iff a \mathrel{R} x \text{ OR } a \mathrel{R} y.$$

4.3.5. The topology on a spectrum

We define three possible properties of a topology τ on a set A.

1. The topology τ is **compact** if, for every subset $\mathscr{X}$ of τ such that

$$\bigcap \mathscr{X} = \varnothing,$$

there is a finite subset $\{X_0, \dots, X_{n-1}\}$ of $\mathscr{X}$ such that

$$X_0 \cap \dots \cap X_{n-1} = \varnothing.$$

If τ has basis $\mathscr{B}$, it is enough to assume $\mathscr{X} \subseteq \mathscr{B}$. We may use the definition in the contrapositive form. A subset $\mathscr{X}$ of τ has the **finite intersection property** if its every finite subset has nonempty intersection. Then τ is compact if and only if its every subset with the finite intersection property has nonempty intersection.

2. Two points of A are **topologically indistinguishable** if every member of τ contains either both or neither of the points. It is enough if this is true for every member of a given basis. The topology τ is **Kolmogorov,** or T_0, if *no* two distinct points of A are topologically indistinguishable.

3. The topology τ is **Hausdorff** if for all distinct elements x_0 and x_1 of A there are elements F_0 and F_1 of τ such that

$$x_0 \notin F_0, \qquad x_1 \notin F_1, \qquad F_0 \cup F_1 = A.$$

Again it is enough to require F_0 and F_1 to belong to a given basis.

Given an element a of a commutative ring R, let us use the notation

$$\mathrm{Z}(a) = \{\mathfrak{p} \in \operatorname{Spec}(R) \colon a \in \mathfrak{p}\}.$$

This gives us the following.

Theorem 113. *Let R be a commutative ring.*

1. *The set $\{\mathrm{Z}(x) \colon x \in R\}$ is a basis for a topology on* $\operatorname{Spec}(R)$*, since*

$$\varnothing = \mathrm{Z}(1), \qquad \mathrm{Z}(x) \cup \mathrm{Z}(y) = \mathrm{Z}(xy). \tag{4.3}$$

2. *The topology is Kolmogorov.*

PI 3. *By the Prime Ideal Theorem (page 101),* Spec(R) *is nonempty, and its topology is compact.*

4. *If R is a Boolean ring, the topology is Hausdorff, and the complement of every* $\mathrm{Z}(x)$ *is* $\mathrm{Z}(1+x)$.

Proof. 1. Let a Galois correspondence $X \mapsto X^*$ and $Y \mapsto Y^\dagger$ between $\mathscr{P}(R)$ and $\mathscr{P}(\mathrm{Spec}(R))$ be determined by the relation $\in$ between R and Spec(R). Then

$$\{x\}^* = \mathrm{Z}(x).$$

Since elements $\mathfrak{p}$ of Spec(R) are prime ideals, we have

$$xy \in \mathfrak{p} \iff x \in \mathfrak{p} \text{ OR } y \in \mathfrak{p},$$
$$\mathfrak{p} \in \mathrm{Z}(xy) \iff \mathfrak{p} \in \mathrm{Z}(x) \text{ OR } \mathfrak{p} \in \mathrm{Z}(y),$$

and so (4.3) holds. By Theorem 112, the sets $\mathrm{Z}(x)$ compose a basis of a topology on Spec(R).

2. If $\mathfrak{p}$ and $\mathfrak{q}$ are distinct elements of Spec(R), we may assume $a \in \mathfrak{p} \smallsetminus \mathfrak{q}$, and so $\mathrm{Z}(a)$ contains $\mathfrak{p}$, but not $\mathfrak{q}$.

3. Suppose $A \subseteq R$. Then

$$\bigcap_{x\in A} \mathrm{Z}(x) = \{\mathfrak{p} \in \mathrm{Spec}(R) \colon A \subseteq \mathfrak{p}\}.$$

If (A) is a proper ideal of R, then by Theorem 99 (page 102) it is included in a prime ideal, which belongs to Spec(R) and therefore to $\bigcap_{x\in A} \mathrm{Z}(x)$. It follows that, if

$$\bigcap_{x\in A} \mathrm{Z}(x) = \varnothing,$$

then (A) must contain 1. In this case, by Theorem 77 (page 85), there is x in $\bigoplus_{a\in A} R$ such that

$$1 = \sum_{a\in A} x_a a.$$

Then

$$\bigcap_{a \in \operatorname{supp}(x)} \mathrm{Z}(a) = \varnothing.$$

Since $\operatorname{supp}(x)$ is a finite subset of A, the topology of $\operatorname{Spec}(R)$ is compact.

4. In a Boolean ring R, since $1+1=0$ (Theorem 86, page 95), every element of $\operatorname{Spec}(R)$ contains exactly one of x and $1+x$ (Theorem 87, page 95), so $\operatorname{Spec}(R)$ is the disjoint union of $\mathrm{Z}(x)$ and $\mathrm{Z}(1+x)$. If $\mathfrak{p}$ and $\mathfrak{q}$ are distinct elements of $\operatorname{Spec}(R)$, then we may assume $\mathfrak{q} \smallsetminus \mathfrak{p}$ has an element a, and then

$$\mathfrak{p} \notin \mathrm{Z}(a), \qquad \mathfrak{q} \notin \mathrm{Z}(1+a), \qquad \mathrm{Z}(a) \cup \mathrm{Z}(1+a) = \operatorname{Spec}(R). \qquad \square$$

The topology on $\operatorname{Spec}(R)$ given by the theorem is the **Zariski topology** on $\operatorname{Spec}(R)$. The corresponding closed subsets of R are just the intersections of collections of prime ideals of R; we shall characterize such intersections in Corollary 117.1 in the next section.

4.4. Radical ideals

We develop an analogue of Theorems 78 (page 86) and 85 (page 94). An element a of a commutative ring R is called **nilpotent** if some power a^n of the element is 0. In particular, 0 itself is nilpotent. The ring R is called **reduced** if it has no nonzero nilpotents. For example, every Boolean ring is reduced. An ideal I of R is called **radical** if

$$x^2 \in I \implies x \in I.$$

Every prime ideal of every commutative ring is radical. Indeed, every *intersection* of prime ideals is radical. Thus, under the Galois correspondence induced by the relation $\in$ between R and $\operatorname{Spec}(R)$, all of the closed subsets of R are radical ideals. We shall establish the converse. Meanwhile, we have the analogue promised above.

Theorem 114. *Let R be a commutative ring.*

1. The ideal (0) of R is radical if and only if R is reduced.

2. *An ideal I of R is radical if and only if the quotient R/I is reduced.*

Thus

$$\begin{aligned}\text{radical ideal : reduced ring} &:: \text{prime ideal : integral domain}\\ &:: \text{maximal ideal : field.}\end{aligned}$$

But the following easy result does not hold for maximal ideals or prime ideals. Recall from page 106 that a *Moore family* on a set is just a family of subsets that is closed under arbitrary intersections. Then the following is an analogue of Theorem 72 (page 82).

Theorem 115. *The radical ideals of a commutative ring R compose a Moore family on R.*

By this and Theorem 102 (page 107), the Moore family of radical ideals of R induces a closure operation

$$X \mapsto \sqrt{}(X)$$

on R. If I is an ideal of R, then $\sqrt{}I$ is called the **radical** of I: it is the smallest radical ideal that includes I. Then I is radical if and only if $I = \sqrt{}I$.

Given a subset X of R, we characterized (X) in Theorem 77 (page 85). Now we can characterize $\sqrt{}(X)$:

Theorem 116. *If I is an ideal of the commutative ring R, then*

$$\sqrt{}I = \bigcup_{n\in\mathbb{N}} \{x \in R\colon x^n \in I\}.$$

But the following characterization will be of more theoretical interest.

Theorem 117. *By Zorn's Lemma, for all subsets A of a commutative ring R,* **AC**

$$\sqrt{}(A) = \bigcap\{\mathfrak{p} \in \operatorname{Spec}(R)\colon A \subseteq \mathfrak{p}\}. \tag{4.4}$$

Proof. Since prime ideals are radical, and $\sqrt{}(A)$ is the smallest radical ideal that includes A, it is clear that

$$\sqrt{}(A) \subseteq \bigcap\{\mathfrak{p} \in \operatorname{Spec}(R) \colon A \subseteq \mathfrak{p}\}.$$

To prove the reverse inclusion, suppose $x \in R \smallsetminus \sqrt{}(A)$; we show the intersection in (4.4) does not contain x either. Using Zorn's Lemma, we let $\mathfrak{b}$ be an ideal of R that is maximal with respect to including $\sqrt{}(A)$, but not containing any power of x. Say y and z are not in $\mathfrak{b}$. By maximality, we have

$$x \in \mathfrak{b} + (y), \qquad\qquad x \in \mathfrak{b} + (z),$$

and therefore, by multiplying,

$$x^2 \in \mathfrak{b} + (yz),$$

so $yz \notin \mathfrak{b}$ (since $x^2 \notin \mathfrak{b}$). Thus $\mathfrak{b}$ is prime, so it belongs to the intersection in (4.4). Therefore this intersection does not contain x. Thus

$$\sqrt{}(A) \supseteq \bigcap\{\mathfrak{p} \in \operatorname{Spec}(R) \colon A \subseteq \mathfrak{p}\}. \qquad \square$$

When R is a Boolean ring, $\sqrt{}(A)$ is just (A), and also the theorem needs only the Prime Ideal Theorem, because in this case, for a prime ideal *not* to contain x is the same as containing $1 + x$.

Corollary 117.1. *For every commutative ring R, under the Galois correspondence induced by the relation $\in$ between R and $\operatorname{Spec}(R)$, the closed subsets of R are precisely the radical ideals.*

For all commutative rings R, Theorem 73 (page 83) guarantees us a homomorphism

$$x \mapsto \big(x + \mathfrak{p} \colon \mathfrak{p} \in \operatorname{Spec}(R)\big) \tag{4.5}$$

from R to $\prod_{\mathfrak{p} \in \operatorname{Spec}(R)} R/\mathfrak{p}$.

Theorem 118. *A commutative ring R is reduced if, and by Zorn's Lemma only if, the homomorphism in* (4.5) *is an embedding.* **AC**

Proof. The homomorphism is an embedding if and only if

$$\bigcap \operatorname{Spec}(R) = (0).$$

By the last theorem, $\bigcap \operatorname{Spec}(R) = \sqrt{}(0)$. By Theorem 114, R is reduced if and only if $(0) = \sqrt{}(0)$. □

The **clopen** subsets of a topological space are the subsets that are both closed and open. The following, based originally on [57], is an analogue of Cayley's Theorem for groups (page 65) and Theorem 58 for associative rings (page 73).

Theorem 119 (Stone Representation Theorem for Boolean Rings). *Suppose R is a Boolean ring.*

1. *By the Prime Ideal Theorem, the Boolean ring R embeds in the Boolean ring $\mathscr{P}(\operatorname{Spec}(R))$ under the map* **PI**

$$x \mapsto \{\mathfrak{p} \in \operatorname{Spec}(R) \colon x \notin \mathfrak{p}\}. \tag{4.6}$$

2. *This map is $x \mapsto \mathrm{Z}(1 + x)$.*
3. *The image of this map is the set of clopen subsets of* $\operatorname{Spec}(R)$.

Proof. The map in (4.6) is part of the commutative diagram in Figure 4.1. We can spell out the details as follows. By Theorem 87 (page 95), for each $\mathfrak{p}$ in $\operatorname{Spec}(R)$, the quotient $R/\mathfrak{p}$ is isomorphic to the field $\mathbb{F}_2$, and so $\prod_{\mathfrak{p}\in\operatorname{Spec}(R)} R/\mathfrak{p}$ is isomorphic to ${\mathbb{F}_2}^{\operatorname{Spec}(R)}$. The inverse of this isomorphism is easier to write down: it is

$$(e_{\mathfrak{p}} \colon \mathfrak{p} \in \operatorname{Spec}(R)) \mapsto (e_{\mathfrak{p}} + \mathfrak{p} \colon \mathfrak{p} \in \operatorname{Spec}(R)).$$

The power ${\mathbb{F}_2}^{\operatorname{Spec}(R)}$ is in turn isomorphic to $\mathscr{P}(\operatorname{Spec}(R))$ under $x \mapsto \operatorname{supp}(x)$ by Theorem 80 (page 87). Then $x \mapsto \operatorname{supp}(x)$ is also an isomorphism from $\prod_{\mathfrak{p}\in\operatorname{Spec}(R)} R/\mathfrak{p}$ to $\mathscr{P}(\operatorname{Spec}(R))$. Preceding this with the embedding of R in $\prod_{\mathfrak{p}\in\operatorname{Spec}(R)} R/\mathfrak{p}$ given by the last theorem, we obtain the map in (4.6).

$$\begin{tikzcd}
R \arrow[rr, "x\mapsto\{\mathfrak{p}\in\operatorname{Spec}(R)\colon x\notin\mathfrak{p}\}"'] \arrow[d, "x\mapsto(x+\mathfrak{p}\colon \mathfrak{p}\in\operatorname{Spec}(R))"] & & \mathscr{P}(\operatorname{Spec}(R)) \\
\prod\limits_{\mathfrak{p}\in\operatorname{Spec}(R)} R/\mathfrak{p} \arrow[rru, "x\mapsto\operatorname{supp}(x)" description] \arrow[rr, "(e_{\mathfrak{p}}+\mathfrak{p}\colon \mathfrak{p}\in\operatorname{Spec}(R))\mapsto(e_{\mathfrak{p}}\colon \mathfrak{p}\in\operatorname{Spec}(R))"] & & {\mathbb{F}_2}^{\operatorname{Spec}(R)} \arrow[u, "x\mapsto\operatorname{supp}(x)"']
\end{tikzcd}$$

Figure 4.1.: Stone Representation Theorem

By Theorem 113 (page 114), this map is $x \mapsto \mathrm{Z}(1+x)$, and all of the sets $\mathrm{Z}(x)$ are clopen. Conversely, suppose a closed subset F of $\operatorname{Spec}(R)$ is also open. We have

$$F = \bigcap_{x\in I} \mathrm{Z}(x),$$

where $I = \bigcap F$. Being a closed subset of a compact space, the complement of F in $\operatorname{Spec}(R)$ is compact. Therefore I has a finite subset $\{x_0,\dots,x_{n-1}\}$ such that

$$\begin{aligned}
F &= \mathrm{Z}(x_0)\cap\dots\cap\mathrm{Z}(x_{n-1}) = \mathrm{Z}(x_0\cdots x_{n-1}),\\
\operatorname{Spec}(R)\smallsetminus F &= \mathrm{Z}(1+x_0)\cup\dots\cup\mathrm{Z}(1+x_{n-1})\\
&= \mathrm{Z}((1+x_0)\cdots(1+x_{n-1})),
\end{aligned}$$

$$F = \mathrm{Z}(1 + (1 + x_0) \cdots (1 + x_{n-1})). \qquad \square$$

We shall see this theorem in another form as Theorem 147 (page 151). Meanwhile, for an arbitrary commutative ring R, since each quotient $R/\mathfrak{p}$ is an integral domain, it will be seen to embed in a field (see page 123), and so, by Theorem 118, every reduced ring will embed in a product of fields.

4.5. Localization

It will be useful now to generalize the construction of $\mathbb{Q}$ from $\mathbb{Z}$ that is suggested by Theorem 34 (page 54). A subset of a commutative ring is called **multiplicative** if it is nonempty and closed under multiplication. For example, $\mathbb{Z} \smallsetminus \{0\}$ is a multiplicative subset of $\mathbb{Z}$, and more generally, we have the following.

Theorem 120. *An ideal $\mathfrak{p}$ of a commutative ring R is prime if and only if the complement $R \smallsetminus \mathfrak{p}$ is multiplicative.*

For example, by Theorem 84 (page 93), the elements of a nontrivial commutative ring that are neither 0 nor zero-divisors compose a multiplicative subset. Other examples of multiplicative subsets of a commutative ring R are $\{1\}$ and and $R^{\times}$. However, the complements of prime ideals are the only examples of multiplicative subsets that will interest us.

Lemma 11. *If S is a multiplicative subset of a commutative ring R, then on $R \times S$ there is an equivalence relation $\sim$ given by*

$$(a, b) \sim (c, d) \iff (ad - bc) \cdot e = 0 \text{ for some } e \text{ in } S. \tag{4.7}$$

Proof. Reflexivity and symmetry are obvious. For transitivity, note that, if $(a, b) \sim (c, d)$ and $(c, d) \sim (e, f)$, so that, for some g and h in S,

$$0 = (ad - bc)g = adg - bcg, \qquad 0 = (cf - de)h = cfh - deh,$$

then $(a,b) \sim (e,f)$ since

$$\begin{aligned}(af-be)cdgh &= afcdgh - becdgh \\ &= adgcfh - bcgdeh = bcgcfh - bcgcfh = 0. \qquad \square\end{aligned}$$

In the notation of the lemma, the equivalence class of the element (a,b) of $R \times S$ is denoted by one of

$$a/b, \qquad \frac{a}{b},$$

and the quotient $(R \times S)/{\sim}$ is denoted by one of

$$S^{-1}R, \qquad R[S^{-1}].$$

If $0 \in S$, then $S^{-1}R$ has exactly one element, which is $0/0$. If R is an integral domain and $0 \notin S$, then the relation $\sim$ in the theorem is given simply by

$$(a,b) \sim (c,d) \iff ad = bc.$$

However, we shall be interested in commutative rings that are not integral domains.

Theorem 121. *Suppose R is a commutative ring with multiplicative subset S.*

1. *In $S^{-1}R$, if $c \in S$,*
$$\frac{a}{b} = \frac{ac}{bc}.$$
2. *$S^{-1}R$ is a commutative ring in which the operations are given by*
$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}, \qquad \frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd}.$$
3. *There is a ring-homomorphism φ from R to $S^{-1}R$ where, for every a in S,*
$$\varphi(x) = \frac{xa}{a}.$$
In particular, if $1 \in S$, then $\varphi(x) = x/1$.

4. *The homomorphism φ is injective if and only if S contains neither* 0 *nor zero-divisors.*

Suppose in particular R is an integral domain and $0 \notin S$.

5. *$S^{-1}R$ is an integral domain (and φ is an embedding).*
6. *If $S = R \smallsetminus \{0\}$, then $S^{-1}R$ is a field, and if ψ is an embedding of R in a field K, then there is an embedding $\tilde{\psi}$ of $S^{-1}R$ in K such that $\tilde{\psi} \circ \varphi = \psi$. (See Figure 4.2.)*

$$\begin{tikzcd}
R \arrow[r, "\psi"] \arrow[d, "\varphi"'] & K \\
S^{-1}R \arrow[ur, dotted, "\tilde{\psi}"'] &
\end{tikzcd}$$

Figure 4.2.: The universal property of the quotient field

Corollary 121.1. *A commutative ring is an integral domain if and only if it is a subring of a field.*

See page 192 for a model-theoretic consequence of the corollary.

When S is the complement of a prime ideal $\mathfrak{p}$, then $S^{-1}R$ is called the **localization** of R at $\mathfrak{p}$ and can be denoted by

$$R_{\mathfrak{p}}.$$

If R is an integral domain, so that (0) is prime, then localization $R_{(0)}$ (which is a field by the theorem) is the **quotient-field** of R. In this case, the last part of the theorem describes the quotient field in terms of a *universal property* in the sense of page 235. However, it is important to note that, if R is not an integral domain, then the homomorphism $x \mapsto x/1$ from R to $R_{\mathfrak{p}}$ might not be an embedding. The following will be generalized as Theorem 126 (page 126).

Theorem 122. *For every Boolean ring R, for every $\mathfrak{p}$ in* $\operatorname{Spec}(R)$*, the homomorphism*

$$x \mapsto \frac{x}{1}$$

from R to $R_{\mathfrak{p}}$ is surjective and has kernel $\mathfrak{p}$. Thus

$$R_{\mathfrak{p}} \cong R/\mathfrak{p}$$

(which is isomorphic to $\mathbb{F}_2$ by Theorem 87, page 95).

Proof. If $a \in R$ and $b \in R \smallsetminus \mathfrak{p}$, then $a/b = a/1$ since $(a - ab) \cdot b = 0$. Thus $x \mapsto x/1$ is surjective. If $a \in \mathfrak{p}$, then $1 + a \in R \smallsetminus \mathfrak{p}$, and $a \cdot (1 + a) = 0$, so $a/1 = 0/1$. Thus the kernel of $x \mapsto x/1$ includes $\mathfrak{p}$. Therefore the kernel must *be* $\mathfrak{p}$, since this ideal is maximal by Theorem 87, and $R_{\mathfrak{p}}$ is not trivial. □

A **local ring** is a commutative ring with a unique maximal ideal. The connection between localizations and local rings is made by the theorem below.

Lemma 12. *An ideal $\mathfrak{m}$ of a commutative ring R is a unique maximal ideal of R if and only if*

$$R^{\times} = R \smallsetminus \mathfrak{m}.$$

Theorem 123. *The localization $R_{\mathfrak{p}}$ of a commutative ring R at a prime ideal $\mathfrak{p}$ is a local ring whose unique maximal ideal is*

$$\mathfrak{p}R_{\mathfrak{p}},$$

namely the ideal generated by the image of $\mathfrak{p}$.

Proof. The ideal $\mathfrak{p}R_{\mathfrak{p}}$ consists of those a/b such that $a \in \mathfrak{p}$. In this case, if $c/d = a/b$, then $cb = da$, which is in $\mathfrak{p}$, so $c \in \mathfrak{p}$ since $\mathfrak{p}$ is prime and $b \notin \mathfrak{p}$. Hence for all x/y in $R_{\mathfrak{p}}$,

$$\begin{aligned} x/y \notin R_{\mathfrak{p}}\mathfrak{p} &\iff x \notin \mathfrak{p} \\ &\iff x/y \text{ has an inverse, namely } y/x. \end{aligned}$$

By the lemma, we are done. □

We can now refer to $R_{\mathfrak{p}}$ (where $\mathfrak{p}$ is prime) as the local ring of R at $\mathfrak{p}$.

4.6. Regular rings

By Theorem 87 (page 95), the Boolean rings are commutative rings whose prime ideals are maximal. There is a larger class of commutative rings whose prime ideals are maximal. Indeed, by the Stone Representation Theorem (page 119), every Boolean ring embeds in a power set $\mathscr{P}(\Omega)$ and hence in a power ${\mathbb{F}_2}^{\Omega}$. This power is a special case of the direct product $\prod_{i\in\Omega} K_i$, where each K_i is a field. For every x in the ring $\prod_{i\in\Omega} K_i$ there is y in the ring such that

$$xyx = x.$$

Indeed, we can just let y be x^*, defined as on page 89. Therefore the ring $\prod_{i\in\Omega} K_i$ is called a **(von Neumann) regular ring.**[8] Thus Boolean rings are also regular rings in this sense, since in a Boolean ring

$$x \cdot 1 \cdot x = x.$$

A regular ring can also be understood as a ring in which, for all x,

$$x \in (x^2).$$

We have the following easily.

Theorem 124. *Every regular ring is reduced.*

Proof. Suppose R is regular and $x^2 = 0$. But $x = x^2y$ for some y, and so $x = 0$. □

We can establish the following generalization of the first part of Theorem 87 (page 95).

Theorem 125. *In regular rings, all prime ideals are maximal.*

Proof. If R is a regular ring, and $\mathfrak{p}$ is a prime ideal, then for all x in R, for some y in R,

$$(xy - 1) \cdot x = 0,$$

[8] In general, a regular ring need not be commutative; see [34, IX.3, ex. 5, p. 442].

and so at least one of $xy - 1$ and x is in $\mathfrak{p}$. Hence if $x + \mathfrak{p}$ is not 0 in $R/\mathfrak{p}$, then $x + \mathfrak{p}$ has the inverse $y + \mathfrak{p}$. Thus $R/\mathfrak{p}$ is a field, so $\mathfrak{p}$ is maximal. □

We now generalize Theorem 122 (page 123).

Theorem 126. *For every regular ring R, for every $\mathfrak{p}$ in* $\operatorname{Spec}(R)$*, the homomorphism*

$$x \mapsto \frac{x}{1}$$

from R to $R_{\mathfrak{p}}$ is surjective and has kernel $\mathfrak{p}$. Thus

$$R_{\mathfrak{p}} \cong R/\mathfrak{p}.$$

Proof. If $a \in R$ and $b \in R \smallsetminus \mathfrak{p}$, and $bcb = b$, then the elements a/b and $ac/1$ of $R_{\mathfrak{p}}$ are equal since

$$(a - bac)b = ab - abcb = ab - ab = 0.$$

Thus the homomorphism $x \mapsto x/1$ from R to $R_{\mathfrak{p}}$ guaranteed by Theorem 121 is surjective. By the last theorem, $\mathfrak{p}$ is maximal, and hence $R_{\mathfrak{p}}$ is a field. As in that theorem, supposing $x \in \mathfrak{p}$, we have

$$(xy - 1) \cdot x = 0$$

for some y, but $1 - xy \notin \mathfrak{p}$. This shows $x/1 = 0/1$. Thus the kernel of $x \mapsto x/1$ includes $\mathfrak{p}$. Having a prime ideal, R is not the trivial ring, so $R_{\mathfrak{p}}$ is not trivial, and thus the kernel of $x \mapsto x/1$ cannot be all of R. Therefore the kernel is $\mathfrak{p}$, since this is a maximal ideal. □

The foregoing three theorems turn out to *characterize* regular rings. That is, every ring of which the conclusions of these theorems hold must be regular. In fact a somewhat stronger statement is true; this is the next theorem below.

For any commutative ring R, the ideal $\surd(0)$ consists precisely of the nilpotent elements of R and is according called the **nilradical** of R. By Theorem 117 (page 117),

$$\surd(0) = \bigcap \operatorname{Spec}(R).$$

By Theorem 114 (page 116), this ideal is just (0) if and only if R is reduced.

Theorem 127. *By the Maximal Ideal Theorem, the following are equivalent conditions on a ring R.*[9] **AC**

1. *R is regular.*
2. *Every prime ideal of R is maximal, and R is reduced.*
3. *The localization $R_{\mathfrak{m}}$ is a field for all maximal ideals $\mathfrak{m}$ of R.*

Proof. 1. We have established (1)⇒(2) in Theorems 124 and 125.

2. We prove (2)⇒(3). Suppose every prime ideal of R is maximal, and R is reduced. Let $\mathfrak{m}$ be a maximal ideal of R. By Theorem 123 (page 124), $\mathfrak{m}R_{\mathfrak{m}}$ is the unique maximal ideal of $R_{\mathfrak{m}}$. By Zorn's Lemma, **AC** every prime ideal $\mathfrak{P}$ of $R_{\mathfrak{m}}$ is included in a maximal ideal; but then this must be $\mathfrak{m}R_{\mathfrak{m}}$. Now, the intersection $\mathfrak{m}R_{\mathfrak{m}} \cap R$ is a proper ideal of R that includes $\mathfrak{m}$, so it is $\mathfrak{m}$. Hence $\mathfrak{P} \cap R$ is a prime ideal of R that is included in $\mathfrak{m}$, so it is $\mathfrak{m}$, and therefore $\mathfrak{P} = \mathfrak{m}R_{\mathfrak{m}}$. Thus this maximal ideal is the unique prime ideal of $R_{\mathfrak{m}}$. This ideal is therefore $\bigcap \operatorname{Spec}(R_{\mathfrak{m}})$, which is the nilradical of the ring. Thus for all r/s in $\mathfrak{m}R_{\mathfrak{m}}$, for some n in $\mathbb{N}$, we have $(r/s)^n = 0$, so $r^n/s^n = 0$, and therefore $tr^n = 0$ for some t in $R \smallsetminus \mathfrak{m}$. In this case, $(tr)^n = 0$, so $tr = 0$, and therefore $r/s = 0$. In short, $\mathfrak{m}R_{\mathfrak{m}} = (0)$. Therefore $R_{\mathfrak{m}}$ is a field.

3. Finally, we show (3)⇒(1). Suppose $R_{\mathfrak{m}}$ is a field for all maximal ideals $\mathfrak{m}$ of R. If $x \in R$, define

$$I = \{r \in R \colon rx \in (x^2)\}.$$

This is an ideal of R containing x. We shall show that it contains 1. We do this by showing that it is not included in any maximal ideal $\mathfrak{m}$. If $x \notin \mathfrak{m}$, then, since $x \in I$, we have $I \nsubseteq \mathfrak{m}$. If $x \in \mathfrak{m}$, then

[9] The equivalence of these conditions is part of [25, Thm 1.16, p. 7]. This theorem adds a fourth equivalent condition: "All simple R-modules are injective." The proofs given involve module theory, except the proof that, if all prime ideals are maximal, and the ring is reduced, then each localization at a maximal ideal is a field. That proof is reproduced below.

$x/1 \notin (R_{\mathfrak{m}})^{\times}$, so, since $R_{\mathfrak{m}}$ is a field, we have $x/1 = 0/1$, and hence

$$rx = 0$$

for some r in $R \smallsetminus \mathfrak{m}$; but $r \in I$. Again $I \nsubseteq \mathfrak{m}$. Thus I must be (1), so $x \in (x^2)$. Therefore R is regular. □

We again consider the regular rings that are products $\prod \mathscr{K}$, where $\mathscr{K}$ is an indexed family $(K_i \colon i \in \Omega)$ of fields. Here we have $xx^*x = x$ when x^* is defined as in (3.8) on page 89. Hence every sub-ring of $\prod \mathscr{K}$ that is closed under the operation $x \mapsto x^*$ is also a regular ring.

We now prove the converse: every regular ring is isomorphic to a sub-ring, closed under $x \mapsto x^*$, of a product of fields. Since regular rings are reduced (Theorem 124), the homomorphism

$$x \mapsto \bigl(x + \mathfrak{p} \colon \mathfrak{p} \in \operatorname{Spec}(R)\bigr) \tag{4.8}$$

from R to $\prod_{\mathfrak{p} \in \operatorname{Spec}(R)} R/\mathfrak{p}$ (given also in (4.5), page 118) is an embedding by Theorem 118 (page 118). Moreover, the quotients $R/\mathfrak{p}$ are fields by Theorem 125 (and Theorem 78, page 86).

Theorem 128. *For every regular ring R, the image of the embedding in (4.8) of R in the product $\prod_{\mathfrak{p} \in \operatorname{Spec}(R)} R/\mathfrak{p}$ of fields is closed under $x \mapsto x^*$.*

Proof. Let the embedding be called f. Given x in R, we have to show that $f(x)^*$ is in the image of f. Now, there is y in R such that $xyx = x$, and therefore

$$f(x)f(y)f(x) = f(x).$$

For each $\mathfrak{p}$ in $\operatorname{Spec}(R)$, by applying the coordinate projection $\pi_{\mathfrak{p}}$, we obtain

$$(x + \mathfrak{p})(y + \mathfrak{p})(x + \mathfrak{p}) = x + \mathfrak{p}.$$

If $x + \mathfrak{p} \neq 0$, we can cancel it, obtaining

$$y + \mathfrak{p} = (x + \mathfrak{p})^{-1} = (x + \mathfrak{p})^*.$$

However, possibly $x + \mathfrak{p} = 0$, while $y + \mathfrak{p} \neq 0$, so that $f(y) \neq f(x)^*$. In this case, letting $z = yxy$, we have

$$xzx = xyxyx = xyx = x, \quad zxz = yxyxyxy = yxyxy = yxy = z.$$

In short, $xzx = x$ and $zxz = z$. Then

$$x \in \mathfrak{p} \iff z \in \mathfrak{p}, \qquad x \notin \mathfrak{p} \implies z + \mathfrak{p} = (x + \mathfrak{p})^{-1},$$

so $(x + \mathfrak{p})^* = z + \mathfrak{p}$. Thus $f(z) = f(x)^*$. □

4.7. Products of spaces

Being a group by Theorem 56, the direct product of a family $(A_i : i \in \Omega)$ of groups is nonempty. Indeed, it contains the identity $(1^{A_i} : i \in \Omega)$. (This is true, even if Ω is empty.) If each A_i is merely a nonempty *set,* we define their **Cartesian product** in the same way as a direct product of groups or rings. However, it is not obvious that the product of a family of nonempty sets will itself be nonempty.

Theorem 129 (Cartesian Product). *By the Axiom of Choice, the product of an indexed family of nonempty sets is nonempty.* **AC**

Theorem 130. *The Cartesian Product Theorem implies the Axiom of Choice.*

If now $\mathscr{A}$ is a family $(A_i : i \in \Omega)$ of topological spaces, the **product topology** on the Cartesian product $\prod \mathscr{A}$ is the weakest topology in which the coordinate projections are **continuous.** This means that for every j in Ω, for every closed subset F of A_j, the subset $\{x \in \prod \mathscr{A} : x_j \in A_j\}$ of $\prod \mathscr{A}$ must be closed; and such subsets compose a sub-basis of the product topology. Thus all finite unions of such sets are closed, and such sets compose a *basis* of the product topology, so that all intersections of arbitrary collections of such unions are closed, and no other subsets of $\prod \mathscr{A}$ are closed.

Theorem 131 (Tychonoff). *By the Axiom of Choice, the product of a family of nonempty compact topological spaces is nonempty and compact in the product topology.* **AC**

Proof. Suppose $\mathscr{A}$ is a family $(A_i\colon i \in \Omega)$ of nonempty compact topological spaces, and $\mathscr{X}$ is a family of closed subsets of $\prod \mathscr{A}$ whose every finite subset has nonempty intersection. We want to show $\bigcap \mathscr{X} \neq \varnothing$. Each element of $\mathscr{X}$ is the intersection of sets belonging to the basis just described; so we may assume that each element of $\mathscr{X}$ belongs to this basis. Moreover, suppose $F \in \mathscr{X}$, and F is a union $F_0 \cup \cdots \cup F_n$ of sets from the sub-basis just described. Then for some i in $n+1$, F_i has nonempty intersection with each element of $\mathscr{X}$. In this case, we can replace F with F_i in $\mathscr{X}$.

Using the Axiom of Choice then, we may assume that every element of $\mathscr{X}$ is a sub-basic set. One way to spell this out is as follows (we shall see a neater way later). We have noted that the topology on $\prod \mathscr{A}$ has, as a sub-basis, the sets ${\pi_i}^{-1}[F]$, where $i \in \Omega$ and F is a closed subset of A_i. Then the topology on $\prod \mathscr{A}$ has, as a *basis,* the sets

$$\pi_{\sigma(0)}{}^{-1}[F_0] \cup \cdots \cup \pi_{\sigma(n)}{}^{-1}[F_n],$$

where $n \in \omega$, and σ is an *injective* function from $n+1$ into Ω, and each F_i is a closed subset of $A_{\sigma(i)}$. Now consider the family of subsets $\mathscr{Y}$ of $\prod \mathscr{A}$ that have the finite intersection property, while each element is either an element of $\mathscr{X}$ or else an element of a finite set of sub-basic sets whose union is in $\mathscr{X}$. The family is ordered in an obvious way, so that $\mathscr{Y}_0 < \mathscr{Y}_1$ if and only if each element of $\mathscr{Y}_1$ is either an element of $\mathscr{Y}_0$ or else an element of a finite set of sub-basic sets whose union is in $\mathscr{Y}_0$. Suppose we are given a chain of the family, and $\mathscr{Y}$ belongs to the chain. Then the chain has an upper bound consisting of each sub-basic set that belongs to $\mathscr{Y}$, as well as, for each union $F_0 \cup \cdots \cup F_n$ of sub-basic sets in $\mathscr{Y}$, either this union itself, if it belongs to every member of the chain, or else F_i, if this belongs to some member of the chain. By Zorn's Lemma (more precisely, its corollary), our family has a maximal element. By what we noted, this maximal element must consist precisely of sub-basic sets.

We may thus assume that every set in $\mathscr{X}$ is a nonempty sub-basic closed set. Then, by the compactness of each A_i, we may assume that, for some indexed family $(F_i\colon i \in \Omega)$, each F_i being a nonempty subset of A_i,

$$\mathscr{X} = \{{\pi_i}^{-1}[F_i]\colon i \in \Omega\}.$$

Then $\bigcap \mathscr{X} = \prod_{i\in\Omega} F_i$, which is nonempty by the Cartesian Product Theorem. □

The converse was published by Kelley in 1950 [37]:[10]

Theorem 132. *The Tychonoff Theorem implies the Axiom of Choice.*

Proof. Let $(A_i\colon i \in \Omega)$ be an indexed family of nonempty sets, and let b not belong to any A_i. If $i \in \Omega$, let

$$\tau_i = \{\varnothing, A_i, A_i \cup \{b\}\};$$

this is a topology on $A_i \cup \{b\}$. Every finite subset of the family

$$\left\{{\pi_i}^{-1}[A_i]\colon i \in \Omega\right\}$$

of closed subsets $\prod_{i\in\Omega}(A_i \cup \{b\})$ has nonempty intersection. Indeed, by induction, for every n in ω, for every subset Ω_0 of size n, we have

$$\prod_{i\in\Omega_0} A_i \neq \varnothing,$$

so the product contains some $(a_i\colon i \in \Omega)$. Then $\bigcap_{i\in\Omega_0} {\pi_i}^{-1}[A_i]$ contains c, where

$$c_i = \begin{cases} a_i, & \text{if } i \in \Omega_0, \\ b, & \text{if } i \in \Omega \smallsetminus \Omega_0. \end{cases}$$

By the Tychonoff Theorem, $\bigcap_{i\in\Omega} {\pi_i}^{-1}[A_i]$ must be nonempty; but this intersection is $\prod_{i\in\Omega} A_i$. □

[10] Actually Kelley's proof had an error, which however is easily corrected, as Ło's and Ryll-Nardzewski observed in 1951 [42]. Kelley's error reduced his claim to being that the Tychonoff Theorem for sets in which the one-element sets compose a basis for the topology implies the Axiom of Choice. Schechter [52] shows that this hypothesis is equivalent to the Boolean Prime Ideal Theorem.

5. Model theory without the Prime Ideal Theorem

Model theory was described on page 48 as the study of structures as such. More precisely, model theory takes into account the *logic* in which the properties of structures are stated and derived. Usually this logic is *first order* logic, which means its variables stand for individual elements of the universe of a structure. Second order logic has variables for relations on the universe. For example, the induction axiom for the natural numbers (page 36) is a second order statement, when considered as a statement about elements and subsets of $\mathbb{N}$. When considered as a part of Theorem 25 (page 43), where it is a statement about all sets and in particular the set ω, it is first order. Indeed, for set theory itself, there is no distinction between first and second order.

In a logic, certain strings of symbols are called *formulas,* and some formulas can be combined to make other formulas. If only finitely many formulas can ever be combined to make new formulas, the logic is *finitary.* First order logic is implicitly finitary. (By "implicitly" I mean that the finitary aspect is not made explicit in the name "first order". One can develop infinitary logics in which variables stand only for individuals.)

The most important theorem of first order model theory is the Compactness Theorem. Its proof needs the Prime Ideal Theorem. However, a lot of model theory can be developed without Compactness. Indeed, in Hodges's encyclopedic volume *Model Theory* [31], Compactness is introduced only in the sixth of the 12 chapters.

The present chapter of the present text develops what we shall need of model theory that does not require the Prime Ideal Theorem. Compactness and related results that do require the Prime Ideal Theorem and even the full Axiom of Choice will be established in the next chapter.

5.1. Logic

For study of arbitrary structures as defined in §2.6 (page 45), we now generalize the logic developed for set theory in §2.2 (page 17). This logic was based on the signature whose only symbol is $\in$. However, we allowed constants standing for arbitrary sets: we had to do this in order to define truth and falsity of sentences (as on page 20).

Likewise, given a structure $\mathfrak{A}$ with signature $\mathscr{S}$, we may augment $\mathscr{S}$ with a constant for each element of A. If we denote the augmented signature by $\mathscr{S}(A)$, then we can *expand* $\mathfrak{A}$ (in the sense of page 48) in an obvious way to a structure denoted by

$$\mathfrak{A}_A,$$

whose signature is $\mathscr{S}(A)$: each a in A, considered as a constant in $\mathscr{S}(A)$, is interpreted in $\mathfrak{A}_A$ as the element a of A.

5.1.1. Terms

In the logic of set theory, a *term* is a variable or constant. In the logic of an arbitrary signature $\mathscr{S}$, there might be n-ary operation symbols for positive n, and so the definition of **term** is broader and is made recursively. We start with a countably infinite set of variables.

1. Each variable is a term of $\mathscr{S}$.
2. For each n in ω, if F is an n-ary operation symbol in $\mathscr{S}$, and $(t_i \colon i \in n)$ is an indexed family of terms of $\mathscr{S}$, then the string

$$Ft_0 \cdots t_{n-1}$$

is a term of $\mathscr{S}$.

As a special case of the second condition, every constant in $\mathscr{S}$ as a term. Thus, if we omit the first condition, we still have a nontrivial definition, at least if $\mathscr{S}$ contains constants. What we have then is the definition of a **closed term:** a term without variables.

There is an analogue of the lemma on page 21:

Theorem 133. *No proper initial segment of a term is a term.*

Then we obtain the analogue of Theorem 1 (page 20):

Theorem 134 (Unique Readability)**.** *A term can be constructed in only one way: If $Ft_0 \cdots t_{n-1}$ and $Fu_0 \cdots u_{m-1}$ are the same term, where the t_i and u_i are terms, then $n = m$, and each t_i is the same term as u_i.*

Informally, if F is a binary operation symbol, and G is a singulary operation symbol, and t and u are terms, then for the terms Ftu and Gt we may write, respectively,

$$(t\,F\,u), \qquad\qquad t^G.$$

If t is a closed term of $\mathscr{S}$, and $\mathfrak{A} \in \mathbf{Str}_{\mathscr{S}}$, then t has an interpretation

$$t^{\mathfrak{A}}$$

in $\mathfrak{A}$, and this interpretation is an element of A. The definition is recursive, like the definition of closed terms themselves; and the definition is justified by Theorem 134. If t is $Ft_0 \cdots t_{n-1}$, where each t_i is a closed term, then

$$t^{\mathfrak{A}} = F^{\mathfrak{A}}({t_0}^{\mathfrak{A}}, \dots, {t_{n-1}}^{\mathfrak{A}}).$$

This covers the special case where t is a constant, so that $n = 0$.

We define the interpretation of an arbitrary term t of $\mathscr{S}$ as follows. Let us denote by

$$\operatorname{var}(t)$$

the set of variables occurring in t. For each $\mathfrak{A}$ in $\mathbf{Str}_{\mathscr{S}}$, if $\boldsymbol{a}$ is the tuple $(a_x \colon x \in \operatorname{var}(t))$ in $A^{\operatorname{var}(t)}$, we obtain the closed term

$$t(\boldsymbol{a})$$

of $\mathscr{S}(A)$ by replacing each occurrence of x in t with the constant a_x, for each x in $\operatorname{var}(t)$. Then we can denote by

$$t^{\mathfrak{A}}$$

the function $\boldsymbol{a} \mapsto t(\boldsymbol{a})^{\mathfrak{A}_A}$ from $A^{\operatorname{var}(t)}$ to A.

We defined *homomorphisms* on page 47. Given the recursive definition of terms, we have the following by induction:

Theorem 135. *Suppose $\mathfrak{A}$ and $\mathfrak{B}$ are in $\mathbf{Str}_{\mathscr{S}}$, If $h\colon \mathfrak{A} \to \mathfrak{B}$, then for each term t of $\mathscr{S}$ and each $\boldsymbol{a}$ from $A^{\operatorname{var}(t)}$,*

$$h(t^{\mathfrak{A}}(\boldsymbol{a})) = t^{\mathfrak{B}}(h(\boldsymbol{a})).$$

The converse fails if $\mathscr{S}$ has predicates. For this case, we consider *atomic formulas.*

5.1.2. Atomic formulas

In our logic of set theory, an *atomic formula* is just a string $t \in u$, where t and u are terms. We introduced the expression $t = u$ as an abbreviation of a certain formula. However, we shall now count this as one of the atomic formulas. Thus, for an arbitrary signature $\mathscr{S}$, the **atomic formulas** are of two kinds:

$$t = u,$$

where t and u are terms of $\mathscr{S}$, and

$$Rt_0 \cdots t_{n-1},$$

for each n in ω, where $(t_i \colon i < n)$ is an indexed family of terms of $\mathscr{S}$, and R is an n-ary predicate of $\mathscr{S}$. If R is a binary predicate of $\mathscr{S}$, and t and u are terms, then for the term Rtu we may write

$$t \mathrel{R} u.$$

It is an obvious consequence of Theorem 133 that atomic formulas are uniquely readable.

An atomic formula in which no variable occurs—an atomic formula in which the terms are closed—is an **atomic sentence.** If $\mathfrak{A} \in \mathscr{S}$, then every atomic sentence of $\mathscr{S}(A)$ is **true** or **false** in $\mathfrak{A}$ according to the obvious definition:

1. $t = u$ is true in $\mathfrak{A}$ if and only if

$$t^{\mathfrak{A}_A} = u^{\mathfrak{A}_A}.$$

2. $Rt_0 \cdots t_{n-1}$ is true in $\mathfrak{A}$ if and only if

$$({t_0}^{\mathfrak{A}_A}, \dots, {t_{n-1}}^{\mathfrak{A}_A}) \in R^{\mathfrak{A}}.$$

If σ is an atomic sentence that is true in $\mathfrak{A}$, we may write

$$\mathfrak{A} \vDash \sigma.$$

If φ is an atomic formula of $\mathscr{S}$, then, as with terms, we can denote by

$$\operatorname{var}(\varphi)$$

the set of variables occurring in φ; and then if $\boldsymbol{a}$ is the tuple $(a_x \colon x \in \operatorname{var}(\varphi))$ in $A^{\operatorname{var}(\varphi)}$, we can denote by

$$\varphi(\boldsymbol{a})$$

the result of replacing each occurrence of x in φ with a_x, for each x in $\operatorname{var}(\varphi)$. Now we have a convertible version of Theorem 135:

Theorem 136. *Suppose $\mathfrak{A}$ and $\mathfrak{B}$ are in $\mathbf{Str}_{\mathscr{S}}$, and $h \colon A \to B$.*

1. *h is a homomorphism from $\mathfrak{A}$ to $\mathfrak{B}$ if and only if, for all atomic formulas φ of $\mathscr{S}$, for all $\boldsymbol{a}$ in $A^{\operatorname{var}(\varphi)}$,*

$$\mathfrak{A} \vDash \varphi(\boldsymbol{a}) \implies \mathfrak{B} \vDash \varphi(h(\boldsymbol{a})).$$

2. *h is an embedding of $\mathfrak{A}$ in $\mathfrak{B}$ if and only if, for all atomic formulas φ of $\mathscr{S}$, for all $\boldsymbol{a}$ in $A^{\operatorname{var}(\varphi)}$,*

$$\mathfrak{A} \vDash \varphi(\boldsymbol{a}) \iff \mathfrak{B} \vDash \varphi(h(\boldsymbol{a})).$$

5.1.3. Formulas

Now arbitrary **formulas** are built up recursively, precisely as in the logic of set theory on page 17:

1. Atomic formulas are formulas.
2. If φ is a formula, then so is its negation $\neg\varphi$.
3. If φ and ψ are formulas, then so are

   a) the disjunction $(\varphi \vee \psi)$,
   b) the conjunction $(\varphi \wedge \psi)$,
   c) the implication $(\varphi \Rightarrow \psi)$, and
   d) the equivalence $(\varphi \Leftrightarrow \psi)$.
4. If φ is a formula and x is variable, then
   a) the instantiation $\exists x\ \varphi$ and
   b) the generalization $\forall x\ \varphi$

   are both formulas.

Again we have:

Theorem 137 (Unique Readability)**.** *A given formula can be built up from atomic formulas in only one way.*

Now the set $\mathrm{fv}(\varphi)$ of **free variables** of a formula φ can be defined recursively:

1. If φ is atomic, then $\mathrm{fv}(\varphi) = \mathrm{var}(\varphi)$.
2. $\mathrm{fv}(\neg\varphi) = \mathrm{fv}(\varphi)$.
3. $\mathrm{fv}((\varphi * \psi)) = \mathrm{fv}(\varphi) \cup \mathrm{fv}(\psi)$.
4. $\mathrm{fv}(\exists x\ \varphi) = \mathrm{fv}(\forall x\ \varphi) = \mathrm{fv}(\varphi) \smallsetminus \{x\}$.

A **sentence** is a formula with no free variables.

If x is a variable and t is a term, we define recursively the result φ^x_t of replacing each **free occurrence** of x in φ with t:

1. If φ is atomic, then φ^x_t is just the result of replacing *every* occurrence of x in φ with t.
2. $(\neg\varphi)^x_t$ is $\neg(\varphi^x_t)$.
3. $(\varphi * \psi)^x_t$ is $(\varphi^x_t * \psi^x_t)$.
4. If x is not y, then $(\exists y\ \varphi)^x_t$ is $\exists y\ \varphi^x_t$, and $(\forall y\ \varphi)^x_t$ is $\forall y\ \varphi^x_t$.
5. $(\exists x\ \varphi)^x_t$ is $\exists x\ \varphi$, and $(\forall x\ \varphi)^x_t$ is $\forall x\ \varphi$.

If $i \mapsto x(i)$ is a bijection from some n in $\boldsymbol{\omega}$ to $\mathrm{fv}(\varphi)$, and $\boldsymbol{a} \in A^{\mathrm{fv}(\varphi)}$, then we can define

$$\varphi(\boldsymbol{a})$$

as

$$(\dots(\varphi^{x(0)}_{a_{x(0)}})^{x(1)}_{a_{x(1)}}\dots)^{x(n-1)}_{a_{x(n-1)}}.$$

This definition is independent of the particular choice of the bijection $i \mapsto x(i)$. (This would not be true if $\boldsymbol{a}$ were a tuple of arbitrary terms.)

Now we can define **truth** and **falsity** of sentences in structures. That is, let $\mathfrak{A} \in \mathbf{Str}_{\mathscr{S}}$. For every formula φ of $\mathscr{S}$, for every $\boldsymbol{a}$ in $A^{\mathrm{fv}(\varphi)}$, we define whether $\varphi(\boldsymbol{a})$ is true or false in $\mathfrak{A}$. If φ is atomic, we have done this. We proceed as on page 20:

1. $\neg\varphi(\boldsymbol{a})$ is true in $\mathfrak{A}$ if and only if $\varphi(\boldsymbol{a})$ is false in $\mathfrak{A}$.
2. The truth or falsity of $(\varphi * \psi)(\boldsymbol{a})$ in $\mathfrak{A}$ depends on the truth or falsity of $\varphi(\boldsymbol{a})$ and $\psi(\boldsymbol{a})$ in $\mathfrak{A}$ according to the usual rules of propositional logic.
3. $(\exists x\ \varphi)(\boldsymbol{a})$ is true in $\mathfrak{A}$ if and only if, for some b in A, $\varphi^x_b(\boldsymbol{a})$ is true in $\mathfrak{A}$.
4. $(\forall x\ \varphi)(\boldsymbol{a})$ is true in $\mathfrak{A}$ if and only if, for all b in A, $\varphi^x_b(\boldsymbol{a})$ is true in $\mathfrak{A}$.

Again, if a sentence σ is true in $\mathfrak{A}$, we write

$$\mathfrak{A} \vDash \sigma. \tag{5.1}$$

5.2. Theories and models

The set of all sentences of a signature $\mathscr{S}$ can be denoted by

$$\mathrm{Sen}(\mathscr{S}).$$

This is the universe of an algebra

$$(\mathrm{Sen}(\mathscr{S}), \neg, \vee, \wedge).$$

The relation of **truth,** symbolized as in (5.1) by $\vDash$, is a relation between $\mathbf{Str}_{\mathscr{S}}$ and $\mathrm{Sen}(\mathscr{S})$. This relation establishes a *Galois correspondence* just as in Theorem 105 (page 110), even though $\mathbf{Str}_{\mathscr{S}}$ is a proper class. With respect to this Galois correspondence, the closed subsets of $\mathrm{Sen}(\mathscr{S})$ are called **theories.** The class $\mathbf{Str}_{\mathscr{S}}$ has closed *subclasses,* called **elementary classes.** The polarities constituting the Galois correspondence can be written respectively as

$$\mathcal{K} \mapsto \mathrm{Th}(\mathcal{K}), \qquad\qquad \Gamma \mapsto \mathbf{Mod}(\Gamma).$$

Here $\mathrm{Th}(\mathcal{K})$ is the **theory of** the class $\mathcal{K}$ of structures, and the elementary class $\mathbf{Mod}(\Gamma)$ is in particular the class of **models of** the set Γ of sentences of $\mathscr{S}$.

We may modify the notation and terminology in an obvious way, so that if $\mathcal{K} = \{\mathfrak{A}\}$, and $\Gamma = \{\sigma\}$, then

$$\mathrm{Th}(\mathfrak{A}) = \mathrm{Th}(\mathcal{K}), \qquad\qquad \mathbf{Mod}(\sigma) = \mathbf{Mod}(\Gamma),$$

and these are respectively the **theory of** $\mathfrak{A}$ and the class of **models of** σ. Then

$$\begin{aligned}\mathrm{Th}(\mathfrak{A}) &= \{\sigma \in \mathrm{Sen}(\mathscr{S}) \colon \mathfrak{A} \vDash \sigma\},\\ \mathbf{Mod}(\sigma) &= \{\mathfrak{A} \in \mathbf{Str}_{\mathscr{S}} \colon \mathfrak{A} \vDash \sigma\}.\end{aligned}$$

For arbitrary subclasses $\mathcal{K}$ of $\mathbf{Str}_{\mathscr{S}}$ and subsets Γ of $\mathrm{Sen}(\mathscr{S})$, we now have

$$\mathrm{Th}(\mathcal{K}) = \bigcap_{\mathfrak{A}\in\mathcal{K}} \mathrm{Th}(\mathfrak{A}), \qquad \mathbf{Mod}(\Gamma) = \bigcap_{\sigma\in\Gamma} \mathbf{Mod}(\sigma). \tag{5.2}$$

If $\mathfrak{A} \in \mathbf{Mod}(\Gamma)$, that is, if $\mathfrak{A}$ is a model of Γ, we may write

$$\mathfrak{A} \vDash \Gamma.$$

Then also[1]

$$\mathbf{Mod}(\Gamma) = \{\mathfrak{A} \in \mathbf{Str}_{\mathscr{S}} \colon \mathfrak{A} \vDash \Gamma\}. \tag{5.3}$$

An arbitrary theory is called a **complete theory** if, for every sentence σ of its signature, the theory contains either σ or its negation $\neg\sigma$, but not both.

Theorem 138. *Let $\mathscr{S}$ be a signature.*

1. *The only theory of $\mathscr{S}$ that, for some σ in* $\mathrm{Sen}(\mathscr{S})$, *contains both σ and $\neg\sigma$ is* $\mathrm{Sen}(\mathscr{S})$ *itself, which is* $\mathrm{Th}(\varnothing)$.
2. *Every complete theory of $\mathscr{S}$ is* $\mathrm{Th}(\mathfrak{A})$ *for some $\mathfrak{A}$ in* $\mathbf{Str}_{\mathscr{S}}$.

[1] We could also write $\mathcal{K} \vDash \sigma$ instead of $\sigma \in \mathrm{Th}(\mathcal{K})$, so that $\mathrm{Th}(\mathcal{K}) = \{\sigma \in \mathrm{Sen}(\mathscr{S}) \colon \mathcal{K} \vDash \sigma\}$; but we shall not actually use this notation.

3. The elementary classes of $\mathscr{S}$ compose a topology on $\mathbf{Str}_{\mathscr{S}}$.

Proof. 1. If T is a theory containing both σ and $\neg\sigma$, then T has no models, and so T must be $\mathrm{Th}(\varnothing)$.

2. If T is complete, then by the first part T cannot be $\mathrm{Th}(\varnothing)$, so it has a model $\mathfrak{A}$, and $T \subseteq \mathrm{Th}(\mathfrak{A})$. Since T is complete, this inclusion must be an equation.

3. By Theorem 112 (page 113), since

$$\varnothing = \mathbf{Mod}(\exists x\ x \neq x), \qquad \mathbf{Mod}(\sigma) \cup \mathbf{Mod}(\tau) = \mathbf{Mod}(\sigma \lor \tau),$$

the classes $\mathbf{Mod}(\sigma)$ compose a basis of a topology on $\mathbf{Str}_{\mathscr{S}}$. □

5.3. Elementary equivalence

Given a signature $\mathscr{S}$, in $\mathbf{Str}_{\mathscr{S}}$ we define

$$\mathfrak{A} \equiv \mathfrak{B} \iff \mathrm{Th}(\mathfrak{A}) = \mathrm{Th}(\mathfrak{B}).$$

The relation $\equiv$ is called **elementary equivalence.** If $\mathfrak{A}$ and $\mathfrak{B}$ are isomorphic, then they are elementarily equivalent. We shall see that the converse fails. Indeed, what makes model theory interesting is that non-isomorphic structures can be elementarily equivalent. Meanwhile, recalling the notion of topological indistinguishability from page 114, we have the following.

Theorem 139. *On $\mathbf{Str}_{\mathscr{S}}$, the relation of topological indistinguishability (with respect to the topology consisting of the elementary classes) is just elementary equivalence.*

5.3.1. Kolmogorov quotients

Suppose A and B are topological spaces, and f is a function from A to B. Then f is called

- **continuous,** if $f^{-1}[Y]$ is closed for every closed subset Y of B;
- **closed,** if $f[X]$ is closed for every closed subset X of A.

A **homeomorphism** is a continuous bijection with continuous inverse. Letting $\sim$ be the relation of topological indistinguishability on A, we can give the quotient $A/{\sim}$ the **quotient topology,** so that a subset $\{x^{\sim}\colon x \in X\}$ of $A/{\sim}$ is closed if and only if its union $\bigcup_{x\in X} x^{\sim}$ is a closed subset of A.

Theorem 140. *Let A be a topological space.*

1. *The quotient topology on $A/{\sim}$ is indeed a topology, even a Kolmogorov topology.*
2. *The quotient map $x \mapsto x^{\sim}$ from A to $A/{\sim}$ is surjective, continuous, and closed.*
3. *If B is a Kolmogorov space, and f is a continuous function from A to B, then there is a unique function h from $A/{\sim}$ to B such that, for all x in A, $h(x^{\sim}) = f(x)$.*

Suppose now f is a surjective continuous closed function from A to a Kolmogorov space B, and for every Kolmogorov space C and every continuous function g from A to C, there is a unique continuous function h from B to C such that

$$g = h \circ f.$$

See Figure 5.1. Then B is a **Kolmogorov quotient** of A with respect to f.

$$\begin{tikzcd}
A \arrow[r, "g"] \arrow[d, "f"'] & C \\
B \arrow[ur, dashed, "\exists!\, h"'] &
\end{tikzcd}$$

Figure 5.1.: Kolmogorov quotient

Theorem 141. *If B_0 and B_1 are Kolmogorov quotients of A with respect to f_0 and f_1 respectively, then the unique homomorphism h_0 from B_0 to B_1 such that $f_1 = h_0 \circ f_0$ is a homeomorphism onto B_1, its inverse being the unique homomorphism h_1 from B_1 to B_0 such that $f_0 = h_1 \circ f_1$.*

Proof. See Figure 5.2. The composition $h_1 \circ h_0$ must be the unique homomorphism h from B_0 to itself such that $f_0 = h \circ f_0$. Since id_{B_0} is such a homomorphism, we have

$$h_1 \circ h_0 = \mathrm{id}_{B_0} .$$

By symmetry, $h_0 \circ h_1 = \mathrm{id}_{B_1}$. □

$$\begin{tikzcd}
 & A \arrow[dl, "f_0"'] \arrow[dr, "f_1"] & \\
B_0 \arrow[rr, bend left, "h_0"] & & B_1 \arrow[ll, bend left, "h_1"]
\end{tikzcd}$$

Figure 5.2.: Two Kolmogorov quotients

By the theorem, any two Kolmogorov quotients of a space are *equivalent.*

Theorem 142. *If f is a continuous, closed, surjective function from A onto B, and for all x and y in A,*

$$x \sim y \iff f(x) = f(y),$$

then B is a Kolmogorov quotient of A with respect to f.

Proof. Suppose $f(x)$ and $f(y)$ are topologically equivalent. Since f is closed, we have $x \sim y$, and therefore $f(x) = f(y)$. Thus B is Kolmogorov. There is a well-defined map $x^{\sim}x \mapsto f(x)$ from $A/{\sim}$ to B. This map continuous, closed, and surjective onto B; so it is a homeomorphism onto B. □

$$\begin{tikzcd}
\mathbf{Str}_{\mathscr{S}} \arrow[r, leftrightsquigarrow, "\models"] \arrow[d, "\mathfrak{A}\mapsto\mathrm{Th}(\mathfrak{A})"'] & \mathrm{Sen}(\mathscr{S}) \arrow[dl, leftrightsquigarrow, "\ni"] \\
\mathrm{S}_0(\mathscr{S}) &
\end{tikzcd}$$

Figure 5.3.: Kolmogorov quotient of $\mathbf{Str}_{\mathscr{S}}$

5.3.2. The space of complete theories

Let us denote the set of complete theories of $\mathscr{S}$ by

$$\mathrm{S}_0(\mathscr{S}).$$

(The subscript 0 indicates that the formulas in a theory have no free variables.) The following begins to resemble Theorem 113 (page 114):

Theorem 143. *For every signature $\mathscr{S}$, with respect to the relation $\in$ between $\mathrm{Sen}(\mathscr{S})$ and $\mathrm{S}_0(\mathscr{S})$,*

1) *the closed subsets of $\mathrm{Sen}(\mathscr{S})$ are precisely the theories of $\mathscr{S}$;*
2) *the closed subsets of $\mathrm{S}_0(\mathscr{S})$ compose a Hausdorff topology;*
3) *the map $\mathfrak{A} \mapsto \mathrm{Th}(\mathfrak{A})$ from $\mathbf{Str}_{\mathscr{S}}$ to $\mathrm{S}_0(\mathscr{S})$ is a continuous surjection, and $\mathrm{S}_0(\mathscr{S})$ is a Kolmogorov quotient of $\mathbf{Str}_{\mathscr{S}}$ with respect to this map.*

The situation of the theorem might be depicted as in Figure 5.3.

The *Compactness Theorem* is that the topology on $\mathrm{S}_0(\mathscr{S})$ is compact. In fact we are going to be able to replace $\mathrm{Sen}(\mathscr{S})$ with a Boolean ring R such that $\mathrm{S}_0(\mathscr{S})$ is homeomorphic to $\mathrm{Spec}(R)$. But this will take some work, which in one approach involves ultraproducts. The Boolean ring will be best thought of as a *Boolean algebra* as developed in the next section.

5.4. Boolean Algebras

We showed in Corollary 80.1 (page 88) that, for any set Ω, the power set $\mathscr{P}(\Omega)$ is the universe of a Boolean ring $(\mathscr{P}(\Omega), \varnothing, \vartriangle, \Omega, \cap)$. By the Stone Representation Theorem (page 119), every Boolean ring embeds in such a ring.

The structure $(\mathscr{P}(\Omega), \varnothing, \Omega, {}^{\mathrm{c}}, \cup, \cap)$ is an example of a *Boolean algebra.* Here ${}^{\mathrm{c}}$ is the singulary operation $X \mapsto \Omega \smallsetminus X$.

5.4.1. Abstract Boolean algebras

Abstractly considered, a **Boolean algebra** is a structure

$$(B, \bot, \top, {}^{-}, \vee, \wedge),$$

meeting the following conditions.

1. The binary operations $\vee$ and $\wedge$ are **commutative:**

$$x \vee y = y \vee x, \qquad\qquad x \wedge y = y \wedge x.$$

2. The elements $\bot$ and $\top$ are **identities** for $\vee$ and $\wedge$ respectively:

$$x \vee \bot = x, \qquad\qquad x \wedge \top = x.$$

3. $\vee$ and $\wedge$ are mutually **distributive:**

$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z), \quad x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z).$$

4. The element $\bar{x}$ is a **complement** of x:

$$x \vee \bar{x} = \top, \qquad\qquad x \wedge \bar{x} = \bot.$$

These axioms are symmetrical in the sense that, if $(B, \bot, \top, {}^{-}, \vee, \wedge)$ is a Boolean algebra, then so is $(B, \top, \bot, {}^{-}, \wedge, \vee)$. Then the latter algebra can be called the **dual** of the former. Just to give them names, we may say that $x \vee y$ is the **join** of x and y, and $x \wedge y$ is their **meet.**

The identities in the following theorem are sometimes given as additional axioms for Boolean algebras; but Huntington [36, 35] shows that the axioms above are sufficient.

Theorem 144. *In any Boolean algebra:*

$$x \vee x = x, \qquad x \wedge x = x, \tag{5.4}$$

$$x \vee \top = \top, \qquad x \wedge \bot = \bot, \tag{5.5}$$

$$x \vee (x \wedge y) = x, \qquad x \wedge (x \vee y) = x, \tag{5.6}$$

$$\bar{\bar{x}} = x, \tag{5.7}$$

$$\overline{x \vee y} = \bar{x} \wedge \bar{y}, \qquad \overline{x \wedge y} = \bar{x} \vee \bar{y}, \tag{5.8}$$

$$(x \vee y) \vee z = x \vee (y \vee z), \qquad (x \wedge y) \wedge z = x \wedge (y \wedge z). \tag{5.9}$$

Proof. By symmetry, it is enough to establish one of each pair of identities. We do this in turn. For (5.4) and (5.5), we have

$$\begin{aligned} x \vee x &= (x \vee x) \wedge \top \\ &= (x \vee x) \wedge (x \vee \bar{x}) \\ &= x \vee (x \wedge \bar{x}) \\ &= x \vee \bot \\ &= x, \end{aligned} \qquad \begin{aligned} x \vee \top &= (x \vee \top) \wedge \top \\ &= (x \vee \top) \wedge (x \vee \bar{x}) \\ &= x \vee (\top \wedge \bar{x}) \\ &= x \vee \bar{x} \\ &= \top, \end{aligned}$$

and for (5.6) and (5.7),

$$\begin{aligned} x \vee (x \wedge y) &= (x \wedge \top) \vee (x \wedge y) \\ &= x \wedge (\top \vee y) \\ &= x \wedge \top \\ &= x, \end{aligned} \qquad \begin{aligned} \bar{\bar{x}} &= \top \wedge \bar{\bar{x}} \\ &= (\bar{x} \vee x) \wedge \bar{\bar{x}} \\ &= (\bar{x} \wedge \bar{\bar{x}}) \vee (x \wedge \bar{\bar{x}}) \\ &= \bot \vee (x \wedge \bar{\bar{x}}) \\ &= (x \wedge \bar{x}) \vee (x \wedge \bar{\bar{x}}) \\ &= x \wedge (\bar{x} \vee \bar{\bar{x}}) \\ &= x \wedge \top \\ &= x. \end{aligned}$$

In showing (5.7), what we show is that the two complements of $\bar{x}$, namely $\bar{\bar{x}}$ and x, are equal. In the same way, any two complements of an element of a Boolean algebra must be equal. We shall use this to establish (5.8). First we establish a special case of associativity:

$$x \vee (\bar{x} \vee y) = \top \wedge ((x \vee (\bar{x} \vee y)))$$

$$\begin{aligned}&= (x \vee \bar{x}) \wedge ((x \vee (\bar{x} \vee y)))\\&= x \vee (\bar{x} \wedge (\bar{x} \vee y))\\&= x \vee \bar{x}\\&= \top,\end{aligned}$$

and likewise $x \wedge (\bar{x} \wedge y) = \bot$. Then

$$\begin{aligned}(x \vee y) \vee (\bar{x} \wedge \bar{y}) &= ((x \vee y) \vee \bar{x}) \wedge ((x \vee y) \vee \bar{y})\\&= \top \wedge \top\\&= \top,\\(x \vee y) \wedge (\bar{x} \wedge \bar{y}) &= (x \wedge (\bar{x} \wedge \bar{y})) \vee (y \wedge (\bar{x} \wedge \bar{y}))\\&= \bot \vee \bot\\&= \bot,\end{aligned}$$

so by uniqueness of complements we must have (5.8). Finally, let T stand for $(x \vee y) \vee z$, and U for $x \vee (y \vee z)$. Then

$$U \vee \bar{T} = U \vee ((\bar{x} \wedge \bar{y}) \wedge \bar{z}) = ((U \vee \bar{x}) \wedge (U \vee \bar{y})) \wedge (U \vee \bar{z}).$$

But the factors here are all $\top$, since

$$\begin{aligned}U \vee \bar{x} &= (x \vee (y \vee z)) \vee \bar{x} = \top,\\U \vee \bar{y} &= (U \vee \bar{y}) \wedge \top\\&= (U \vee \bar{y}) \wedge (y \vee \bar{y})\\&= (U \wedge y) \vee \bar{y}\\&= ((x \vee (y \vee z)) \wedge y) \vee \bar{y}\\&= ((x \wedge y) \vee ((y \vee z) \wedge y)) \vee \bar{y}\\&= ((x \wedge y) \vee y) \vee \bar{y}\\&= y \vee \bar{y}\\&= \top,\end{aligned}$$

and likewise $U \vee \bar{z} = \top$. Thus

$$U \vee \bar{T} = (\top \wedge \top) \wedge \top = \top \wedge \top = \top.$$

Dually, $U \wedge \bar{T} = \bot$. Then $U = \bar{\bar{T}} = T$, that is, (5.9). □

Huntington shows also that each of the eight axioms for Boolean algebras is logically independent from the others. He does this by exhibiting, for each of the axioms, a structure in which the axiom is false, but the remaining seven axioms are true. In each case, the universe of the structure has two elements.[2]

5.4.2. Boolean operations

Theorem 145. *Boolean algebras and Boolean rings are the same thing in the following sense:*

1. *If $\mathfrak{A}$ is a Boolean algebra $(B, \bot, \top, {}^-, \vee, \wedge)$, and we define*

$$x + y = (x \wedge \bar{y}) \vee (\bar{x} \wedge y), \tag{5.10}$$

then $(B, \bot, \top, +, \wedge)$ is a Boolean ring $R(\mathfrak{A})$.

2. *If $\mathfrak{R}$ is a Boolean ring $(B, 0, 1, +, \cdot)$, and we define*

$$x \vee y = x + y + xy, \qquad \bar{x} = 1 + x, \tag{5.11}$$

then $(B, 0, 1, {}^-, \vee, \cdot)$ is a Boolean algebra $A(\mathfrak{R})$.

3. *$R(A(\mathfrak{R})) = \mathfrak{R}$ and $A(R(\mathfrak{A})) = \mathfrak{A}$.*

Proof. 1. If the Boolean algebra $(B, \bot, \top, {}^-, \vee, \wedge)$ is given, then the addition defined by (5.10) is obviously commutative. For associativity, we compute

$$\begin{aligned}(x + y) + z \\ &= (((x \wedge \bar{y}) \vee (\bar{x} \wedge y)) \wedge \bar{z}) \vee (((\bar{x} \vee y) \wedge (x \vee \bar{y})) \wedge z) \\ &\quad = (x \wedge \bar{y} \wedge \bar{z}) \vee (\bar{x} \wedge y \wedge \bar{z}) \vee (\bar{x} \wedge \bar{y} \wedge z) \vee (x \wedge y \wedge z),\end{aligned}$$

which is symmetric in x, y, and z. Also $x + x = \bot$. We already know $x \wedge x = x$. Then $(B, \bot, \top, +, \wedge)$ is a Boolean ring.

[2] Huntington treats our two axioms of the complement as a single axiom, but with the hypothesis that the identities are unique. This hypothesis can itself be proved by $\bot' = \bot' \vee \bot = \bot$ and so forth. In our formalism, the universe of a Boolean algebra is automatically closed under the operations $\vee$ and $\wedge$; but Huntington treats this closure as two separate axioms. Finally, Huntington requires Boolean algebras to have two distinct elements. Thus Huntington has ten axioms for Boolean algebras, and he shows them to be logically independent.

2. If the Boolean ring $(B, 0, 1, +, \cdot)$ is given, then the joining operation defined in (5.11) is obviously commutative. Also $x \vee 0 = x$. For distributivity, we have

$$\begin{aligned}(x \vee y)(x \vee z) &= (x + y + xy)(x + z + xz)\\ &= x^2 + xz + x^2z + xy + yz + xyz + x^2y + xyz + x^2yz\\ &= x + xz + xz + xy + yz + xyz + xy + xyz + xyz\\ &= x + yz + xyz\\ &= x \vee (yz),\end{aligned}$$

while

$$\begin{aligned}xy \vee xz &= xy + xz + x^2yz\\ &= xy + xz + xyz\\ &= x(y + z + yz)\\ &= x(y \vee z).\end{aligned}$$

Finally,

$$\begin{gathered}x \wedge \bar{x} = x(1 + x) = x + x^2 = x + x = 0,\\ x \vee \bar{x} = x + 1 + x + x(1 + x) = 1.\end{gathered}$$

Thus $(B, 0, 1, {}^-, \vee, \cdot)$ is a Boolean algebra.

3. In $A(\mathfrak{R})$, we compute

$$\begin{aligned}(x \cdot \bar{y}) \vee (\bar{x} \cdot y) &= (x \cdot (1 + y)) \vee ((1 + x) \cdot y)\\ &= (x + xy) \vee (y + xy)\\ &= x + xy + y + xy + (x + xy)(y + xy)\\ &= x + y;\end{aligned}$$

so this is the sum of x and y in $R(A(\mathfrak{R}))$ as well. In $R(\mathfrak{A})$,

$$x + y + (x \wedge y)$$

$$\begin{aligned}
&= ((x \wedge \bar{y}) \vee (\bar{x} \wedge y)) + (x \wedge y)\\
&= (((x \wedge \bar{y}) \vee (\bar{x} \wedge y)) \wedge \overline{x \wedge y}) \vee (\overline{(x \wedge \bar{y}) \vee (\bar{x} \wedge y)} \wedge x \wedge y)\\
&= (((x \wedge \bar{y}) \vee (\bar{x} \wedge y)) \wedge (\bar{x} \vee \bar{y})) \vee ((\bar{x} \vee y) \wedge (x \vee \bar{y}) \wedge x \wedge y)\\
&= (((x \wedge \bar{y}) \vee (\bar{x} \wedge y)) \wedge (\bar{x} \vee \bar{y})) \vee (x \wedge y)\\
&= ((x \wedge \bar{y} \wedge (\bar{x} \vee \bar{y})) \vee (\bar{x} \wedge y \wedge (\bar{x} \vee \bar{y}))) \vee (x \wedge y)\\
&= (x \wedge \bar{y}) \vee (\bar{x} \wedge y) \vee (x \wedge y)\\
&= x \vee y;
\end{aligned}$$

so this is the join of x and y in $A(R(\mathfrak{A}))$ as well. We finish by noting

$$\top + x = (\top \wedge \bar{x}) \vee (\bar{\top} \wedge x) = \bar{x} \vee (\bot \wedge x) = \bar{x} \vee \bot = \bar{x}. \qquad \square$$

We now have, by the Stone Representation Theorem (page 119), that every Boolean algebra embeds in the Boolean algebra $\mathscr{P}(\Omega)$ for some set Ω. A **Boolean operation** on $\mathscr{P}(\Omega)$ is just an operation on $\mathscr{P}(\Omega)$ that is the interpretation of a term in the signature of rings or Boolean algebras.

5.4.3. Filters

In $\mathscr{P}(\Omega)$ we have

$$X \cap Y = X \iff X \subseteq Y \iff X \cup Y = Y.$$

Then in an abstract Boolean algebra we can define an ordering $<$ by either of the equivalences

$$x \wedge y = x \iff x \leqslant y \iff x \vee y = y.$$

By Corollary 80.1 (page 88), A subset I of a Boolean algebra A is an ideal of the corresponding Boolean ring if and only if

$$\begin{gathered}
\bot \in I,\\
x \in I \;\&\; y \in I \implies x \vee y \in I,\\
y \in I \;\&\; x \leqslant y \implies x \in I.
\end{gathered}$$

By Theorem 87 (page 95), an ideal I of A is maximal if and only if

$$x \in A \smallsetminus I \iff \bar{x} \in I.$$

By the **De Morgan Laws** (5.8), the operation $x \mapsto \bar{x}$ is an isomorphism from a Boolean algebra to its dual. A subset of a Boolean algebra is called a **filter** if it is an ideal of the ring corresponding to the dual algebra, or equivalently if its image under $x \mapsto \bar{x}$ is an ideal of the ring corresponding to the original algebra. Thus a subset F of a Boolean algebra A is a filter if and only if

$$\begin{gathered}\top \in F,\\ x \in F \ \&\ y \in F \implies x \wedge y \in F,\\ x \in F \ \&\ x \leqslant y \implies y \in F.\end{gathered}$$

See Figure 5.4.

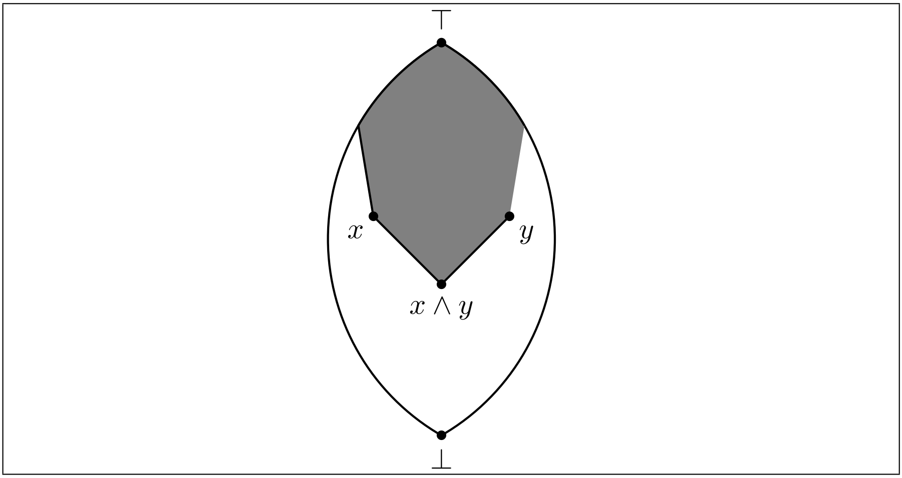

Figure 5.4.: A filter of a Boolean algebra

A maximal proper filter is called an **ultrafilter.**

Theorem 146. *A subset U of a Boolean algebra A is an ultrafilter if and only if*

$$x \in U \ \&\ y \in U \implies x \wedge y \in U, \qquad x \in A \smallsetminus U \iff \bar{x} \in U.$$

We may denote the set of all ultrafilters of A by

$$\mathrm{Sto}(A).$$

This is called the **Stone space** of A, because of the following, which is closer than Theorem 119 (page 119) is to the original form of Stone's theorem [57]. Given x in A, we shall use the notation

$$[x] = \{U \in \mathrm{Sto}(A) \colon x \in U\}$$

(but this is *not* an equivalence class as on page 76).

Theorem 147 (Stone Representation Theorem for Boolean Algebras)**.** *Let A be Boolean algebra.*

1. *The subset $\{[x] \colon x \in A\}$ of $\mathscr{P}(\mathrm{Sto}(A))$ is a basis for a compact Hausdorff topology on $\mathrm{Sto}(A)$.*
2. *The set $\{[x] \colon x \in A\}$ is precisely the set of clopen subsets of $\mathrm{Sto}(A)$ in this topology.*
3. *The map $x \mapsto [x]$ is an embedding of A in $\mathscr{P}(\mathrm{Sto}(A))$.*

5.5. Logical equivalence

Given a signature $\mathscr{S}$, in $\mathrm{Sen}(\mathscr{S})$ we define

$$\sigma \sim \tau \iff \mathbf{Mod}(\sigma) = \mathbf{Mod}(\tau).$$

The relation $\sim$ is called **logical equivalence.** Logically equivalent sentences are just sentences with the same models. We may use the notation

$$\sigma^{\sim} = \{\tau \in \mathrm{Sen}(\mathscr{S}) \colon \sigma \sim \tau\}$$

as on page 76. We also define

$$\mathrm{Lin}_0(\mathscr{S}) = \mathrm{Sen}(\mathscr{S})/{\sim};$$

this is the set of logical equivalence classes $\sigma^{\sim}$ of sentences σ of $\mathscr{S}$. (Here Lin stands for Lindenbaum; see below.) In model theory, while we are interested in the distinction between non-isomorphic structures

$$\begin{tikzcd}
\mathbf{Str}_{\mathscr{S}} \arrow[r, leftrightsquigarrow, "\vDash"] \arrow[d, "\mathfrak{A}\mapsto\mathrm{Th}(\mathfrak{A})"'] & \mathrm{Sen}(\mathscr{S}) \arrow[d, "\sigma\mapsto\sigma^{\sim}"] \\
\mathrm{S}_0(\mathscr{S}) \arrow[r, leftrightsquigarrow] \arrow[ur, leftrightsquigarrow, "\ni"'] & \mathrm{Lin}_0(\mathscr{S})
\end{tikzcd}$$

Figure 5.5.: Lindenbaum algebra

that are elementarily equivalent, we are not interested in the distinction between sentences that are different as strings of symbols, but are logically equivalent. However, the distinction is essential to logic as such. In any case, we can enlarge Figure 5.3 (page 143) to Figure 5.5. We have not got a symbol for the induced relation

$$\{(\mathrm{Th}(\mathfrak{A}), \sigma^{\sim})\colon (\mathfrak{A}, \sigma) \in \mathbf{Str}_{\mathscr{S}} \times \mathrm{Sen}(\mathscr{S}) \ \ \& \ \ \mathfrak{A} \vDash \sigma\},$$

which is

$$\{(T, \sigma^{\sim})\colon (T, \sigma) \in \mathrm{S}_0(\mathscr{S}) \times \mathrm{Sen}(\mathscr{S}) \ \ \& \ \ \sigma \in T\},$$

between $\mathrm{S}_0(\mathscr{S})$ and $\mathrm{Lin}_0(\mathscr{S})$.

A sentence like $\exists x \ x \neq x$ with no models is a **contradiction;** A sentence like $\forall x \ x = x$ of which every structure is a model is a **validity.** In the next theorem, we use the symbols

$$\bot, \qquad\qquad\qquad \top$$

to denote a contradiction and validity, respectively. The notion of a *congruence* on an algebra was defined on page 77.

Theorem 148. *For every signature $\mathscr{S}$, the relation $\mathscr{S}$ of logical equivalence is a congruence on the algebra*

$$(\mathrm{Sen}(\mathscr{S}), \bot, \top, \neg, \vee, \wedge).$$

The corresponding quotient algebra is a Boolean algebra.

Proof. Suppose $\sigma \sim \sigma_1$ and $\tau \sim \tau_1$. Then $\sigma \lor \tau \sim \sigma_1 \lor \tau_1$, because

$$\begin{aligned}\mathfrak{A} \vDash \sigma \lor \tau &\iff \mathfrak{A} \vDash \sigma \text{ OR } \mathfrak{A} \vDash \tau\\ &\iff \mathfrak{A} \vDash \sigma_1 \text{ OR } \mathfrak{A} \vDash \tau_1\\ &\iff \mathfrak{A} \vDash \sigma_1 \lor \tau_1.\end{aligned}$$

Similarly $\neg\sigma \sim \neg\sigma_1$ and $\sigma \land \tau \sim \sigma_1 \land \tau_1$. Thus $\sim$ is a congruence-relation. The quotient algebra is a Boolean algebra because $\sigma \lor \tau \sim \tau \lor \sigma$ and so forth. □

The Boolean algebra of the theorem is called the **Lindenbaum algebra** of sentences of $\mathscr{S}$, "in memory of a close colleague of Tarski who died at the hands of the Nazis" [31, p. 319]. In $\mathrm{Lin}_0(\mathscr{S})$, we now have $\sigma^{\sim} \leqslant \tau^{\sim}$ if and only if the sentence $\sigma \Rightarrow \tau$ is a validity, or equivalently

$$\mathfrak{A} \vDash \sigma \implies \mathfrak{A} \vDash \tau.$$

If T is a theory of $\mathscr{S}$, and $\mathscr{S} \in T$, and $\mathscr{S} \sim \tau$, then $\tau \in T$; thus

$$\{\tau \in \mathrm{Sen}(\mathscr{S})\colon \sigma \sim \tau\} = \{\tau \in T\colon \sigma \sim \tau\}.$$

In particular, the quotient $T/{\sim}$ is a subset of $\mathrm{Sen}(\mathscr{S})/{\sim}$.

If now τ is a topology on a set B, and $A \subseteq B$, then $\{A \cap F\colon F \in \tau\}$ is a topology on A, namely the **subspace topology,** and as being equipped with this topology, B is a subspace of A. In this case, B is **dense** in A if every nonempty open subset of A contains a point of B.

Theorem 149. *For every signature* $\mathscr{S}$*,*

1) *for every theory* T *of* $\mathscr{S}$*, the quotient* $T/{\sim}$ *is a filter of* $\mathrm{Lin}_0(\mathscr{S})$*;*
2) *for every complete theory* T *of* $\mathscr{S}$*, the quotient* $T/{\sim}$ *is an ultrafilter of* $\mathrm{Lin}_0(\mathscr{S})$*;*
3) *the map* $T \mapsto T/{\sim}$ *from* $\mathrm{S}_0(\mathscr{S})$ *to* $\mathrm{Sto}(\mathrm{Lin}_0(\mathscr{S}))$ *is a homeomorphism onto its image;*
4) *this image is dense in* $\mathrm{Sto}(\mathrm{Lin}_0(\mathscr{S}))$*.*

Proof. To establish density of the image of $\mathrm{S}_0(\mathscr{S})$ in $\mathrm{Sto}(\mathrm{Lin}_0(\mathscr{S}))$, we note that every nonempty open subset of $\mathrm{Sto}(\mathrm{Lin}_0(\mathscr{S}))$ includes $[\sigma^{\sim}]$ for some σ that is not a contradiction; but then σ has a model $\mathfrak{A}$, and so $[\sigma^{\sim}]$ contains $\mathrm{Th}(\mathfrak{A})$. □

$$\begin{tikzcd}
\mathbf{Str}_{\mathscr{S}} \arrow[r, leftrightsquigarrow, "\vDash"] \arrow[d, "\mathfrak{A}\mapsto\mathrm{Th}(\mathfrak{A})"'] & \mathrm{Sen}(\mathscr{S}) \arrow[d, "\sigma\mapsto\sigma^{\sim}"] \arrow[dl, leftrightsquigarrow, "\ni"] \\
\mathrm{S}_0(\mathscr{S}) \arrow[r, leftrightsquigarrow] \arrow[d, tail, "T\mapsto T/{\sim}"'] & \mathrm{Lin}_0(\mathscr{S}) \arrow[dl, leftrightsquigarrow, "\ni"] \\
\mathrm{Sto}(\mathrm{Lin}_0(\mathscr{S})) &
\end{tikzcd}$$

Figure 5.6.: Stone space of Lindenbaum algebra

We can enlarge Figure 5.5 to Figure 5.6.

Corollary 149.1. *For every signature $\mathscr{S}$, the following statements are equivalent:*

- $\mathrm{S}_0(\mathscr{S})$ *is compact.*
- *The image of* $\mathrm{S}_0(\mathscr{S})$ *under* $T \mapsto T/{\sim}$ *is* $\mathrm{Sto}(\mathrm{Lin}_0(\mathscr{S}))$.
- *This image is a closed subspace of* $\mathrm{Sto}(\mathrm{Lin}_0(\mathscr{S}))$.

Proof. All closed subspaces of a compact space are compact. The only dense closed subspace of a topological space is the whole space. In a Hausdorff space, all compact subspaces are closed. □

We shall therefore be able to understand the Compactness Theorem as any one of these three equivalent statements. However, we cannot prove the Compactness Theorem itself without more work. So far, all we have used for our theorems is that $\mathbf{Str}_{\mathscr{S}}$ is a class $\boldsymbol{M}$, and $\mathrm{Sen}(\mathscr{S})$ is the universe of an algebra $(S, \bot, \top, \neg, \vee, \wedge)$, and $\vDash$ is a relation from

$\boldsymbol{M}$ to S, where for all A in $\boldsymbol{M}$, and all s and t in S,

$$\begin{gathered} A \nvDash \bot, \qquad\qquad A \vDash \top, \\ A \vDash \neg s \iff A \nvDash s, \\ A \vDash s \lor t \iff A \vDash s \text{ OR } A \vDash t, \\ A \vDash s \land t \iff A \vDash s \mathrel{\&} A \vDash t. \end{gathered}$$

Hence for example we can replace $\mathbf{Str}_{\mathscr{S}}$ with an arbitrary subclass. In particular, for each non-contradictory σ in $\mathrm{Sen}(\mathscr{S})$, we can choose (using the Axiom of Choice) a model $\mathfrak{A}_\sigma$ of σ, and then we can replace $\mathbf{Str}_{\mathscr{S}}$ with $\{\mathfrak{A}_\sigma \colon \sigma \in \mathrm{Sen}(\mathscr{S})\}$. The relation $\sim$ of logical equivalence will be unchanged; but possibly not every element of $\mathrm{Sto}(\mathrm{Lin}_0(\mathscr{S}))$ is $\mathrm{Th}(\mathfrak{A}_\sigma)$ for some σ.

5.6. Definable relations

We are usually interested in $\mathbf{Mod}(T)$ for particular theories T of a signature $\mathscr{S}$. One way to study this is to study the *definable relations* of models of T. Suppose $\mathfrak{A} \vDash T$, and φ is an n-ary formula of $\mathscr{S}$. Then the subset

$$\{\boldsymbol{a} \in A^n \colon \mathfrak{A} \vDash \varphi(\boldsymbol{a})\}$$

of A^n is said to be **defined** by φ and can be denoted by one of

$$\varphi^{\mathfrak{A}}, \qquad\qquad \varphi(\mathfrak{A}).$$

This set is then a **0-definable relation** of $\mathfrak{A}$. If $B \subseteq A$, and φ is a formula of $\mathscr{S}(B)$, then $\varphi^{\mathfrak{A}}$ is a B**-definable relation** of $\mathfrak{A}$.

If σ is a sentence, then $\sigma^{\mathfrak{A}} \in \{0, 1\}$, and

$$\sigma^{\mathfrak{A}} = 1 \iff \mathfrak{A} \vDash \sigma.$$

We can then extend the notion of logical equivalence to arbitrary formulas φ and ψ of $\mathscr{S}$ having the same free variables: these two formulas are **logically equivalent,** and we write

$$\varphi \sim \psi,$$

if for all $\mathfrak{A}$ in $\mathbf{Str}_{\mathscr{S}}$,

$$\varphi^{\mathfrak{A}} = \psi^{\mathfrak{A}}. \tag{5.12}$$

If V is a finite set of variables, we may denote by

$$\mathrm{Fm}_V(\mathscr{S})$$

the set of formulas φ of $\mathscr{S}$ such that $\mathrm{fv}(\varphi) = V$; then we let

$$\mathrm{Lin}_V(\mathscr{S}) = \mathrm{Fm}_V(\mathscr{S})/{\sim}.$$

Then $\mathrm{Lin}_V(\mathscr{S})$ is the universe of a Boolean algebra, just as in Theorem 148 (page 152). Alternatively, if a bijection $i \mapsto v_i$ from n in ω to V is understood to have been chosen, we may replace the subscript V with n.

Further modifications are possible. If T is some theory of $\mathscr{S}$, we say that φ and ψ are **equivalent in** T (or *modulo* T, or *with respect to* T) if (5.12) holds for all $\mathfrak{A}$ in $\mathbf{Mod}(T)$. Then we obtain the algebras $\mathrm{Lin}_n(T)$.

5.7. Substructures

A formula is **quantifier-free** if neither of the symbols $\exists$ and $\forall$ occurs in it. There is a recursive definition of the quantifier-free formulas: just delete condition 4 from the recursive definition of formulas on page 136. If $\mathfrak{A}$ is a structure of signature $\mathscr{S}$, then the **diagram** of $\mathfrak{A}$ is the set

$$\mathrm{diag}(\mathfrak{A})$$

of quantifier-free sentences of $\mathscr{S}(A)$ that are true in $\mathfrak{A}$. Now we can give a variation of Theorem 136 (page 136):

Theorem 150. *Suppose $\mathfrak{A}$ and $\mathfrak{B}$ are in $\mathbf{Str}_{\mathscr{S}}$, and $h\colon A \to B$. The following are equivalent:*

1. *h is an embedding of $\mathfrak{A}$ in $\mathfrak{B}$.*
2. *For all quantifier-free formulas φ of $\mathscr{S}$, for all $\boldsymbol{a}$ in $A^{\mathrm{var}(\varphi)}$,*

$$\mathfrak{A} \models \varphi(\boldsymbol{a}) \implies \mathfrak{B} \models \varphi(h(\boldsymbol{a})).$$

3. *For all quantifier-free formulas φ of $\mathscr{S}$, for all $\boldsymbol{a}$ in $A^{\operatorname{var}(\varphi)}$,*

$$\mathfrak{A} \vDash \varphi(\boldsymbol{a}) \iff \mathfrak{B} \vDash \varphi(h(\boldsymbol{a})). \tag{5.13}$$

4. *When $\mathfrak{B}^*$ is the expansion of $\mathfrak{B}$ to $\mathscr{S}(A)$ such that, for each a in A,*

$$a^{\mathfrak{B}^*} = h(a),$$

then

$$\mathfrak{B}^* \vDash \operatorname{diag}(\mathfrak{A}).$$

For the theorem, it would be enough to define $\operatorname{diag}(\mathfrak{A})$ to consist of the atomic and negated atomic sentences of $\mathscr{S}(A)$ that are true in $\mathfrak{A}$; and indeed sometimes this is the definition used. We shall use the following in proving the Compactness Theorem by Henkin's method (page 168).

Corollary 150.1. *Suppose $\mathfrak{A}$ and $\mathfrak{B}$ are in $\mathbf{Str}_{\mathscr{S}}$. If $\mathfrak{B}$ expands to a model $\mathfrak{B}^*$ of $\operatorname{diag}(\mathfrak{A})$, and every element of B is $a^{\mathfrak{B}^*}$ for some a in A, then*

$$\mathfrak{A} \cong \mathfrak{B},$$

and indeed $a \mapsto a^{\mathfrak{B}^*}$ is an isomorphism from $\mathfrak{A}$ to $\mathfrak{B}$.

A theory T of $\mathscr{S}$ is **axiomatized** by a subset Γ of $\operatorname{Sen}(\mathscr{S})$ if T is the closure of Γ, that is,

$$T = \operatorname{Th}(\mathbf{Mod}(\Gamma));$$

equivalently, every model of Γ is a model of T.

If $\mathfrak{A}$ is a structure of signature $\mathscr{S}$, then, by the last theorem, the class of structures of $\mathscr{S}(A)$ in which $\mathfrak{A}_A$ embeds is elementary, and its theory is axiomatized by $\operatorname{diag}(\mathfrak{A})$. However, the class of structures of $\mathscr{S}$ in which $\mathfrak{A}$ embeds is not generally elementary.

A **universal** formula is a formula of the form

$$\forall x_0 \cdots \forall x_{n-1}\, \varphi, \tag{5.14}$$

where φ is quantifier-free. The universal formula in (5.14) might be abbreviated as

$$\forall \boldsymbol{x}\ \varphi.$$

If T is a theory, then we denote by

$$T_{\forall}$$

the theory axiomatized by the universal sentences in T.

Lemma 13. *For every theory T, the theory $T_{\forall}$ is included in the theory of substructures of models of T, that is,*

$$\mathfrak{A} \subseteq \mathfrak{B}\ \&\ \mathfrak{B} \vDash T \implies \mathfrak{A} \vDash T_{\forall}.$$

Proof. Suppose $\mathfrak{A} \subseteq \mathfrak{B}$, and $\mathfrak{B} \vDash T$, and φ is quantifier-free, and $\forall \boldsymbol{x}\ \varphi$ is in T. For every $\boldsymbol{a}$ in $A^{\mathrm{fv}(\varphi)}$, we have $\boldsymbol{a} \in B^{\mathrm{fv}(\varphi)}$, so $\mathfrak{B} \vDash \varphi(\boldsymbol{a})$ and therefore, by the last theorem, $\mathfrak{A} \vDash \varphi(\boldsymbol{a})$. Thus $\mathfrak{A} \vDash \forall \boldsymbol{x}\ \varphi$. □

The converse is given in Theorem 176 on page 192 below.

In Theorem 150, if (5.13) holds for *all* formulas φ of $\mathscr{S}$ and all $\boldsymbol{a}$ in $A^{\mathrm{var}(\varphi)}$, then h is called an **elementary embedding** of $\mathfrak{A}$ in $\mathfrak{B}$. In this case, if $\mathfrak{A} \subseteq \mathfrak{B}$, and h is the inclusion of A in B, then $\mathfrak{A}$ is called an **elementary substructure** of $\mathfrak{B}$, and we write

$$\mathfrak{A} \preccurlyeq \mathfrak{B}.$$

The structures in which $\mathfrak{A}$ embeds elementarily are precisely the reducts to $\mathscr{S}$ of the models of $\mathrm{Th}(\mathfrak{A}_A)$.

A theory T of a signature $\mathscr{S}$ is called **model-complete** if for all models $\mathfrak{A}$ of T, the theory of $\mathscr{S}(A)$ axiomatized by $T \cup \mathrm{diag}(\mathfrak{A})$ is complete.

Theorem 151. *A theory T is model-complete if and only if, for all $\mathfrak{A}$ and $\mathfrak{B}$ in* $\mathbf{Mod}(T)$,

$$\mathfrak{A} \subseteq \mathfrak{B} \implies \mathfrak{A} \preccurlyeq \mathfrak{B}.$$

Proof. Each condition is equivalent to the condition that, for all models $\mathfrak{A}$ of T, $T \cup \mathrm{diag}(\mathfrak{A})$ axiomatizes $\mathrm{Th}(\mathfrak{A}_A)$. □

The Löwenheim–Skolem Theorem below is a generalization of the theorem published by Löwenheim in 1915 [45] and improved by Skolem in 1920 [56]: a sentence with a model has a countable model. Skolem's argument uses what we shall call the *Skolem normal form* of the given sentence; we shall discuss this in §8.6 (page 212). Meanwhile, an example of a sentence in Skolem normal form is

$$\forall x\ \exists y\ x\ R\ y,$$

where R is a binary predicate. If this sentence has a model $\mathfrak{A}$, then, by the Axiom of Choice, there is a singulary operation $x \mapsto x^*$ on A such that, for all b in A,

$$\mathfrak{A} \vDash b\ R\ b^*.$$

Given b in A, we can define $(b_k \colon k \in \omega)$ recursively by

$$b_0 = b, \qquad\qquad b_{k+1} = {b_k}^*.$$

Then $\{b_k \colon k \in \omega\}$ is countable and is the universe of a substructure of $\mathfrak{A}$ in which $\forall x\ \exists y\ x\ R\ y$ is true. Our own proof of the general result will follow the lines of Skolem's idea. But we shall use the following theorem in order to be able to work with arbitrary sentences. We shall use the *idea* of the theorem in proving the Compactness Theorem by Henkin's method (page 168).

Theorem 152 (Tarski–Vaught Test). *Suppose $\mathfrak{A} \subseteq \mathfrak{B}$, both having signature $\mathscr{S}$. Then $\mathfrak{A} \preccurlyeq \mathfrak{B}$, provided that, for all singulary formulas φ of $\mathscr{S}(A)$,*

$$\mathfrak{B} \vDash \exists x\ \varphi \implies \text{for some } c \text{ in } A,\ \mathfrak{B} \vDash \varphi(c),$$

that is,

$$\varphi^{\mathfrak{B}} \neq \varnothing \implies \varphi^{\mathfrak{B}} \cap A \neq \varnothing.$$

Proof. Under the given condition, we show by induction that for all formulas φ of $\mathscr{S}$, if $\boldsymbol{a} \in A^{\mathrm{fv}(\varphi)}$, then

$$\mathfrak{A} \vDash \varphi(\boldsymbol{a}) \iff \mathfrak{B} \vDash \varphi(\boldsymbol{a}).$$

This is given to be the case when φ is atomic (or more generally quantifier-free), and it is easily preserved under negation and conjunction. Suppose it holds when φ is a formula ψ. By hypothesis, for all $\boldsymbol{a}$ in $A^{\mathrm{fv}(\exists x\ \psi)}$, the following are equivalent:

$$\begin{gathered}
\mathfrak{B} \vDash (\exists y\ \psi)(\boldsymbol{a}), \\
\text{for some } b \text{ in } A,\ \mathfrak{B} \vDash \psi^y_b(\boldsymbol{a}), \\
\text{for some } b \text{ in } A,\ \mathfrak{A} \vDash \psi^y_b(\boldsymbol{a}), \\
\mathfrak{A} \vDash (\exists y\ \psi)(\boldsymbol{a}).
\end{gathered}$$

This completes the induction. □

6. Compactness and Łoś's Theorem

In a signature $\mathscr{S}$, if Γ is a set of sentences whose every finite subset has a model, we shall show that Γ itself has a model. This will be the **Compactness Theorem.**

The Compactness Theorem for countable signatures was obtained by Gödel in his doctoral dissertation and published in 1930 as a kind of corollary [24, Thm X, p. 590] of his *Completeness Theorem,* which we shall take up in Chapter 8 (page 197). According to Chang and Keisler [8, p. 604], Malcev established the Compactness Theorem for arbitrary signatures in 1936; but in Hodges's estimation [31, p. 318], the statement and proof had "shortcomings."

Henkin gave a new proof of the Compactness Theorem in his own doctoral dissertation and published it in 1949 [28, 30]. An alternative proof by means of *ultraproducts* was published in 1962/3 by Frayne, Morel, and Tarski [19, Thm 2.10, p. 216].[1] It is these two proofs that will interest us here.

6.1. Construction of elementary substructures

First we establish a result that does not rely on the Compactness Theorem, but does use the Axiom of Choice.

Theorem 153 (Downward Löwenheim–Skolem). *By the Axiom of Choice, for every signature $\mathscr{S}$, for every structure $\mathfrak{B}$ of $\mathscr{S}$, for every subset X of B, there is a structure $\mathfrak{A}$ such that* **AC**

$$\mathfrak{A} \preccurlyeq \mathfrak{B}, \qquad X \subseteq A, \qquad |A| \leqslant \max(|X|, |\mathscr{S}|, \omega).$$

Proof. Suppose $Y \subseteq B$. By the Axiom of Choice, there is a bijection **AC**

[1] Apparently this proof was announced in 1958. For this and other historical notes on the ultraproduct method, see the introduction to [19] and its correction [20].

$\varphi \mapsto b_\varphi$ from $\mathrm{Fm}_{\{x\}}(\mathscr{S}(Y))$ to B such that, for all φ in $\mathrm{Fm}_{\{x\}}(\mathscr{S}(Y))$,

$$\mathfrak{B} \vDash \exists x\ \varphi \Rightarrow \varphi(b_\varphi).$$

If $a \in B$, and φ is the formula $x = a$, then $b_\varphi = a$. Thus, when we define

$$Y' = \{b_\varphi \colon \varphi \in \mathrm{Fm}_{\{x\}}(\mathscr{S}(Y))\},$$

we have $Y \subseteq Y'$. Then

$$|Y'| \leqslant \max(|Y|, |\mathscr{S}|, \omega).$$

Now we can define $(X_n \colon n \in \omega)$ recursively by

$$X_0 = X, \qquad\qquad X_{k+1} = X_k{}',$$

and we can let

$$A = \bigcup_{n\in\omega} X_n.$$

By considering formulas $F\boldsymbol{x} = y$, we see that A is the universe of a substructure $\mathfrak{A}$ of $\mathfrak{B}$. It is of the required cardinality, and by the Tarski–Vaught Test, it is an elementary substructure of $\mathfrak{B}$. □

In the theorem, if $\max(|S|, \omega) \leqslant |X|$, then $|A| = |X|$. The *proof* of the theorem does not use cardinalities as such, but makes essential use of the Axiom of Choice, and by this, all sets have cardinalities anyway. The "upward" version of the theorem occurs on page 191.

6.2. Models from theories

Recall from page 139 that a *theory* of a signature $\mathscr{S}$ is just the set T of sentences of $\mathscr{S}$ that are true in each of some given class of structures of $\mathscr{S}$. In this case, by Theorem 149 (page 153), the set $\{\sigma^\sim \colon \sigma \in T\}$ of logical equivalence classes of elements of T is a filter of the Lindenbaum algebra $\mathrm{Lin}_0(\mathscr{S})$.

If $\Gamma \subseteq \mathrm{Sen}(\mathscr{S})$, and every finite subset of Γ has a model, then the set $\{\sigma^\sim \colon \sigma \in \Gamma\}$ generates a *proper* filter of $\mathrm{Lin}_0(\mathscr{S})$. In this setting, the

Compactness Theorem is that every such filter is $\{\sigma^{\sim} \colon \sigma \in \mathrm{Th}(\mathcal{K})\}$ for some class $\mathcal{K}$ of structures of $\mathscr{S}$.

Thus the Compactness Theorem is the converse of the theorem that $T/{\sim}$ is a filter when T is a theory (Theorem 149, page 153). However, we shall not use this formulation in our first proofs. Given a set of sentences whose every finite subset has a model, we just want to show that the whole set has a model.

Suppose $\Gamma \subseteq \mathrm{Sen}(\mathscr{S})$, and Γ does have a model. By the Downward Löwenheim–Skolem Theorem, Γ has a model of size no greater than $\max(|\mathscr{S}|, \omega)$. Let A be a set of new constants of size $\max(|\mathscr{S}|, \omega)$. Then we can find a structure $\mathfrak{B}$ of $\mathscr{S}(A)$ that is a model of Γ and whose every element is the interpretation of a closed term of $\mathscr{S}(A)$. (For example, every element of B could be $a^{\mathfrak{B}}$ for some a in A.) Let $T = \mathrm{Th}(\mathfrak{B})$. If C is the set of closed terms of $\mathscr{S}(A)$, and E is the equivalence relation on C given by

$$t \mathrel{E} u \iff (t = u) \in T, \tag{6.1}$$

then there is a well-defined bijection $t/E \mapsto t^{\mathfrak{B}}$ from C/E onto B. Then $\mathfrak{B}$ is determined up to isomorphism by T; we may say $\mathfrak{B}$ is a **canonical model** of T.

Thus, if we are going to be able to prove the Compactness Theorem at all, then, given a subset Γ of $\mathrm{Sen}(\mathscr{S})$ whose every finite subset has a model, we must be able to embed Γ in a theory with a canonical model; and the signature of that theory can be $\mathscr{S}(A)$, where $|A| = \max(|\mathscr{S}|, \omega)$.

An arbitrary complete theory need not have a canonical model.

Theorem 154. *A complete theory T of a signature $\mathscr{S}$ has a canonical model if and only if, for every singulary formula $\varphi(x)$ of $\mathscr{S}$, for some closed term t of $\mathscr{S}$,*

$$(\exists x\ \varphi) \in T \implies \varphi(t) \in T. \tag{6.2}$$

Proof. Suppose T has a canonical model $\mathfrak{B}$. If $(\exists x\ \varphi) \in T$, then $\mathfrak{B} \vDash \exists x\ \varphi$, so for some b in B, $\mathfrak{B} \vDash \varphi(b)$. But then $b = t^{\mathfrak{B}}$ for some closed term t of $\mathscr{S}$, so $\mathfrak{B} \vDash \varphi(t)$ and therefore $\varphi(t) \in T$.

Suppose conversely (6.2) holds for all singulary $\varphi(x)$ of $\mathscr{S}$. Since T is a complete theory, it is $\mathrm{Th}(\mathfrak{M})$ for some structure $\mathfrak{M}$ of $\mathscr{S}$. Let C be the set of closed terms of $\mathscr{S}$ and let $B = \{t^{\mathfrak{M}} \colon t \in C\}$. Then B is the universe of a substructure $\mathfrak{B}$ of $\mathfrak{M}$, and moreover $\mathfrak{B} \preccurlyeq \mathfrak{M}$ by the Tarski–Vaught Test (page 159). Then $\mathfrak{B}$ is a canonical model of T. □

The following theorem can be seen as a combination of Hodges's [31, Thms 2.3.3 & 2.3.4, pp. 44–6]. Hodges refers to the set T of sentences in the theorem as a *Hintikka set,* because of a 1955 paper of Hintikka. According to Hodges, "equivalent ideas appear in" a 1955 paper by Beth and a 1956 paper by Schütte.

Theorem 155. *Let $\mathscr{S}$ be a signature, and suppose T is a subset of* $\mathrm{Sen}(\mathscr{S})$ *such that*

1) every finite subset of T has a model;

2) for all σ in $\mathrm{Sen}(\mathscr{S})$*, either σ or $\neg\sigma$ is in T;*

3) for all singulary formulas $\varphi(x)$ of $\mathscr{S}$, for some closed term t of $\mathscr{S}$, (6.2) *holds.*

Then T is a complete theory with a canonical model.

Proof. Let T be as in the hypothesis. It suffices to show that T has a model $\mathfrak{B}$, since in this case T must be the complete theory $\mathrm{Th}(\mathfrak{B})$, and T will have a canonical model by the previous theorem. In fact the model $\mathfrak{B}$ that we find will be a canonical model.

If, for some n in ω, the sentence

$$\sigma_0 \land \dots \land \sigma_{n-1} \Rightarrow \sigma_n$$

of $\mathscr{S}$ is a validity, then the finite subset $\{\sigma_0, \dots, \sigma_{n-1}, \neg\sigma_n\}$ of $\mathrm{Sen}(\mathscr{S})$ has no model, and therefore

$$\{\sigma_0, \dots, \sigma_{n-1}\} \subseteq T \implies \sigma_n \in T.$$

In case $n = 0$, this means T contains all validities. For instance, for all closed terms t of $\mathscr{S}(A)$,

$$(t = t) \in T. \tag{6.3}$$

Also, for all formulas φ of $\mathscr{S}$, for all closed terms s_x and t_x of $\mathscr{S}$ (where x ranges over $\operatorname{fv}(\varphi)$),

$$\begin{aligned}\bigl\{s_x = t_x \colon x \in \operatorname{fv}(\varphi)\bigr\} \cup \bigl\{\varphi\bigl(s_x \colon x \in \operatorname{fv}(\varphi)\bigr)\bigr\} \subseteq T&\\ \implies \varphi\bigl(t_x \colon x \in \operatorname{fv}(\varphi)\bigr) \in T.&\end{aligned} \tag{6.4}$$

In particular, for all closed terms s, t, and u of $\mathscr{S}$,

$$(s = t) \in T \implies (t = s) \in T, \tag{6.5}$$

$$\{s = t,\ t = u\} \subseteq T \implies (s = u) \in T. \tag{6.6}$$

We now construct the desired model $\mathfrak{B}$ of T. The argument will have these parts:

1. The definition of B.
2. The definition of $F^{\mathfrak{B}}$ for operation symbols F of $\mathscr{S}$.
3. The definition of $R^{\mathfrak{B}}$ for predicates R of $\mathscr{S}$.
4. The proof that $\mathfrak{B} \vDash \sigma$ for all atomic sentences σ in T.
5. The proof that $\mathfrak{B} \vDash T$.

1. By (6.3), (6.5), and (6.6), the relation E given by

$$t \mathrel{E} u \iff (t = u) \in T$$

is an equivalence relation on the set C of closed terms of $\mathscr{S}$. Now we may define $B = C/E$. In general, if $\boldsymbol{t}$ is an element $(t_0, \dots, t_{n-1})$ of C^n, we may use the notation

$$\boldsymbol{t}/E = (t_0/E, \dots, t_{n-1}/E).$$

2. Given an n-ary operation symbol F of $\mathscr{S}$ for some n in ω, given $\boldsymbol{t}$ in C^n, we want to define

$$F^{\mathfrak{B}}(\boldsymbol{t}/E) = (Ft_0 \cdots t_{n-1})/E.$$

This is a valid definition, since if $\boldsymbol{s}/E = \boldsymbol{t}/E$, then T contains the equations

$$t_0 = s_0, \qquad \dots, \qquad t_{n-1} = s_{n-1}, \qquad Ft_0 \cdots t_{n-1} = Ft_0 \cdots t_{n-1},$$

so that, by n applications of (6.4), T contains the equation

$$Ft_0 \cdots t_{n-1} = Fs_0 \cdots s_{n-1}.$$

3. Next, given an n-ary predicate R in $\mathscr{S}$ for some n in ω, we want to define $R^{\mathfrak{B}}$ by the rule

$$\boldsymbol{t}/E \in R^{\mathfrak{B}} \iff (Rt_0 \cdots t_{n-1}) \in T; \tag{6.7}$$

but again we must check that this definition is good. We are free to make the definition

$$R^{\mathfrak{B}} = \{\boldsymbol{t}/E \colon (Rt_0 \cdots t_{n-1}) \in T\};$$

but to have (6.7), we must have

$$\boldsymbol{t}/E = \boldsymbol{s}/E \;\;\&\;\; Rt_0 \cdots t_{n-1} \in T \implies (Rs_0 \cdots s_{n-1}) \in T.$$

We do have this by (6.4).

4. For all atomic sentences σ of $\mathscr{S}$, we show

$$\mathfrak{B} \vDash \sigma \iff \sigma \in T. \tag{6.8}$$

If σ is an equation $s = t$, then

$$\mathfrak{B} \vDash \sigma \iff s^{\mathfrak{B}} = t^{\mathfrak{B}} \iff s \mathrel{E} t \iff \sigma \in T,$$

while if σ is $Rt_0 \cdots t_{n-1}$, then

$$\mathfrak{B} \vDash \sigma \iff \boldsymbol{t}^{\mathfrak{B}} \in R^{\mathfrak{B}} \iff \boldsymbol{t}/E \in R^{\mathfrak{B}} \iff \sigma \in T.$$

5. Since for all sentences σ and τ of $\mathscr{S}$,

$$\begin{gathered} \sigma \in T \iff \neg\sigma \notin T, \\ \{\sigma, \tau\} \subseteq T \implies \sigma \land \tau \in T, \end{gathered}$$

and for all singulary formulas φ of $\mathscr{S}$,

$$\exists x\; \varphi \in T \iff \text{ for some } t \text{ in } C,\ \varphi(t) \in T,$$

we can conclude that (6.8) holds for arbitrary sentences σ of $\mathscr{S}$. In particular, $\mathfrak{B} \vDash T$. □

Given a set Γ of sentences of $\mathscr{S}$ whose every finite subset has a model, we shall embed Γ in a set T as in the last theorem in two different ways, by the Henkin method and the ultraproduct method. These methods differ specifically as follows.

The Henkin method. If A is a set of new constants, then by
AC Zorn's Lemma, there will be a maximal subset T of $\operatorname{Sen}(\mathscr{S}(A))$ such that

i) $\Gamma \subseteq T$,
ii) every finite subset of T has a model, and
iii) For every singulary $\varphi(x)$ of $\mathscr{S}(A)$, for some closed term t of $\mathscr{S}(A)$, (6.2) holds.

If, further, $|A| = \max(|\mathscr{S}|, \omega)$, then T will satisfy the remaining hypothesis of the last theorem. In case $\mathscr{S}$ is countable, then T can be found, without using the Axiom of Choice, by listing the sentences of $\mathscr{S}(A)$ and deciding, one by one, whether a sentence or its negation should belong to T. Alternatively, $T/{\sim}$ can be found through the compactness of the Stone space of the Lindenbaum algebra of (logical equivalence classes of) sentences of an appropriate signature.

The ultraproduct method. For every finite subset Δ of Γ, using
the Axiom of Choice if necessary, we pick a model $\mathfrak{A}_\Delta$ of Δ. The **AC**
universe of each $\mathfrak{A}_\Delta$ being A_Δ, we let

$$A = \prod_{\Delta \in \mathscr{P}_\omega(\Gamma)} A_\Delta.$$

Considering A as a set of new constants, we expand each $\mathfrak{A}_\Delta$ to a structure $\mathfrak{A}_\Delta{}^*$ of $\mathscr{S}(A)$ by defining, for each a in A,

$$a^{\mathfrak{A}_\Delta{}^*} = a_\Delta.$$

i) Letting $\mathscr{U}$ be an ultrafilter of $\mathscr{P}(\mathscr{P}_\omega(\Gamma))$, we define

$$T = \bigl\{\sigma \in \operatorname{Sen}(\mathscr{S}(A))\colon \{\Delta \in \mathscr{P}_\omega(\Gamma)\colon \mathfrak{A}_\Delta{}^* \vDash \sigma\} \in \mathscr{U}\bigr\}.$$

Then T will be a complete theory with a canonical model; such a model is called an **ultraproduct** of the structures $\mathfrak{A}_\Delta{}^*$.

ii) If, further, $\mathscr{U}$ is chosen so as to contain every subset

$$\{\Delta \in \mathscr{P}_\omega(\Gamma)\colon \sigma \in \Delta\}$$

of $\mathscr{P}_\omega(\Gamma)$, where $\sigma \in \Gamma$, then we shall have $\Gamma \subseteq T$.

We now work out the details.

6.3. Henkin's method

The following does not require the Axiom of Choice.

Theorem 156 (Countable Compactness). *Suppose $\mathscr{S}$ is a countable signature, and Γ is a set of sentences of $\mathscr{S}$ whose every finite subset has a model. Then Γ has a model.*

Proof. Let A be a set $\{a_n\colon n \in \omega\}$ of constants not belonging to $\mathscr{S}$. It is possible to define a surjective function $n \mapsto \sigma_n$ from ω to $\mathrm{Sen}(\mathscr{S}(A))$. We shall recursively define a function $n \mapsto \Gamma_n$ from ω to $\mathscr{P}(\mathrm{Sen}(\mathscr{S}(A)))$ such that the union $\bigcup_{n\in\omega}\Gamma_n$ is a theory T as in Theorem 155.

We start by letting

$$\Gamma_0 = \Gamma.$$

Then every finite subset of Γ_0 has a model. Suppose Γ_n has been defined so that

1) it is the union of Γ_0 with a finite set, and
2) its every finite subset has a model.

Note that this is indeed the case when $n = 0$. If it is the case for some n, then one of $\Gamma_n \cup \{\sigma_n\}$ and $\Gamma_n \cup \{\neg\sigma_n\}$ has the same properties. Indeed, only the second property could fail. If $\Gamma_n \cup \{\sigma_n\}$ does not have the property, then for some finite subset Δ of Γ_n, there is no model of $\Delta \cup \{\sigma_n\}$. Thus in every model of Δ, the sentence $\neg\sigma_n$ is true. If Θ is another finite subset of Γ_n, then $\Delta \cup \Theta$ has a model, and this will also be a model of $\Theta \cup \{\neg\sigma\}$. Thus $\Gamma_n \cup \{\neg\sigma\}$ is as desired. In any case, we define Γ_{n+1} as follows.

- If the set $\Gamma_n \cup \{\neg\sigma_n\}$ has the two desired properties, we let Γ_{n+1} be this set.

- Suppose $\Gamma_n \cup \{\neg\sigma_n\}$ does not have the properties, so that $\Gamma_n \cup \{\sigma\}$ must have them.
  - If σ_n is not existential, we let $\Gamma_{n+1} = \Gamma_n \cup \{\sigma_n\}$.
  - If σ_n is $\exists x\ \varphi$ for some formula φ, we let

    $$\Gamma_{n+1} = \Gamma_n \cup \{\exists x\ \varphi\} \cup \{\varphi^x_{a_k}\}, \tag{6.9}$$

    where k is the least ℓ such that a_ℓ does not occur in any sentence in $\Gamma_n \cup \{\exists x\ \varphi\}$. Since this set is assumed to be finite, such ℓ exist.

Then Γ_{n+1} has the desired properties that Γ_n is assumed to have.

By induction, all Γ_n do have the properties. We can now let

$$T = \bigcup_{n\in\omega} \Gamma_n.$$

If $\{\tau_0, \dots, \tau_{n-1}\}$ is a finite subset of T, then each τ_k belongs to some $\Gamma_{f(k)}$, and so they all belong to $\Gamma_{\max\{f(k)\colon k<n\}}$, and therefore they have a common model. In short, every finite subset of T has a model. Also, for all sentences σ of $\mathscr{S}(A)$, either σ or $\neg\sigma$ is in T. Finally, by construction, if $\exists x\ \varphi$ is in T, then φ^x_c is in T for some constant c. Thus Theorem 155 applies, and so T has a model. In particular, this model is a model of Γ. □

In a variant of the foregoing proof, Theorem 155 is not used as it is. We first assume that there is no finite bound on the sizes of models of finite subsets of Γ. Then we let $\Gamma_0 = \Gamma \cup \{a_i \neq a_j \colon i < j < \omega\}$. We obtain Γ_{n+1} from Γ_n as before, except that, if we make the definition (6.9), then we have let k be the least ℓ such that a_ℓ does not occur in any sentence in $(\Gamma_n \cup \{\exists x\ \varphi\}) \smallsetminus \Gamma_0$. We define T as before, but now we can obtain a model of T whose universe is just A.

In this alternative approach, The remaining case is handled differently. Suppose Γ has a finite subset Δ_0 such that there is a finite upper bound on the size of models of Δ_0. Since Γ is countable, we can form a chain

$$\Delta_0 \subseteq \Delta_1 \subseteq \Delta_2 \subseteq \cdots$$

of finite subsets of Γ whose union is Γ. Then there is a corresponding chain

$$\mathscr{S}_0 \subseteq \mathscr{S}_1 \subseteq \mathscr{S}_2 \subseteq \cdots$$

of finite signatures such that $\Delta_n \subseteq \mathrm{Sen}(\mathscr{S}_n)$ for each n in ω. For each n in ω then, there are only finitely many nonisomorphic structures of $\mathscr{S}_n$ that are models of Δ_n. We may assume that the universe of each of them is a von Neumann natural number. As n varies, these models of Δ_n are (partially) ordered by the relation of being a reduct. That is, if $m < n$ and $\mathfrak{A} \vDash \Delta_m$, while $\mathfrak{B} \vDash \Delta_n$,

$$\mathfrak{A} < \mathfrak{B} \iff \mathfrak{A} = \mathfrak{B} \restriction \mathscr{S}_m.$$

With this ordering, these structures compose a *tree* in the sense of page 211. The tree is an infinite ω-tree, so by König's Lemma (Theorem 189, page 212) it has an infinite branch; the union of this branch is a model of Γ.

Still without using the Axiom of Choice, we can obtain Compactness for an uncountable signature, provided the signature itself is given to us as being well ordered (otherwise we can apply the Well Ordering Theorem, page 98, which uses the Axiom of Choice).

Theorem 157 (Well Ordered Compactness). *Suppose $\mathscr{S}$ is a signature $\{s_\alpha \colon \alpha < \kappa\}$ for some cardinal κ, and Γ is a set of sentences of $\mathscr{S}$ whose every finite subset has a model. Then Γ has a model.*

Proof. Let A be a set $\{a_\alpha \colon \alpha < \kappa\}$ of constants not belonging to $\mathscr{S}$. It is possible to define a surjective function $\alpha \mapsto \sigma_\alpha$ from ω to $\mathrm{Sen}(\mathscr{S}(A))$. We shall recursively define a function $\alpha \mapsto \Gamma_\alpha$ from ω to $\mathscr{P}(\mathrm{Sen}(\mathscr{S}(A)))$ such that the union $\bigcup_{n\in\omega} \Gamma_n$ is a set T as in Theorem 155. Suppose, for some α such that $\alpha < \kappa$, a subset Γ_β of $\mathrm{Sen}(\mathscr{S}(A))$ has been defined whenever $\beta < \alpha$ so that

$$\gamma < \beta < \alpha \implies \Gamma_\gamma \subseteq \Gamma_\beta,$$

and also, whenever $\beta < \alpha$, the set Γ_β is a subset Δ of $\mathrm{Sen}(\mathscr{S}(A))$ such that

1) $|\Delta \smallsetminus \Gamma_0| < \kappa$, and

2) every finite subset of Δ has a model.

Then one of the two sets

$$\bigcup_{\beta<\alpha} \Gamma_\beta \cup \{\sigma_\beta\}, \qquad\qquad \bigcup_{\beta<\alpha} \Gamma_\beta \cup \{\neg\sigma_\beta\}$$

is also such a subset Δ of $\mathrm{Sen}(\mathscr{S}(A))$. Now we can obtain Γ_α from $\bigcup_{\beta<\alpha} \Gamma_\beta$, just as before we obtained Γ_{n+1} from Γ_n in the previous proof. □

As we observed above (page 163), a more algebraic formulation of the Compactness Theorem is that every filter of $\mathrm{Lin}_0(\mathscr{S})$ is $T/{\sim}$ for some theory T. To prove this, by the Boolean Prime Ideal Theorem, **PI** it is enough to show that every ultrafilter is $T/{\sim}$ for some complete theory T. We can proceed as follows, using Theorem 149 (page 153) and its corollary.

Theorem 158 (Compactness). *By the Boolean Prime Ideal Theorem, for all signatures $\mathscr{S}$, the map injective map $T \mapsto T/{\sim}$ from $\mathrm{S}_0(\mathscr{S})$ to $\mathrm{Sto}(\mathrm{Lin}_0(\mathscr{S}))$ is surjective.*

Proof. Suppose Γ is a subset of $\mathrm{Sen}(\mathscr{S})$ such that

$$\{\sigma^{\sim} \colon \sigma \in \Gamma\} \in \mathrm{Sto}(\mathrm{Lin}_0(\mathscr{S})).$$

We want to show Γ has a model $\mathfrak{A}$, since in that case

$$\mathrm{Th}(\mathfrak{A})/{\sim} = \{\sigma^{\sim} \colon \sigma \in \Gamma\}.$$

Let A be a set of constants not in $\mathscr{S}$. It will be enough to embed Γ in a subset of $\mathrm{Sen}(\mathscr{S}(A))$ that satisfies the hypothesis of Theorem 155 (page 164). Such a subset of $\mathrm{Sen}(\mathscr{S}(A))$ is precisely a set $\{\sigma \in \mathrm{Sen}(\mathscr{S}(A)) \colon \sigma^{\sim} \in \mathscr{U}\}$, where $\mathscr{U}$ is an element of $\mathrm{Sto}(\mathrm{Lin}_0(\mathscr{S}))$ that belongs to the intersection

$$\bigcap_{\varphi \in \mathrm{Fm}_{\{x\}}(\mathscr{S}(A))} \left([(\neg\exists x\ \varphi)^{\sim}] \cup \bigcup_{c \in A} [\varphi(c)^{\sim}] \right) \cap \bigcap_{\sigma \in \Gamma} [\sigma^{\sim}]. \tag{6.10}$$

Suppose there is an embedding $\varphi \mapsto c_\varphi$ of $\mathrm{Fm}_{\{x\}}(\mathscr{S}(A))$ in A. Then by the compactness of $\mathrm{Sto}(\mathrm{Lin}_0(\mathscr{S}(A)))$, we have

$$\bigcap_{\varphi\in\mathrm{Fm}_{\{x\}}(\mathscr{S}(A))}[(\exists x\ \varphi \Rightarrow \varphi(c_\varphi))^\sim] \cap \bigcap_{\sigma\in\Gamma}[\sigma^\sim] \neq \varnothing;$$

but the intersection here is a subset of the intersection in (6.10). Thus it is enough to define A as $\bigcup_{k\in\omega} A_k$, where

$$A_0 = \varnothing, \qquad\qquad A_1 = \left\{c_\varphi \colon \varphi \in \mathrm{Fm}_{\{x\}}(\mathscr{S})\right\},$$

and

$$A_{k+2} = \left\{c_\varphi \colon \varphi \in \mathrm{Fm}_{\{x\}}(\mathscr{S}(A_{k+1})) \smallsetminus \mathrm{Fm}_{\{x\}}(\mathscr{S}(A_k))\right\}. \qquad \square$$

Thus the Compactness Theorem as such needs only the Boolean Prime Ideal Theorem. We shall prove the converse as Theorem 168 (page 184).

6.4. Products

The following can be proved as a consequence of Theorem 158 and the Tarski–Vaught Test (page 159). But it is also just a reformulation of Theorem 155 (page 164).

Theorem 159. *Let $\mathscr{S}$ be a signature, and suppose T is a subset of* $\mathrm{Sen}(\mathscr{S})$ *such that*

1) $\{\sigma^\sim \colon \sigma \in T\} \in \mathrm{Sto}(\mathrm{Lin}_0(\mathscr{S}))$;
2) *for all sentences σ and τ of $\mathscr{S}$, if $\sigma \in T$ and $\sigma \sim \tau$, then $\tau \in T$;*
3) *for all singulary formulas $\varphi(x)$ of $\mathscr{S}$, for some closed term t of $\mathscr{S}$,* (6.2) *holds.*

Then T is a complete theory with a canonical model.

We now establish what amounts to a special case of this. Supposing $\mathscr{A}$ is an indexed family $(\mathfrak{A}_i \colon i \in \Omega)$ of structures with a common signature $\mathscr{S}$, we let

$$A = \prod_{i\in\Omega} A_i. \tag{6.11}$$

An element $(a_i \colon i \in \omega)$ of A may be written just as a. Understanding A as a set of new constants, we can expand each $\mathfrak{A}_i$ to a structure $\mathfrak{A}_i{}^*$ in the signature $\mathscr{S}(A)$ so that, for each a in A,

$$a^{\mathfrak{A}_i{}^*} = a_i. \tag{6.12}$$

Given σ in $\operatorname{Sen}(\mathscr{S}(A))$, we define

$$\|\sigma\|_{\mathscr{A}} = \{i \in \Omega \colon \mathfrak{A}_i{}^* \vDash \sigma\}.$$

We combine the $\mathfrak{A}_i$ into a single structure as follows. The usual reference for the theorem is Łoś's 1955 paper [44], although the theorem is not given clearly there.

Theorem 160 (Łoś's Theorem). *Suppose $\mathscr{A}$ is an indexed family*

$$(\mathfrak{A}_i \colon i \in \Omega)$$

of nonempty structures of $\mathscr{S}$, and A is as in (6.11). *Let $\mathscr{U}$ be an ultrafilter of $\mathscr{P}(\Omega)$. There is an equivalence relation E on A given by*

$$a \mathrel{E} b \iff \|a = b\|_{\mathscr{A}} \in \mathscr{U}.$$

By the Axiom of Choice, there is a structure $\mathfrak{B}$ of $\mathscr{S}(A)$ with universe A/E, such that, for all a in A, **AC**

$$a^{\mathfrak{B}} = \{b \in A \colon a \mathrel{E} b\},$$

and for all σ in $\operatorname{Sen}(\mathscr{S}(A))$,

$$\mathfrak{B} \vDash \sigma \iff \|\sigma\|_{\mathscr{A}} \in \mathscr{U}. \tag{6.13}$$

Proof. We have

$$\|\sigma \wedge \tau\|_{\mathscr{A}} = \|\sigma\|_{\mathscr{A}} \cap \|\tau\|_{\mathscr{A}}, \tag{6.14}$$

$$\|\neg\sigma\|_{\mathscr{A}} = \Omega \smallsetminus \|\sigma\|_{\mathscr{A}}. \tag{6.15}$$

Let

$$T = \{\sigma \in \operatorname{Sen}(\mathscr{S}(A)) \colon \|\sigma\|_{\mathscr{A}} \in \mathscr{U}\}. \tag{6.16}$$

By Theorem 146 (page 150), $T/{\sim}$ is an ultrafilter of $\mathrm{Lin}_0(\mathscr{S}(A))$, since

$$\begin{aligned}
\sigma \in T \;\&\; \tau \in T &\implies \|\sigma\|_{\mathscr{A}} \in \mathscr{U} \land \|\tau\|_{\mathscr{A}} \in \mathscr{U} && \text{[by (6.16)]}\\
&\implies \|\sigma\|_{\mathscr{A}} \cap \|\tau\|_{\mathscr{A}} \in \mathscr{U} && \text{[by Thm 146]}\\
&\implies \|\sigma \land \tau\|_{\mathscr{A}} \in \mathscr{U} && \text{[by (6.14)]}\\
&\implies \sigma \land \tau \in T && \text{[by (6.16)]}
\end{aligned}$$

and

$$\begin{aligned}
\sigma \in \mathrm{Sen}(\mathscr{S}(A)) \smallsetminus T &\iff \|\sigma\|_{\mathscr{A}} \in \mathscr{P}(\Omega) \smallsetminus \mathscr{U} && \text{[by (6.16)]}\\
&\iff \Omega \smallsetminus \|\sigma\|_{\mathscr{A}} \in \mathscr{U} && \text{[by Thm 146]}\\
&\iff \|\neg\sigma\|_{\mathscr{A}} \in \mathscr{U} && \text{[by (6.15)]}\\
&\iff \neg\sigma \in T && \text{[by (6.16)].}
\end{aligned}$$

AC Moreover, for every ψ in $\mathrm{Fm}_{\{x\}}(\mathscr{S}(A))$, by the Axiom of Choice, we can find a in A such that, for each i in Ω,

$$\mathfrak{A}_i \vDash \exists x\ \psi \iff \mathfrak{A}_i \vDash \psi(a_i). \tag{6.17}$$

Then

$$\|\exists x\ \psi\|_{\mathscr{A}} = \|\psi(a)\|_{\mathscr{A}},$$

so T is as in the previous theorem. □

The structure $\mathfrak{B}$ found in the theorem is an **ultraproduct** of the indexed family $\mathscr{A}$ or $(\mathfrak{A}_i \colon i \in \Omega)$ and can be denoted by one of

$$\prod \mathscr{A}/\mathscr{U}, \qquad \prod_{i\in\Omega} \mathfrak{A}_i/\mathscr{U}. \tag{6.18}$$

We may also denote an equivalence class $\{b \in A \colon a\ E\ b\}$ by

$$a/\mathscr{U}.$$

If $\mathscr{U}$ is merely a filter of $\mathscr{P}(\Omega)$, the quotient in (6.18) is still defined, but is called a **reduced product** of the indexed family.

In the proof of Łoś's Theorem, we need the Axiom of Choice only in the last step, involving quantifiers. If the ultrafilter $\mathscr{U}$ is principal, namely $\{X \in \mathscr{P}(\Omega) \colon i \in X\}$ for some i in Ω, then $\mathfrak{B} \cong \mathfrak{A}_i{}^*$. Thus Łoś's Theorem by itself does not imply even the Boolean Prime Ideal Theorem. However, these two theorems together imply the Axiom of Choice (Theorem 171, page 188).

Meanwhile, we can formulate the Compactness Theorem as a weaker version of Łoś's Theorem (with the Boolean Prime Ideal Theorem):

Theorem 161 (Compactness). *Suppose Γ is a set of sentences of a signature $\mathscr{S}$, and every finite subset Δ of Γ has a model $\mathfrak{A}_\Delta$. Let*

$$\mathscr{A} = (\mathfrak{A}_\Delta \colon \Delta \in \mathscr{P}_\omega(\Gamma)), \qquad A = \prod_{\Delta \in \mathscr{P}_\omega(\Gamma)} A_\Delta.$$

There is a complete theory of $\mathscr{S}(A)$ that includes Γ, namely

$$\{\sigma \in \mathrm{Sen}(\mathscr{S}(A)) \colon \|\sigma\|_{\mathscr{A}} \in \mathscr{U}\},$$

where $\mathscr{U}$ is an ultrafilter of $\mathscr{P}(\mathscr{P}_\omega(\Gamma))$ that contains each of the sets

$$\{\Delta \in \mathscr{P}_\omega(\Gamma) \colon \sigma \in \Delta\},$$

where $\sigma \in \Gamma$.

Proof. The indicated theory is the theory of the ultraproduct $\prod \mathscr{A}/\mathscr{U}$. □

6.5. Cardinality

By the theorem below, a non-principal ultrapower $\mathfrak{C}$ of a countably infinite structure $\mathfrak{A}$ is uncountable. By the Downward Löwenheim–Skolem Theorem (page 161), in a countable signature, there will then be a countable structure $\mathfrak{B}$ such that

$$\mathfrak{A} \prec \mathfrak{B} \prec C.$$

Indeed, $\mathfrak{B}$ may be chosen to include $A \cup \{x\}$ for some x in $C \smallsetminus A$. Even though $\mathfrak{A}$ is then a proper substructure of $\mathfrak{B}$, these two may be isomorphic. However, this is not the case when $\mathfrak{A}$ is $(\mathbb{N}, +, \cdot)$. Thus *countable non-standard models of arithmetic* exist. A more illuminating construction of such models is given in §7.7 below.

The following is a special case of [31, Thm 9.5.4(a)] (and is said to be found in Frayne, Morel, and Scott [19][2]).

Theorem 162. *For all signatures $\mathscr{S}$, for all $\mathfrak{A}$ in $\mathbf{Str}_{\mathscr{S}}$, for all singulary formulas φ of $\mathscr{S}(A)$, for all non-principal ultrafilters $\mathscr{U}$ of $\mathscr{P}(\omega)$,*

$$\omega \leqslant |\varphi(\mathfrak{A})| \implies |\varphi(\mathfrak{A}^{\omega}/\mathscr{U})| = |\varphi(\mathfrak{A})|^{\omega}.$$

In particular, if $\mathfrak{A}$ is countable, then all infinite definable relations of $\mathfrak{A}^{\omega}/\mathscr{U}$ have the cardinality of the continuum.

Proof. Suppose $a \in A^{\omega}$ and $a/\mathscr{U} \in \varphi(\mathfrak{A}^{\omega}/\mathscr{U})$. then by Łoś's Theorem

$$\|\varphi(a)\|_{\mathscr{A}} \in \mathscr{U}.$$

Then we may assume $\|\varphi(a)\|_{\mathscr{A}} = \omega$. That is, we can find a' in $\varphi(\mathfrak{A})^{\omega}$ such that $a/\mathscr{U} = a'/\mathscr{U}$. By the Axiom of Choice then, there is an injection $a/\mathscr{U} \mapsto a'$ from $\varphi(\mathfrak{A}^{\omega}/\mathscr{U})$ to $\varphi(\mathfrak{A})^{\omega}$. This shows **AC**

$$|\varphi(\mathfrak{A}^{\omega}/\mathscr{U})| \leqslant |\varphi(\mathfrak{A})|^{\omega}.$$

For the reverse inequality when $\varphi(\mathfrak{A})$ is infinite, it is enough to find a function $a \mapsto a^*$ from $\varphi(\mathfrak{A})^{\omega}$ to itself such that

$$a \neq b \implies a^*/\mathscr{U} \neq b^*/\mathscr{U},$$

so that $a \mapsto a^*/\mathscr{U}$ will be an embedding of $\varphi(\mathfrak{A})^{\omega}$ in $\varphi(\mathfrak{A}^{\omega}/\mathscr{U})$. We want

$$a \neq b \implies \|a^* \neq b^*\|_{\mathscr{A}} \in \mathscr{U}.$$

[2] I have a printout of this article, but have not sorted through all of its many basic results to find this one. It should be noted that the article has a correction' [20], which merely refines the account of Tarski's contribution to the subject (as well as taking some of the credit away from Frayne).

Now, $a \neq b$ means $a_m \neq b_m$ for some m in ω. For each m in ω, we have $\omega \smallsetminus m \in \mathscr{U}$. Thus it is enough if

$$a_m \neq b_m \implies \omega \smallsetminus m \subseteq \|a^* \neq b^*\|_{\mathscr{A}},$$

that is,

$$a_m \neq b_m \ \& \ m \leqslant n \implies a^*{}_n \neq b^*{}_n. \tag{6.19}$$

For this, it is enough if, for each n in ω, a^*n is an injective function of $(a_0, \dots, a_n)$. So let μ_n be an injection from $\varphi(\mathfrak{A})^{n+1}$ to $\varphi(\mathfrak{A})$ (which exists because $\varphi(\mathfrak{A})$ is infinite), and define

$$a^*{}_n = \mu_n(a_0, \dots, a_i). \qquad \square$$

Let us try to generalize this argument, replacing ω with an arbitrary infinite index-set Ω. Instead of elements m and n of ω, we work with elements i and j of Ω. We replace the element $\omega \smallsetminus m$ of the ultrafilter of $\mathscr{P}(\omega)$ with some element X_i of the ultrafilter of $\mathscr{P}(\Omega)$. The old condition $m \leqslant n$ is now $j \in X_i$, so that (6.19) becomes

$$a_i \neq b_i \ \& \ j \in X_i \implies a^*{}_j \neq b^*{}_j,$$

and $(a_0, \dots, a_n)$, which is $(a_m \colon m \leqslant n)$, becomes $(a_i \colon j \in X_i)$. So $a^*{}_j$ should be an injective function of this. As before, it is enough if the sets

$$\{i \in \Omega \colon j \in X_i\}$$

are finite. An ultrafilter of $\mathscr{P}(\Omega)$ is called **regular** if it has such elements X_i for all i in Ω.

Theorem 163. *There are regular ultrafilters of $\mathscr{P}(\Omega)$ for every infinite set Ω.*

Proof. Let Ω be an infinite set. Then Ω is equipollent with $\mathscr{P}_\omega(\Omega)$. So it is enough to show that there are regular ultrafilters of $\mathscr{P}(\mathscr{P}_\omega(\Omega))$. To do this, if $i \in \mathscr{P}_\omega(\Omega)$, we need only define

$$X_i = \{j \in \mathscr{P}_\omega(\Omega) \colon i \subseteq j\}.$$

Since $X_i \cap X_j = X_{i \cup j}$, the X_i do generate a filter of $\mathscr{P}(\mathscr{P}_\omega(\Omega))$. The filter is proper, since $i \in X_i$, so none of the X_i is empty. Moreover,

$$\{i \in \mathscr{P}_\omega(\Omega) \colon j \in X_i\} = \{i \in \mathscr{P}_\omega(\Omega) \colon i \subseteq j\} = \mathscr{P}(j),$$

which is finite. So there are regular proper filters, and hence regular ultrafilters, of $\mathscr{P}(\mathscr{P}_\omega(\Omega))$. □

Hence, in Theorem 162, ω can be replaced with an arbitrary infinite set.

6.6. Convergence of ultrafilters

There are a number of equivalent formulations of the definition of compactness of a topological space.[3] We shall use one of them to understand Łoś's Theorem (Theorem 160, page 173) more clearly as being a refinement of the Compactness Theorem (as Theorem 161, page 175), given the model theory of the previous chapter (and in particular the Tarski–Vaught Test (page 159).

In a topological space, an **open neighborhood** of a point is an open set that contains the point. Then a **neighborhood** of the point is a set that includes an open neighborhood of the point. For an arbitrary set Ω, a **filter on** the set Ω is just a **filter of** the Boolean algebra $\mathscr{P}(\Omega)$.

Lemma 14. *The set of all neighborhoods of a point of a topological space is a proper filter on the space.*

A filter on a topological space

- **clusters** at a point of the space if the union of this filter with the filter of neighborhoods of the point generates a proper filter (that is, every set in the former filter has nonempty intersection with every set of the latter filter);

[3] See for example Willard [62, Thm 17.4, p. 118].

- **converges** to a point of the space if the filter includes the filter of neighborhoods of the point.

A point where a filter clusters is a **cluster point** of the filter.

A cluster point of a filter need not belong to the intersection of the filter. For example, in $\mathbb{R}$, every point of the interval $[0,1]$ is a cluster point of the filter generated by $(0,1)$.

The **closure** of a subset of a topological space is the smallest closed set that includes the subset.

Lemma 15. *A cluster point of a filter belongs to the closure of every set in the filter.*

Theorem 164. *A topological space is compact if and only if every proper filter on it has a cluster point.*

Proof. Let (A,τ) be a topological space. Suppose first the space is compact, and let $\mathscr{F}$ be a proper filter on A. Then $\mathscr{F}\cap\tau$ is closed under taking finite intersections, and in particular all such intersections are nonempty, so $\bigcap(\mathscr{F}\cap\tau)$ must contain a point p, by the compactness of τ. In particular, if $F\in\tau$ and $p\notin F$, that is, if $A\smallsetminus F$ is an open neighborhood of p, then $F\notin\mathscr{F}$, so $\mathscr{F}\cup\{A\smallsetminus F\}$ generates a proper filter.

Now suppose conversely that every filter on A has a cluster point, and let $\mathscr{F}$ be a subset of τ whose every finite subset has nonempty intersection. Then $\mathscr{F}$ generates a filter on A, and this filter clusters at some point p. In this case, by the lemma, $p\in\bigcap\mathscr{F}$. □

An ultrafilter with a cluster point converges to that point. Thus, on a compact space, every ultrafilter converges.

Theorem 165. *By the Prime Ideal Theorem, if every ultrafilter on a topological space converges, the space is compact.* **PI**

We shall want to allow the possibility that an ultrafilter on a *subspace* of a topological space converges to a point of the larger space. For this, we can use the following observation.

Lemma 16. *If A and B are sets, and $A \subseteq B$, and $\mathscr{U}$ is an ultrafilter on A, then the filter on B that $\mathscr{U}$ generates is*

$$\{X \subseteq B \colon X \cap A \in \mathscr{U}\},$$

and this is an ultrafilter on B.

In the situation of the lemma, if B is actually a topological space, we may say that $\mathscr{U}$ **converges** to a point of B if the ultrafilter on B that $\mathscr{U}$ generates converges to the point.

Theorem 166. *Suppose A and B are Boolean algebras, f is a homomorphism from A to B, and $\mathscr{U}$ is an ultrafilter of B. Then $f^{-1}[\mathscr{U}]$ is an ultrafilter of A.*

Proof. If x and y are in $f^{-1}[\mathscr{U}]$, then $f(x \wedge y) = f(x) \wedge f(y)$, which is in $\mathscr{U}$, so $x \wedge y \in f^{-1}[\mathscr{U}]$. If $z \in A$, then, since $f(\neg z) = \neg f(z)$, we have

$$\neg z \in f^{-1}[\mathscr{U}] \iff \neg f(z) \in \mathscr{U} \iff f(z) \notin \mathscr{U} \iff z \notin f^{-1}[\mathscr{U}].$$

□

Now we can expand the theorem that Stone spaces of Boolean algebras are compact by considering also *subspaces* of Stone spaces.

Theorem 167. *Suppose A is a Boolean algebra, $U \in \mathrm{Sto}(A)$, $\Omega \subseteq \mathrm{Sto}(A)$, and $\mathscr{U}$ is an ultrafilter on Ω. Then $\mathscr{U}$ converges to U if and only if, for all x in A,*

$$U \in [x] \iff [x] \cap \Omega \in \mathscr{U}.$$

The set

$$\{x \in A \colon [x] \cap \Omega \in \mathscr{U}\}$$

is an element of $\mathrm{Sto}(A)$, and therefore $\mathscr{U}$ converges to this point.

Proof. By the Stone Representation Theorem for Algebras (page 151), the map $x \mapsto [x] \cap \Omega$ from A to $\mathscr{P}(\Omega)$ is a homomorphism of Boolean algebras, so by the last theorem, the given set is an ultrafilter U of A. Then

$$U \in [x] \iff x \in U \iff [x] \cap \Omega \in \mathscr{U},$$

so $\mathscr{U}$ converges to U. □

Letting Ω be $\mathrm{Sto}(A)$ itself (and assuming the Prime Ideal Theorem), we obtain a neat proof that $\mathrm{Sto}(A)$ is compact. Similarly, we **PI** shall obtain the Compactness Theorem from Łoś's Theorem. In this context, we take A to be $\mathrm{Lin}_0(\mathscr{S})$ and Ω to be the image of $\mathrm{S}_0(\mathscr{S})$ under $T \mapsto T/{\sim}$.

Corollary 167.1. *By the Prime Ideal Theorem,* $\mathrm{S}_0(\mathscr{S})$ *is compact if and only if, for every ultrafilter* $\mathscr{U}$ *on this space, there is* $\mathfrak{B}$ *in* $\mathbf{Str}_{\mathscr{S}}$ *such that, for every* σ *in* $\mathrm{Sen}(\mathscr{S})$, **PI**

$$\mathfrak{B} \vDash \sigma \iff [\sigma^{\sim}] \in \mathscr{U}. \tag{6.20}$$

Proof. $\mathfrak{B} \vDash \sigma \iff \mathrm{Th}(\mathfrak{B})/{\sim} \in [\sigma^{\sim}]$. □

Now, each T in $\mathrm{S}_0(\mathscr{S})$ has a model $\mathfrak{A}_T$. Thus we obtain an indexed family $(\mathfrak{A}_T \colon T \in \mathrm{S}_0(\mathscr{S}))$ of structures of $\mathscr{S}$, and then we have

$$\begin{aligned}[\sigma^{\sim}] &= \{T \in \mathrm{S}_0(\mathscr{S}) \colon \sigma \in T\}\\ &= \{T \in \mathrm{S}_0(\mathscr{S}) \colon \mathfrak{A}_T \vDash \sigma\}\\ &= \|\sigma\|_{\mathscr{A}}.\end{aligned}$$

We now have that (6.20) is equivalent to (6.13) in Łoś's Theorem. Therefore the compactness of $\mathrm{S}_0(\mathscr{S})$ follows from this theorem if we let

$$\mathfrak{B} = \prod_{T \in \mathrm{S}_0(\mathscr{S})} \mathfrak{A}_T/\mathscr{U}.$$

Conversely, we shall derive Łoś's Theorem from the Compactness Theorem and the Tarski–Vaught Test—and the Axiom of Choice. Suppose $(\mathfrak{A}_i \colon i \in \Omega)$ is an indexed family of nonempty structures of $\mathscr{S}$. We may assume that the map $i \mapsto \mathrm{Th}(\mathfrak{A}_i)$ from Ω to $\mathrm{S}_0(\mathscr{S})$ is injective. (Otherwise we could enlarge $\mathscr{S}$ to contain a nullary predicate P_i

for each i in Ω, and we could define P_i to be true in $\mathfrak{A}_j$ if and only if $i = j$.) Then we may assume Ω is a subset of $\mathrm{S}_0(\mathscr{S})$.

We can define A as in (6.11) (on page 172) and expand each $\mathfrak{A}_i$ to a structure $\mathfrak{A}_i{}^*$ of $\mathscr{S}(A)$ as in (6.12). Now using the map $i \mapsto \mathrm{Th}(\mathfrak{A}_i{}^*)$, we may assume Ω is a subset of $\mathrm{S}_0(\mathscr{S}(A))$. Suppose $\mathscr{U}$ is an ultrafilter on Ω. By the Compactness Theorem, $\mathscr{U}$ converges to some point $\mathrm{Th}(\mathfrak{C})$ of $\mathrm{S}_0(\mathscr{S}(A))$. This means (6.20) holds, when $\mathfrak{B}$ is $\mathfrak{C}$, for all σ in $\mathrm{Sen}(\mathscr{S}(A))$.

But the structure $\mathfrak{C}$ has a substructure $\mathfrak{B}$ whose universe B is $\{a^{\mathfrak{C}} \colon a \in A\}$. Indeed, for every positive integer n, if F is an n-ary operation symbol of $\mathscr{S}$, and $(a^i \colon i < n) \in A^n$, let

$$b = \bigl(F^{\mathfrak{A}_i}(a^j \colon j < n) \colon i \in \Omega\bigr).$$

Then $\mathfrak{C} \models Fa^0 \cdots a^{n-1} = b$. Thus $\mathfrak{B}$ is well defined and $\mathfrak{B} \subseteq \mathfrak{C}$.

AC Let $\psi(x)$ be a singulary formula of $\mathscr{S}(A)$, and as in the proof of Łoś's Theorem, using the Axiom of Choice, let a in A be such that, for each i in Ω, (6.17) holds. Then

$$\mathfrak{C} \models \exists x\ \psi \iff \mathfrak{C} \models \psi(a).$$

By the Tarski–Vaught Test (page 159), $\mathfrak{B} \preccurlyeq \mathfrak{C}$. Then (6.20) holds as it is, which means Łoś's Theorem holds.

6.7. Closed sets

Another way to think about Łoś's Theorem and the Compactness Theorem is as follows. First note that, by Corollary 149.1 (page 154), for all signatures $\mathscr{S}$, the space $\mathrm{S}_0(\mathscr{S})$ of complete theories of T is compact if and only if its image under $T \mapsto T/{\sim}$ is a closed subspace of $\mathrm{Sto}(\mathrm{Lin}_0(\mathscr{S}))$.

Lemma 17. *If A and B are sets, and $A \subseteq B$, and $\mathscr{U}$ is an ultrafilter on B, then the set*

$$\{X \cap A \colon X \in \mathscr{U}\}$$

is a filter on A, and if it is a proper filter, it is an ultrafilter.

Lemma 18. *Suppose (B, τ) is a topological space, and $A \subseteq B$.*

1. *$A \in \tau$ if and only if no ultrafilter on A converges to a point of $B \smallsetminus A$.*
2. *Suppose further that (B, τ) is Hausdorff. Then $A \in \tau$ if and only if every convergent ultrafilter on A converges to a point of A.*

Proof. Suppose $\mathscr{U}$ is an ultrafilter on A that converges to a point p of $B \smallsetminus A$. For every open neighborhood U of p, we must have $U \cap A \in \mathscr{U}$, and in particular $U \cap A \neq \varnothing$. Thus $B \smallsetminus A$ cannot be an open neighborhood of p, so A is not closed.

Conversely, if A is not closed, then $B \smallsetminus A$ has a point p whose every open neighborhood contains a point of A. Let $\mathscr{U}$ be an ultrafilter on B that includes the filter of neighborhoods of p. Then $\{X \cap A \colon X \in \mathscr{U}\}$ is a proper filter on A and therefore an ultrafilter, but it converges to p. □

Thus $\mathrm{S}_0(\mathscr{S})$ is compact if and only if, on its image under $T \mapsto T/{\sim}$, every ultrafilter converges to an element of this image. But Łoś's Theorem establishes this convergence, as before.

7. Applications

7.1. The Prime Ideal Theorem

We establish now the mutual equivalence of the following.

1. The Boolean Prime Ideal Theorem (page 102).
2. The Prime Ideal Theorem (page 101).
3. The Tychonoff Theorem (page 129) restricted to Hausdorff spaces (page 114).
4. The Compactness Theorem (page 171).

In 1954, Dana Scott [53] announced that the Boolean Prime Ideal Theorem implies the Prime Ideal Theorem.[1] It is not clear what proof he had in mind. Since the Boolean Prime Ideal Theorem implies the Compactness Theorem (171), we can establish the result as follows. The *diagram* of a structure was defined on page 156.

Theorem 168. *The Compactness Theorem implies the Prime Ideal Theorem.*

Proof. Let $\mathscr{S}$ be the signature of commutative rings, let T be the theory of nontrivial commutative rings in this signature, and let $\mathfrak{R} \vDash T$. Let P be a new singulary predicate. Every finite subset Γ of the collection

$$\begin{aligned}T \cup \operatorname{diag}(\mathfrak{R}) \cup \{P0,\ \neg P1\}& \\ \cup \{Pa \land Pb \Rightarrow P(a-b)\colon a \in R\ \&\ b \in R\}& \\ \cup \{P(ab) \Leftrightarrow Pa \lor Pb\colon a \in R\ \&\ b \in R\}&\end{aligned}$$

[1] Scott spoke at "the five hundred third meeting of the American Mathematical Society. . . held at Yosemite National Park on Saturday, May 1, 1954." Thanks to Wilfrid Hodges for giving me the reference, which is not available on MathSciNet.

of sentences of $\mathscr{S}(R) \cup \{P\}$ has a model. Indeed, suppose A is the set of elements of R appearing in Γ. Then A generates a finite sub-ring $\mathfrak{B}$ of $\mathfrak{R}$, and by Theorem 89 (page 97), $\mathfrak{B}$ has a maximal ideal $\mathfrak{m}$, which is prime by Corollary 85.1 (page 94). Then $\mathfrak{B}$ expands to the model $(\mathfrak{B}_A, \mathfrak{m})$ of Γ. By Compactness, the whole collection above has a model $(\mathfrak{S}_R, \mathfrak{p})$, where $\mathfrak{S}$ is a ring with prime ideal $\mathfrak{p}$, and (by Theorem 150, page 156), $\mathfrak{R}$ is a sub-ring of $\mathfrak{S}$. Then $R \cap \mathfrak{p}$ is a prime ideal of $\mathfrak{R}$. □

By Theorems 95 and 96 (page 101), the Axiom of Choice and the Maximal Ideal Theorem are equivalent. By Theorems 131 and 132 (page 129), the Axiom of Choice and the Tychonoff Theorem are equivalent. We shall now establish that a weaker form of the Maximal Ideal Theorem, namely the Prime Ideal Theorem, is equivalent to a weaker form of the Tychonoff Theorem. The remaining theorems of this section are due to Łoś and Ryll-Nardzewski [42, 43].

Theorem 169. *The Boolean Prime Ideal Theorem implies the Tychonoff Theorem for Hausdorff spaces.*

Proof. Suppose $\mathscr{A}$ is an indexed family $(A_i : i \in \Omega)$ of nonempty Hausdorff spaces. We first show that its product is nonempty. Let

$$B = \bigcup_{\Gamma \subseteq \Omega} \prod_{i \in \Gamma} A_i,$$

and if $j \in \Omega$, let

$$B_j = \bigcup_{\{j\} \subseteq \Gamma \subseteq \Omega} \prod_{i \in \Gamma} A_i,$$

The B_j generate a proper filter on B, since for all n in ω, if σ is an injection from n into Ω, then

$$\varnothing \subset \prod_{i<n} A_{\sigma(i)} \subseteq B_{\sigma(0)} \cap \cdots \cap B_{\sigma(n-1)}.$$

Using the Boolean Prime Ideal Theorem, we let $\mathscr{U}$ be an ultrafilter that includes this filter. We shall derive from this an ultrafilter on each A_i. If $p \in A_i$, let

$$C_i(p) = \{a \in B_i \colon a_i = p\}.$$

Then for all p and q in A_i,

$$p \neq q \implies C_i(p) \cap C_i(q) = \varnothing.$$

Thus the map $X \mapsto \bigcup_{p \in X} C_i(p)$ from $\mathscr{P}(A_i)$ to $\mathscr{P}(B)$ is a homomorphism h_i of Boolean algebras. By Theorem 166 (page 180), ${h_i}^{-1}[\mathscr{U}]$ is an ultrafilter $\mathscr{U}_i$ on A_i. Since A_i is compact, $\mathscr{U}_i$ converges to a point of A_i; since A_i is also Hausdorff, $\mathscr{U}_i$ converges to a *unique* point a_i of A_i. Then $(a_i \colon i \in \Omega) \in \prod \mathscr{A}$.

We finished our proof of the general Tychonoff Theorem by noting that the product of nonempty closed subsets of the A_i is nonempty. To reach this point, we used Zorn's Lemma. But when the A_i are Hausdorff, we need only the Boolean Prime Ideal Theorem. Indeed, suppose now $\mathscr{X}$ is a family of closed subsets of $\prod \mathscr{A}$ with the finite intersection property. Then $\mathscr{X}$ generates a proper filter on $\prod \mathscr{A}$, and by Theorem 99 (page 102), this filter is included in an ultrafilter $\mathscr{U}$. (Note that this conclusion requires $\prod \mathscr{A}$ to be nonempty, so that $\mathscr{P}(\prod \mathscr{A})$ is a nontrivial ring.) For each i in Ω, the set

$$\{\pi_i[X] \colon X \in \Omega\}$$

is an ultrafilter on A_i (why?), so it converges to some a_i, which is unique since A_i is Hausdorff. Then $(a_i \colon i \in \Omega) \in \bigcap \mathscr{X}$. Therefore $\prod \mathscr{A}$ is compact. □

Lemma 19. *The Tychonoff Theorem for Hausdorff space implies that, whenever $\mathscr{A}$ is a family $(A_i \colon i \in \Omega)$ of nonempty compact Hausdorff spaces, and moreover there is a symmetric binary relation E on $\bigcup_{i \in \Omega} A_i$ such that*

- *for all distinct i and j in Ω, the subset $\{(x, y) \in A_i \times A_j \colon x \mathrel{E} y\}$ of $A_i \times A_j$ is closed, and also,*

- *for every finite subset Ω_0 of Ω, for some x in $\prod\mathscr{A}$, for all distinct i and j in Ω_0, $x \mathrel{E} y$,*

then the latter condition holds when $\Omega_0 = \Omega$.

Proof. If $X \subseteq \Omega$, let

$$T(X) = \left\{ x \in \prod\mathscr{A} \colon \bigwedge_{\substack{\{i,j\}\subseteq X\\ i\neq j}} x_i \mathrel{E} x_j \right\}.$$

By hypothesis, when X is finite, then $T(X)$ is nonempty. Moreover,

$$T(X) = \bigcap_{\substack{\{i,j\}\subseteq X\\ i\neq j}} T(\{i,j\}),$$

so this is closed. An element of the intersection

$$\bigcap_{\substack{X\subseteq\Omega\\ |X|<\omega}} T(X)$$

would be the desired element of $\prod\mathscr{A}$; since this product is compact, the desired element exists. □

Theorem 170. *The Tychonoff Theorem for Hausdorff spaces implies the Boolean Prime Ideal Theorem.*

Proof. Let R be a Boolean ring, and let Ω be the set of finitely generated nontrivial sub-rings of R. These will be just the nontrivial *finite* sub-rings of R. Then $(\operatorname{Spec}(B) \colon B \in \Omega)$ is a family of nonempty compact Hausdorff spaces: we have this without any special assumption, by Theorem 89, page 97. Then E is as in the hypothesis of the lemma when, if B and C are distinct elements of Ω, and $\mathfrak{p} \in \operatorname{Spec}(B)$ and $\mathfrak{q} \in \operatorname{Spec}(C)$,

$$\mathfrak{p} \mathrel{E} \mathfrak{q} \iff \mathfrak{p} \cap C = \mathfrak{q} \cap B.$$

Let $(\mathfrak{p}_B \colon B \in \Omega)$ be as guaranteed by the lemma. Then $\bigcup_{B\in\Omega} \mathfrak{p}_B$ is a prime ideal of R. □

7.2. The Axiom of Choice

A function f on a set A of nonempty sets is a **choice function** if for all b in A, $f(b) \in b$. Then the Axiom of Choice is equivalent to the statement that every set of nonempty sets has a choice function.

The following result was published by Howard in 1975 [33].

Theorem 171. *The Boolean Prime Ideal Theorem and Łoś's Theorem together imply the Axiom of Choice.*

Proof. Let A be a set of nonempty sets that does not have a choice function. Let $\Omega = \bigcup A \cup A$,[2] and let

$$R = \Big\{(x, y) \in \bigcup A \times A \colon x \in y\Big\} \cup \Big\{(x, x) \colon x \in \bigcup A\Big\}.$$

Then

$$(\Omega, R) \vDash \forall y\ \exists x\ x\ R\ y.$$

The subsets of A on which there *is* a choice function constitute a proper ideal $\mathscr{I}$ on A. Let $\mathscr{U}$ be an ultrafilter on Ω that includes the dual filter $\{\Omega \smallsetminus X \colon X \in \mathscr{I}\}$ on A. Then

$$\prod_{i \in \Omega} (\Omega, R)/\mathscr{U} \vDash \forall y\ \exists x\ x\ R\ y.$$

In particular, for the element $\big((i, i) \colon i \in \Omega\big)$ of Ω^Ω, there exists an element $(a_i \colon i \in \Omega)$ such that

$$\{i \in \Omega \colon a_i\ R\ i\} \in \mathscr{U}.$$

Let $B = \{i \in A \colon a_i\ R\ i\}$. Then $\{(i, a_i) \colon i \in B\}$ is a choice function on B. However, by assumption, there is a choice function also on $A \smallsetminus B$. Hence there is a choice function on A. □

[2] Howard notes that we may assume the elements of A pairwise disjoint, and that we may assume A and $\bigcup A$ are disjoint.

7.3. Arrow's Theorem

This section is inspired by Sasha Borovik's article [6]. We consider an index-set Ω as a set of *voters.* Each voter i in Ω is called on to assign a linear ordering $<_i$ to a set A of *candidates.* These orderings are to be used to assign a linear ordering $<$ to A. This ordering $<$ should be a kind of average of the orderings $<_i$. This suggests that we should take an ultraproduct of the structures $(A, <_i)$. We shall see that, on some reasonable assumptions, we *must* do this.

We want to determine $<$ by first selecting a subset D of $\mathscr{P}(\Omega)$ such that, for all x and y in A, we shall be able to require

$$\{i\colon x <_i y\} \in D \implies x < y.$$

So D will be, so to speak, a collection of 'winning coalitions'. If $X \in D$, then the members of X can determine how the candidates in A shall be ordered (if all members of X agree). Then we must have, first of all,

$$D \neq \varnothing,$$
$$X \in D \implies X^{\mathrm{c}} \notin D.$$

We also require that additional votes for a particular ordering can only help that ordering:

$$X \in D \ \& \ X \subseteq Y \subseteq \Omega \implies Y \in D.$$

Hence in particular $\Omega \in D$. We require voting to be decisive:

$$X \notin D \implies X^{\mathrm{c}} \in D.$$

If A consists of just two candidates, this is all we need. Then D is not necessarily an ultrafilter on Ω; for it need not be closed under intersections. Indeed, in the 'democratic' case, if Ω has a finite number $2n-1$ of members, then D will be $\{X \in \mathscr{P}(\Omega)\colon |X| \geqslant n\}$; this is definitely not closed under intersections unless $n = 1$.

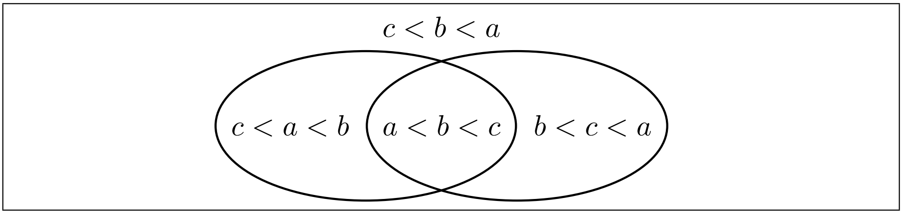

Figure 7.1.: An election with three candidates

But now suppose A contains three distinct candidates, a, b, and c; and let

$$\{i\colon a <_i b\} = A, \qquad\qquad \{i\colon b <_i c\} = B.$$

Suppose both A and B are in D. Then we must conclude $a < b$ and $b < c$ and therefore $a < c$. We have now

$$A \cap B \subseteq \{i\colon a <_i c\}, \qquad\qquad \{i\colon a <_i c\} \in D.$$

However, possibly

$$A \cap B = \{i\colon a <_i c\};$$

this is the case when—as is possible—

$$\begin{gathered}\{i\colon c <_i a <_i b\} = A \smallsetminus B,\\ \{i\colon b <_i c <_i a\} = B \smallsetminus A,\\ \{i\colon c <_i b <_i a\} = (A \cup B)^{\mathrm{c}}.\end{gathered}$$

See Figure 7.1. Thus we must have $A \cap B \in D$. Therefore D is an ultrafilter on Ω. If Ω is finite, then D must be a principal ultrafilter: that is, one voter decides everything, and the system is a dictatorship.

7.4. Completeness of theories

Using the Compactness Theorem, we can establish a complement to Theorem 153 (page 161):

Theorem 172 (Upward Löwenheim–Skolem)**.** *If* $\mathfrak{A}$ *is an infinite structure with signature* $\mathscr{S}$*, and* $\max(|A|, |\sigma|) \leqslant \kappa$*, then there is a structure* $\mathfrak{B}$ *such that*

$$\mathfrak{A} \preccurlyeq \mathfrak{B}, \qquad\qquad |B| = \kappa.$$

Proof. Let C be a set $\{c_\alpha \colon \alpha < \kappa\}$ be a set of new constants, all distinct. By Compactness, the set

$$\mathrm{Th}(\mathfrak{A}_A) \cup \{c_\alpha \neq c_\beta \colon \alpha < \beta < \kappa\}$$

of sentences has a model $\mathfrak{D}_{A \cup C}$. By construction, this model has cardinality at least κ. By the downward version of the theorem, $\mathfrak{D}$ has an elementary substructure $\mathfrak{B}$ of size κ such that $A \subseteq B$. Since also $\mathfrak{A} \preccurlyeq \mathfrak{D}$, the structure $\mathfrak{B}$ is as desired. □

This theorem yields an easy test for completeness of theories. For an infinite cardinal κ, a theory is **κ-categorical** if all of its models of size κ are isomorphic to one another.

Theorem 173 (Łoś–Vaught Test)**.** *If a theory* T *of a signature* $\mathscr{S}$ *has models, but no finite models;* $|\mathscr{S}| \leqslant \kappa$*; and* T *is* κ*-categorical; then* T *is complete.*

Proof. If T contains neither σ nor $\neg\sigma$, then both $T \cup \{\neg\sigma\}$ and $T \cup \{\sigma\}$ have models, which must be infinite. Then by the Löwenheim–Skolem–Tarski theorems (both upward and downward forms may be needed), each of the two sets has a model of cardinality κ; but these two models cannot be isomorphic to one another. □

Algebraically closed fields are defined on page 226.

Theorem 174.

- *The theory of algebraically closed fields of characteristic* 0 *is complete.*
- *For all primes* p*, the theory of algebraically closed fields of characteristic* p *is complete.*

Proof. None of these theories has no finite models. Every algebraically closed field is determined up to isomorphism by its characteristic and its transcendence-degree. If κ is uncountable, then a field with transcendence-degree κ has cardinality κ. Now the Łoś–Vaught Test applies. □

Similarly we have the following (see page 158 above):

Theorem 175. *The theory of algebraically closed fields is model-complete.*

Proof. If T is this theory, $K \vDash T$, and $|K| < \kappa$, then $T \cup \operatorname{diag}(K)$ is κ-categorical, but has no finite models. □

We can also now prove the converse of the lemma on page 158 above.

Theorem 176. *For all theories T, the models of $T_\forall$ are precisely the substructures of models of T.*

Proof. Assuming $\mathfrak{A} \vDash T_\forall$, we want to show $T \cup \operatorname{diag}(\mathfrak{A})$ has a model. By Compactness, and since $\operatorname{diag}(\mathfrak{A})$ is closed under conjunction, it is enough to show $T \cup \{\vartheta(\boldsymbol{a})\}$ has a model whenever ϑ is a quantifier-free formula of $\mathscr{S}$ and $\mathfrak{A} \vDash \vartheta(\boldsymbol{a})$. If it has no model, then $T \vdash \neg\vartheta(\boldsymbol{a})$, so (since no entry of $\boldsymbol{a}$ is in $\mathscr{S}$) $T \vdash \forall \boldsymbol{x}\ \neg\vartheta(\boldsymbol{x})$, and therefore $\mathfrak{A} \vDash \forall \boldsymbol{x}\ \neg\vartheta(\boldsymbol{x})$, which is absurd. □

In particular, when T is just field-theory, then $T_\forall$ is the theory of integral domains, by Corollary 121.1 (page 123).

7.5. Elementary classes

In [44] Łoś defined ultraproducts (but not by that name) in order to state the following algebraic test for being an elementary class of structures.

Theorem 177. *A subclass of $\mathbf{Str}_{\mathscr{S}}$ is elementary if and only if it contains:*

- *every structure that is elementarily equivalent to a member, and*

- *every ultraproduct of members.*

Proof. The 'only if' direction is the easier. An elementary class is the class of models of some theory T. If the class is $\mathcal{K}$, and $\mathfrak{A} \in \mathcal{K}$, and $\mathfrak{A} \equiv \mathfrak{B}$, then $\mathfrak{B} \vDash T$, so $\mathfrak{B} \in \mathcal{K}$. If $\{\mathfrak{A}_i \colon i \in \Omega\} \subseteq \mathcal{K}$, then $\mathfrak{A}_i \vDash T$ in each case, so every ultraproduct of the $\mathfrak{A}_i$ is a model of T, by Łoś's Theorem.

The more difficult direction is 'if'. Suppose $\mathcal{K}$ is a non-elementary subclass of $\mathbf{Str}_{\mathscr{S}}$. Then there is a model $\mathfrak{B}$ of $\mathrm{Th}(\mathcal{K})$ that does not belong to $\mathcal{K}$. However, every element σ of $\mathrm{Th}(\mathfrak{B})$ has a model in $\mathcal{K}$, since otherwise $\neg\sigma$ would be in $\mathrm{Th}(\mathcal{K})$. Therefore every finite subset Δ of $\mathrm{Th}(\mathfrak{B})$ has a model $\mathfrak{A}_\Delta$ in $\mathcal{K}$ (since otherwise the negation of the conjunction of the members of Δ would be in $\mathrm{Th}(\mathcal{K})$). By (the proof of) the Compactness Theorem, some ultraproduct of $(\mathfrak{A}_\Delta \colon \Delta \in \mathscr{P}_\omega(\mathrm{Th}(\mathfrak{B})))$ is elementarily equivalent to $\mathfrak{B}$. □

7.6. Saturation

If V is a finite set of variables, a **V-type** is just a subset of $\mathrm{Fm}_V(\mathscr{S})$. A V-type is **complete** if its image in $\mathrm{Lin}_V(\mathscr{S})$ under $\varphi \mapsto \varphi^\sim$ is an ultrafilter. Usually $V = \{x_0, \dots, x_{n-1}\}$, and then V-types are called n-types. In this case, a subset Γ of $\mathrm{Fm}_V(\mathscr{S})$ is a complete type if and only if

- for all φ in $\mathrm{Fm}_V(\mathscr{S})$, exactly one of φ and $\neg\varphi$ is in Γ, and
- for all finite subsets $\{\varphi_0, \dots, \varphi_{m-1}\}$ of Γ, there is a model of

$$\exists x_0 \cdots \exists x_{n-1} \, (\varphi_0 \wedge \cdots \wedge \varphi_{n-1}).$$

If $\mathfrak{M} \in \mathbf{Str}_{\mathscr{S}}$, and A is a subset of M, then, slightly generalizing the notation introduced on page 133, we denote by

$$\mathfrak{M}_A$$

the structure $\mathfrak{M}$, expanded in the obvious way to the signature $\mathscr{S}(A)$. An n-type Γ of $\mathscr{S}(A)$ is **consistent with $\mathfrak{M}$** if, for all finite subsets

$\{\varphi_0, \dots, \varphi_{m-1}\}$ of Γ,

$$\mathfrak{M} \vDash \exists x_0 \cdots \exists x_{n-1} \, (\varphi_0 \land \cdots \land \varphi_{n-1}),$$

that is,

$$(\exists x_0 \cdots \exists x_{n-1} \, (\varphi_0 \land \cdots \land \varphi_{n-1})) \in \mathrm{Th}(\mathfrak{M}_A).$$

In this case, by Compactness, if $\boldsymbol{c}$ is an n-tuple $(c_i \colon i < n)$ of new constants, then there is a model $\mathfrak{N}$ of $\mathrm{Th}(\mathfrak{M}_M) \cup \{\varphi(\boldsymbol{c}) \colon \varphi \in \Gamma\}$. Then we may assume $\mathfrak{M}_M \subseteq \mathfrak{N} \upharpoonright \mathscr{S}(M)$, and then

$$\mathfrak{M} \preccurlyeq \mathfrak{N} \upharpoonright \mathscr{S}.$$

We say Γ is **realized** in $\mathfrak{N} \upharpoonright \mathscr{S}$ by $({c_0}^{\mathfrak{N}}, \cdots, {c_{n-1}}^{\mathfrak{N}})$.

If $\mathfrak{M}$ is considered as fixed, we may denote by

$$\mathrm{S}_n(A)$$

the set of all complete n-types of $\mathscr{S}(A)$ that are consistent with $\mathfrak{M}$. The elements of A are the **parameters** of elements of $\mathrm{S}_n(A)$.

For every infinite cardinal κ, a structure is called **κ-saturated** if it realizes every type that is consistent with it and that has fewer than κ-many parameters. In particular, a structure is **ω_1-saturated** or **$\aleph_1$-saturated** if it realizes all types in countably many parameters.

Theorem 178. *For every structure $\mathfrak{A}$ with a countable signature, every non-principal ultrapower $\mathfrak{A}^\omega/P$ of $\mathfrak{A}$ is ω_1-saturated.*

Proof. If Φ is a type in countably many parameters, then Φ itself is countable, so we can write it as $\{\varphi_n \colon n \in \omega\}$. Let $\boldsymbol{a}_n$ satisfy $\varphi_0 \land \cdots \land \varphi_n$ in $\mathfrak{A}$. Then

$$k \leqslant n \implies \mathfrak{A} \vDash \varphi_k(\boldsymbol{a}_n).$$

Therefore, if P is a non-principal prime ideal of $\mathscr{P}(\omega)$, then $(\boldsymbol{a}_n \colon n \in \omega)/P$ realizes Φ in $\mathfrak{A}^\omega/P$. □

There is a version [8, Thm 6.1.1, p. 384] of the foregoing for uncountable index-sets (or exponents) Ω; but then P must have a countable subset whose union is Ω (so one should show that such prime ideals can be found).

7.7. A countable non-standard model of arithmetic

By **arithmetic** we mean the theory of $(\omega,+,\cdot)$ or of $(\omega,+,\cdot,0,1,\leqslant)$; it makes little difference, since

1) $\leqslant$ is definable in $(\omega,+,\cdot)$ by the formula $\exists z\ x+z=y$,
2) $\{0\}$ is definable by $\forall y\ y+x=y$,
3) $\{1\}$ is definable by $0<x\land\forall y\ (0=y\lor x\leqslant y)$.

Similarly $\{n\}$ is definable in $(\omega,+,\cdot)$ for all n in ω.

Every ultrapower of $(\omega,+,\cdot)$ is a model of arithmetic. Every *non-principal* ultrapower $\mathfrak{B}$ (determined by a non-principal ultrafilter F on ω) is a *non-standard* model of arithmetic, in the sense that it is not isomorphic to $(\omega,+,\cdot)$, but contains an infinite element c. However, $\mathfrak{B}$ here must be uncountable by Theorem 162. As we noted before this theorem, by the Downward Löwenheim–Skolem–Tarski Theorem (Theorem 153), we can obtain a countable elementary substructure $\mathfrak{A}$ of $\mathfrak{B}$ that includes $\omega\cup\{c\}$, and then $\mathfrak{A}$ will be an elementary extension of $(\omega,+,\cdot)$.

We can construct such a structure $\mathfrak{A}$ more directly as follows. Let A be the set of 0-*definable* singulary operations of $(\omega,+,\cdot)$. This means $f\in A$ if and only if the relation $\{(x,f(x))\colon x\in\omega\}$ is 0-definable (that is, definable without parameters). We can consider A as a subset of ω^ω. Then a constant sequence $(x,x,x,\dots)$ should be understood as the constant function $\{(n,x)\colon n\in\omega\}$ or $n\mapsto x$, which is in A. Thus the diagonal map embeds ω in A. Also A is closed under $+$ and $\cdot$. Therefore A is the universe of a substructure $\mathfrak{A}$ of $\mathfrak{B}$. Also, if $n\in\omega$, and φ is an $(n+1)$-ary formula, and $\boldsymbol{f}$ is an element $(f^0,\dots,f^{n-1})$ of A^n, then A has an element g such that for all i in ω,

$$(\omega,+,\cdot)\vDash\exists y\ \varphi(\boldsymbol{f}(i),y)\iff(\omega,+,\cdot)\vDash\varphi(\boldsymbol{f}(i),g(i)).$$

Indeed, g can be such that $g(i)$ is the *least* b such that $(\omega,+,\cdot)\vDash\varphi(\boldsymbol{f}(i),b)$, if such b exist; and otherwise $g(i)=0$. Then g is defined by the formula

$$(\varphi(\boldsymbol{f}(x),y)\land(\forall z\ (\varphi(\boldsymbol{f}(x),z)\Rightarrow y\leqslant z)))\lor(\forall z\ \neg\varphi(\boldsymbol{f}(x),z)\land y=0).$$

It follows by the Tarski–Vaught Test (page 159) that

$$\mathfrak{A} \preccurlyeq \mathfrak{B};$$

therefore, since $(\omega, +, \cdot) \subseteq \mathfrak{A}$, we have

$$(\omega, +, \cdot) \prec \mathfrak{A}.$$

Indeed, we now have that the following are equivalent:

$$\begin{gathered}
\mathfrak{B} \vDash \exists y\ \varphi(\boldsymbol{f}, y), \\
\{i \colon (\omega, +, \cdot) \vDash \exists y\ \varphi(\boldsymbol{f}(i), y)\} \in F, \\
\{i \colon (\omega, +, \cdot) \vDash \varphi(\boldsymbol{f}(i), g(i))\} \in F, \\
\mathfrak{B} \vDash \varphi(\boldsymbol{f}, g).
\end{gathered}$$

Now the Tarski–Vaught Test applies. This construction of $\mathfrak{A}$ is apparently due to Skolem.[3]

[3] I take it from Bell and Slomson [4, Ch. 12, §2].

8. Completeness of proof systems

Recall from page 152 that a sentence true in all structures of its signature is called *valid.* It is easy in principle to show that a sentence is *not* valid: just exhibit a model of its negation. But if a sentence *is* valid, how can we show this? We cannot simply verify the sentence in each structure of its signature, since there will be infinitely many of these structures, and even to verify a universal sentence $\forall x\ \varphi$ in *one* infinite structure requires checking infinitely many individual cases.

The method of *formal proof* is a way to establish the validity of sentences.

We shall develop a *proof-system* in which every provable sentence is valid. That is the easy part. The harder part is to show that, if a sentence is not provable in our system, then its negation has a model. Equivalently, every validity will have a formal proof in the proof-system. This result is *Gödel's Completeness Theorem.* A model of the negation of an unprovable sentence can be obtained as an ultraproduct, which is why we consider the whole subject here.

Recall again from page 152 that a sentence with no models is a *contradiction.* In our proof-system, there will be a notion of proving some sentences with the use of other sentences as hypotheses. A set of sentences from which a contradiction cannot be proved is **consistent** (with respect to the proof-system). Gödel's methods generalize to show that, at least in countable signatures, every consistent set of sentences has a model. The Compactness Theorem for countable signatures is a corollary of this result. Indeed, if a set Γ of sentences has no model, then by Gödel's Completeness Theorem a contradiction can be proved from Γ. But proofs are finite, and so a contradiction can be proved from a finite subset Γ_0 of Γ, and therefore Γ_0 has no model.

8.1. Formal proofs

It will be convenient to work, not only with sentences, but with arbitrary formulas. A **formal proof** is just a (finite) list of formulas such that each formula on the list is either

1) an *axiom,* or
2) derivable from formulas earlier in the list by means of a *rule of inference.*

We choose the axioms and rules of inference to serve our needs; taken all together, they constitute a **proof-system.** In a formal proof in such a system, the last formula is then said to be **provable** in the system, or to be a **theorem** of the system. Note that in fact *every* formula in a formal proof is provable, because every initial segment of a formal proof is still a formal proof.

A proof-system is **sound** if each of its theorems that is a sentence is valid; **complete,** if each validity is a theorem. In his doctoral dissertation of 1930, Gödel [24] defined a sound proof-system, obtained from the *Principia Mathematica* [61] of Russell and Whitehead, and showed that it was complete.

In formulas as defined on page 136, the logical symbols that can appear are $=$, $\neg$, $\vee$, $\wedge$, $\Rightarrow$, $\Leftrightarrow$, $\exists$, $\forall$, variables, and parentheses. (The other symbols come from the signature being used.) In fact we do not need $\wedge$, $\Rightarrow$, $\Leftrightarrow$ and $\exists$, but can understand them as abbreviations:

$$\begin{aligned}\varphi \wedge \psi &\text{ for } \neg(\neg\varphi \vee \neg\psi),\\ \varphi \Rightarrow \psi &\text{ for } \neg\varphi \vee \psi,\\ \varphi \Leftrightarrow \psi &\text{ for } \neg(\neg\varphi \vee \neg\psi) \vee \neg(\varphi \vee \psi),\\ \exists x\ \varphi &\text{ for } \neg\forall x\ \neg\varphi.\end{aligned}$$

The first four of Gödel's axioms, or rather *schemes* of axioms, are found on page 13, Chapter 1, of the *Principia Mathematica.* Recall that, by our convention on symbolic precedence given on page 18, $\vee$ takes precedence over $\Rightarrow$, and of two instances of $\Rightarrow$, the one on the right takes precedence.[1] By this convention then, the four axiom schemes are as follows.

1) $\varphi \vee \varphi \Rightarrow \varphi$,
2) $\varphi \Rightarrow \varphi \vee \psi$,
3) $\varphi \vee \psi \Rightarrow \psi \vee \varphi$,
4) $(\varphi \Rightarrow \psi) \Rightarrow \chi \vee \varphi \Rightarrow \chi \vee \psi$.

The remaining axiom schemes involve variables explicitly. Given a formula φ and variables x and y, we use the expression

$$\varphi^x_y$$

to denote the result of replacing every free occurrence of x in φ with y. We say that y is **substitutable** for x in φ if there is no subformula $\forall y\ \psi$ of φ in which there is an occurrence of x that is free as an occurrence in φ. For example, suppose φ is $\forall y\ (x \neq y \Rightarrow x = y)$, which is false when $x \neq y$. Then φ^x_y is $\forall y\ (y \neq y \Rightarrow y = y)$, which is valid; but y is not substitutable for x in φ.

Two of Gödel's remaining axioms are found in Chapter 9 of the *Principia Mathematica* (at $*9.2$ and $*9.25$, pp. 138–40).[2]

5) $\forall x\ \varphi \Rightarrow \varphi$.[3]
6) $\forall x\ (\vartheta \vee \varphi) \Rightarrow \vartheta \vee \forall x\ \varphi$, if x does not occur freely on ϑ.

[1] For Russell and Whitehead, the *primitive* Boolean connectives are $\vee$ and $\neg$; the expression $\varphi \Rightarrow \psi$ can then be understood as an abbreviation of $\neg\varphi \vee \psi$. As Gödel notes, after the first four axioms, there was a fifth, namely $\varphi \vee (\psi \vee \chi) \Rightarrow \psi \vee (\varphi \vee \chi)$, but Bernays showed it to be redundant. For us, each of the four axioms represents infinitely many axioms, since φ, ψ, and χ can be any formulas. It should be noted that Russell and Whitehead were involved in *creating* formal logic; in their time, our way of understanding formulas was not yet fully developed. For an amusing fictionalized account of Russell's interactions with Gödel, see *Logicomix* [15].

[2] Gödel's own reference is to the *Principia Mathematica's* Chapter 10, where the axioms are repeated, at $*10.1$ and $*10.12$, pp. 145–6. Gödel's six axioms used *propositional variables* where I put φ, ψ, and χ, and they used a *functional variable* where I put ϑ. Then in addition to the rules of inference given below, there was a rule allowing propositional and functional variables to be replaced by *formulas* in our sense.

[3] Gödel gives a stronger form: $\forall x\ \varphi \Rightarrow \varphi^x_y$, if y is substitutable for x in φ, but we do not need it: see Theorem 185.

Another axiom scheme involves a **change of bound variable:**

7) $\varphi \to \varphi'$, where φ has a subformula ψ in which a variable x does not occur freely, and there is a variable y not occurring in ψ at all, and φ' is the result of replacing each occurrence of x in ψ with y.

Equality is treated in two axiom schemes, found in Chapter 13 of the *Principia Mathematica* (at ∗13.15 and ∗13.101, pp. 177–8):

8) $x = x$,

9) $x = y \Rightarrow \varphi \Rightarrow \varphi^x_y$, if y is substitutable for x in φ.

The rules of inference are three:[4]

Detachment: From φ and $\varphi \Rightarrow \psi$ may be inferred ψ.[5]

Generalization: From ϑ may be inferred $\forall x\ \vartheta$.

Change of Free Variable: From φ may be inferred φ^x_y, provided y is substitutable for x in φ.

In the Rule of Change of Free Variable as stated above, y is **substituted for** x; in the Rule of Generalization, x is **generalized on.** A **generalization** of a formula φ is a formula is a sentence $\forall \boldsymbol{x}\ \varphi$ in which all free variables of φ are generalized on. Then we can generalize the notion of validity by saying that an arbitrary formula is **valid** if some (and hence every) generalization of it is true in every structure of its signature.

Theorem 179 (Soundness)**.** *Every provable formula is valid.*

[4] See the previous footnote on Gödel's additional rule of inference. Gödel apparently expressed the axiom of change of bound variable together with the rule of change of free variable as one rule, stated simply as, 'Individual variables (free or bound) may be replaced by others, so long as this does not cause overlapping of the scopes of variables denoted by the same sign' [24, p. 584]. Concerning all of his rules of inference, Gödel notes, 'Although Whitehead and Russell use these rules throughout their derivations, they do not formulate all of them explicitly.'

[5] Detachment is not Gödel's name for this rule; he (or more precisely his translator) calls it the Inferential Schema.

Proof. Induction. The axioms are valid, and the rules of inference preserve validity. □

We shall want to avoid writing down actual proofs, being content to recognize that they must exist, because of results like the following.

Theorem 180 (Detachment). *If φ and $\varphi \Rightarrow \psi$ are provable, then so is ψ.*

Proof. If $\chi_0, \dots, \chi_{n-1}, \varphi$ and $\chi_n, \dots, \chi_{n+m-1}, \varphi \Rightarrow \psi$ are proofs, then so is

$$\chi_0, \dots, \chi_{n-1}, \chi_n, \dots, \chi_{n+m-1}, \varphi, \varphi \Rightarrow \psi, \psi. \qquad \square$$

8.2. Propositional logic

A completeness theorem for *propositional logic* was already known before Gödel's completeness theorem.[6] **Propositional formulas** are, strictly, not formulas as defined in §5.1.3 (page 136) above; but they can be understood as formulas in which:

1) the place of atomic formulas is taken by **propositional variables;**
2) no quantification symbol ∃ or ∀ is used.

That is, for us, since we treat ∧, ⇒, and ⇔ as abbreviations,

1) every propositional variable is a formula;
2) if F is a propositional formula, so is $\neg F$;
3) if F and G are propositional formulas, so is $(F \lor G)$.

There are no *individual* variables in a propositional formula, but only *propositional* variables. A *structure* for propositional logic assigns a truth-value to each of these propositional variables. Then a propositional formula is true or false in the structure, according to the relevant parts of the definition of truth of sentences (on page 138), which we can express symbolically now as:

$$\mathfrak{A} \models \neg\sigma \iff \mathfrak{A} \not\models \sigma,$$

[6] Gödel's reference for this is Bernays from 1926; but the theorem can be found in Post's 1921 article [49].

$$\mathfrak{A} \vDash \sigma \lor \tau \iff \mathfrak{A} \vDash \sigma \text{ OR } \mathfrak{A} \vDash \tau.$$

We may treat the truth-value *true* as 1, and *false* as 0. Then a propositional formula F in an n-tuple $(P_0, \dots, P_{n-1})$ of propositional variables determines an n-ary operation $\hat{F}$ on 2^n, where if $\boldsymbol{e} \in 2^n$, then $\hat{F}(\boldsymbol{e})$ is the truth-value of F in any propositional structure that assigns the value e_i to P_i when $i < n$. This operation $\hat{F}$ can be described completely in a *truth-table.*

Theorem 181 (Propositional Completeness). *The first four axiom-schemes above (page 199), along with the inference-rule of Detachment, constitute a (sound and) complete proof-system for propositional logic.*

We are not going to prove this, since we already have an algorithm for determining whether a formula is a propositional validity: just write out its truth-table.[7]

Let us for the moment refer to atomic formulas and generalizations as **elementary formulas** [54, p. 26]. In a formula, if every elementary subformula is replaced with a propositional variable, the result is a propositional formula. Then we may refer to one formula φ as a **tautological consequence** of a finite set Γ of formulas if, when all of these formulas are converted to propositional formulas, φ becomes true in every structure in which the formulas of Γ become true. A formula is a **tautology,** simply, if it is a tautological consequence of the empty set of formulas.

Theorem 182 (Tautology). *If every formula in a finite set Γ of formulas is a theorem, and φ is a tautological consequence of Γ, then φ is a theorem.*

Proof. Write Γ as $\{\psi_0, \dots, \psi_{n-1}\}$. Under the hypothesis, the formula

$$\psi_0 \Rightarrow \dots \Rightarrow \psi_{n-1} \Rightarrow \varphi$$

[7] This is not a practical algorithm for long formulas; it may be more efficient to check a proposed formal proof of a formula than to write out the truth table of the formula. On the other hand, we have no algorithm for finding formal proofs. Then again, the proof of the completeness theorem would supply an algorithm.

is a propositional validity, so it is a theorem. By the Detachment Theorem, φ must be a theorem. □

Theorem 183. *The formula*

$$\forall x\,(\varphi \Rightarrow \psi) \Rightarrow \forall x\,\varphi \Rightarrow \forall x\,\psi$$

is always provable.

Proof. Suppose y does not occur in φ or ψ. Then the following formulas are provable:

$$\begin{array}{rl}
\forall x\,(\varphi \Rightarrow \psi) \Rightarrow \varphi \Rightarrow \psi, & \text{[Axiom 5]} \\
\forall x\,\varphi \Rightarrow \varphi, & \text{[Axiom 5]} \\
\forall x\,(\varphi \Rightarrow \psi) \Rightarrow \forall x\,\varphi \Rightarrow \psi, & \text{[Tautology Theorem]} \\
\forall x\,(\varphi \Rightarrow \psi) \Rightarrow \forall x\,\varphi \Rightarrow \psi^x_y, & \text{[Change of Free Variable]} \\
\forall y\,(\forall x\,(\varphi \Rightarrow \psi) \Rightarrow \forall x\,\varphi \Rightarrow \psi^x_y), & \text{[Generalization]} \\
\forall x\,(\varphi \Rightarrow \psi) \Rightarrow \forall y\,(\forall x\,\varphi \Rightarrow \psi^x_y), & \text{[Axiom 6]} \\
\forall x\,(\varphi \Rightarrow \psi) \Rightarrow \forall x\,\varphi \Rightarrow \forall y\,\psi^x_y, & \text{[Axiom 6]} \\
\forall x\,(\varphi \Rightarrow \psi) \Rightarrow \forall x\,\varphi \Rightarrow \forall x\,\psi. & \text{[Change of Bound Variable]}
\end{array}$$ □

Theorem 184. *If x_0, …, x_{n-1} are not free in ϑ, then the formula*

$$\forall x_0\,\cdots\,\forall x_{n-1}\,(\vartheta \lor \varphi) \Rightarrow \vartheta \lor \forall x_0\,\cdots\,\forall x_{n-1}\,\varphi$$

is provable.

Theorem 185. *If each variable y_i is substitutable for x_i in φ, then the formulas*

$$\begin{gathered}
\forall x_0\,\ldots\,\forall x_{n-1}\,\varphi \Rightarrow \varphi^{x_0\cdots x_{n-1}}_{y_0\cdots y_{n-1}}, \\
\varphi^{x_0\cdots x_{n-1}}_{y_0\cdots y_{n-1}} \Rightarrow \exists x_0\,\ldots\,\exists x_{n-1}\,\varphi
\end{gathered}$$

are provable.

Proof. 1. The following are instances of Axiom 5:

$$\begin{gathered}\forall x_0 \ \dots \ \forall x_{n-1}\, \varphi \Rightarrow \forall x_1 \ \dots \ \forall x_{n-1}\, \varphi,\\ \forall x_1 \ \dots \ \forall x_{n-1}\, \varphi \Rightarrow \forall x_2 \ \dots \ \forall x_{n-1}\, \varphi,\\ \dots\dots\dots\dots\dots\dots,\\ \forall x_{n-2}\, \forall x_{n-1}\, \varphi \Rightarrow \forall x_{n-1}\, \varphi,\\ \forall x_0\, \varphi \Rightarrow \varphi.\end{gathered}$$

Then $\forall x_0 \ \dots \ \forall x_{n-1}\, \varphi \Rightarrow \varphi$ is provable by the Tautology Theorem. Since no x_i is free in the subformula $\forall x_0 \ \dots \ \forall x_{n-1}\, \varphi$, by the Rule of Change of Free Variable we can now prove

$$(\forall x_0 \ \dots \ \forall x_{n-1}\, \varphi \Rightarrow \varphi)^{x_0 \cdots x_{n-1}}_{y_0 \cdots y_{n-1}},$$

which is $\forall x_0 \ \dots \ \forall x_{n-1}\, \varphi \Rightarrow \varphi^{x_0 \cdots x_{n-1}}_{y_0 \cdots y_{n-1}}$, as desired.

2. From Axiom 5, using the tautology $(\psi \Rightarrow \neg\chi) \Rightarrow \chi \Rightarrow \neg\psi$ in the form

$$(\forall x\ \neg\varphi \Rightarrow \neg\varphi) \Rightarrow \varphi \Rightarrow \exists x\ \varphi$$

(that is, $\forall x\ \neg\varphi \Rightarrow \neg\varphi) \Rightarrow \varphi \Rightarrow \neg\forall x\ \neg\varphi)$, we obtain the axiom

$$\varphi \Rightarrow \exists x\ \varphi.$$

Now an argument like the previous one yields the claim. □

8.3. Sequents

If φ is a formula, then a formal proof **from** φ as a **hypothesis** is a formal proof in the earlier sense, except

- φ, like an axiom, may be introduced into the proof, but
- *no free variable of φ may be substituted for or generalized on.*

If ψ is provable from φ in this sense, we may write

$$\varphi \vdash \psi. \tag{8.1}$$

If φ is provable, simply, then we may express this by

$$\vdash \varphi.$$

The restriction on the use of Generalization and Change of Variables ensures that the following is true.

Theorem 186. *If $\varphi \vdash \psi$, then the formula $\varphi \Rightarrow \psi$ is valid.*

Proof. Induction on ψ. □

We may call an expression as in (8.1) a **sequent.** It will be useful to note the following, so that we can work with sequents rather than formal proofs themselves.

Theorem 187.

1. *If $\vdash \varphi$ and $\varphi \vdash \psi$, then $\vdash \psi$.*
2. *If $\chi \vdash \varphi$ and $\varphi \vdash \psi$, then $\chi \vdash \psi$.*

Proof. The first claim is easily obtained by concatenating two formal proofs. Thus, if

$$\vartheta_0, \dots, \vartheta_{n-1}, \varphi$$

is a formal proof of φ, and

$$\vartheta_n, \dots, \vartheta_{n+m-1}, \psi$$

is a formal proof of ψ from φ, then

$$\vartheta_0, \dots, \vartheta_{n-1}, \vartheta_n, \dots, \vartheta_{n+m-1}, \psi$$

is a formal proof of ψ.

For the second claim, a similar concatenation may not be a formal proof from χ, if χ has free variables that are not free in φ. For, in this case, a particular proof of ψ from φ might have involved substitution for, or generalization on, some of these variables. But suppose

$$\vartheta_0, \dots, \vartheta_{n-1}, \varphi$$

is a formal proof of φ from χ, and

$$\vartheta_n, \dots, \vartheta_{n+m-1}, \psi$$

is a formal proof of ψ from φ, and the free variables of χ that are not free in φ are $x_0, \dots, x_{n-1}$. Let $y_0, \dots, y_{n-1}$ be distinct variables not appearing at all in any of the formulas in the two formal proofs above, and then, if $k < m$, let

$$\vartheta_{n+k}{}' \text{ be } (\vartheta_{n+k})^{x_0 \cdots x_{n-1}}_{y_0 \cdots y_{n-1}}, \qquad \psi_k \text{ be } \psi^{x_0 \cdots x_{n-1-k}}_{y_0 \cdots y_{n-1-k}}.$$

The sequence

$$\vartheta_n{}', \dots, \vartheta_{n+m-1}{}', \psi_0$$

is a formal proof of ψ_0 from φ in which none of the x_i appear. In particular, none of these variables is substituted for or generalized on in the proof. Therefore

$$\vartheta_0, \dots, \vartheta_{n-1}, \vartheta_n{}', \dots, \vartheta_{n+m-1}{}', \psi_0$$

is a formal proof of ψ_0 from χ. Finally

$$\vartheta_0, \dots, \vartheta_{n-1}, \vartheta_n{}', \dots, \vartheta_{n+m-1}{}', \psi_0, \dots, \psi_{m-1}, \psi$$

is a formal proof of ψ from χ. □

8.4. Completeness by ultraproducts

Suppose σ is an arbitrary sentence. We want to show that either σ is provable, or else its negation has a model.[8] In fact this model will be countable. For now, we make several simplifying assumptions:

1. For some positive integers p and q, for some p-tuple $\boldsymbol{x}$ and q-tuple $\boldsymbol{y}$ of variables, all distinct from one another, for some quantifier-free formula φ,

$$\sigma \quad \text{is} \quad \exists \boldsymbol{x}\, \forall \boldsymbol{y}\, \varphi.$$

[8] The ensuing argument is based mainly on that of Bell and Slomson [4, Ch. 12, §1]. These writers cite J.N. Crossley for the suggestion of introducing ultraproducts to Gödel's original argument. Church [10, §44] explicates Gödel's original argument more faithfully.

2. No operation symbols occur in φ.

3. The sign $=$ of equality does not occur in φ.

The justification of these assumptions does not involve ultraproducts, so it is relegated to §8.6, page 212.

We may assume that all variables come from a countable set V, and that there is a bijection $k \mapsto v_k$ from ω onto V. The power V^p being also countable, we may suppose we have a bijection

$$k \mapsto \boldsymbol{x}_k$$

from ω onto V^p. Then there is an injection

$$k \mapsto \boldsymbol{y}_k$$

from ω into V^q such that $\boldsymbol{y}_k$ has no entries in common with $\boldsymbol{x}_0 \cdots \boldsymbol{x}_k$. We now denote

$$\begin{array}{rcl} \varphi^{\boldsymbol{x}\boldsymbol{y}}_{\boldsymbol{x}_k\boldsymbol{y}_k} & \text{by} & \varphi_k, \\ \varphi_0 \lor \cdots \lor \varphi_k & \text{by} & \vartheta_k, \\ \forall \boldsymbol{x}_k\, \forall \boldsymbol{y}_k\ \cdots\ \forall \boldsymbol{x}_0\, \forall \boldsymbol{y}_0\, \vartheta_k & \text{by} & \tau_k. \end{array}$$

That is, τ_k is a generalization of ϑ_k, and ϑ_k itself is defined recursively in k, thus:

$$\vartheta_0 \text{ is } \varphi_0, \qquad\qquad \vartheta_{k+1} \text{ is } \vartheta_k \lor \varphi_{k+1}.$$

Lemma 20. *In the notation above, for all k in ω, the sentence*

$$\tau_k \Rightarrow \sigma$$

is provable.

Proof. We shall use induction. First, the following are provable:

$$\begin{array}{ll} \tau_0 \Rightarrow \vartheta_0, & \text{[Theorem 185]} \\ \tau_0 \Rightarrow \varphi_0, & [\vartheta_0 \text{ is } \varphi_0] \\ \tau_0 \Rightarrow \varphi, & \text{[Change of Free Variable]} \end{array}$$

$$\begin{aligned}
&\tau_0 \Rightarrow \exists \boldsymbol{y}\ \varphi, && \text{[Theorem 185]}\\
&\tau_0 \Rightarrow \forall \boldsymbol{x}\ \exists \boldsymbol{y}\ \varphi, && \text{[Generalization]}\\
&\tau_0 \Rightarrow \sigma. && [\sigma \text{ is } \forall \boldsymbol{x}\ \exists \boldsymbol{y}\ \varphi]
\end{aligned}$$

For the inductive step, we note first that since no entry of $\boldsymbol{y}_{k+1}$ appears in ϑ_k, we have the theorems

$$\begin{aligned}
&\tau_{k+1} \Rightarrow \forall \boldsymbol{y}_{k+1}\ \vartheta_{k+1}, && \text{[Theorem 185]}\\
&\tau_{k+1} \Rightarrow \forall \boldsymbol{y}_{k+1}\ (\vartheta_k \lor \varphi_{k+1}), && [\vartheta_{k+1} \text{ is } \vartheta_k \lor \varphi_{k+1}]\\
&\tau_{k+1} \Rightarrow \vartheta_k \lor \forall \boldsymbol{y}_{k+1}\ \varphi_{k+1}, && \text{[Theorem 184]}\\
&\tau_{k+1} \Rightarrow \vartheta_k \lor \exists \boldsymbol{x}_{k+1}\ \forall \boldsymbol{y}_{k+1}\ \varphi_{k+1}, && \text{[Theorem 185]}\\
&\tau_{k+1} \Rightarrow \vartheta_k \lor \exists \boldsymbol{x}\ \forall \boldsymbol{y}\ \varphi, && \text{[Change of Bound Variable]}\\
&\tau_{k+1} \Rightarrow \vartheta_k \lor \sigma, && [\sigma \text{ is } \exists \boldsymbol{x}\ \forall \boldsymbol{y}\ \varphi]\\
&\tau_{k+1} \Rightarrow \tau_k \lor \sigma. && \text{[Theorem 184]}
\end{aligned}$$

Thus if $\tau_k \Rightarrow \sigma$ is a theorem, then so is $\tau_{k+1} \Rightarrow \sigma$. This completes the induction. □

Theorem 188 (Completeness)**.** *In the notation above, if σ is not provable, then $\neg\sigma$ has a model.*

Proof. If some τ_k is provable, then so is σ itself, by the lemma, and we are done. So suppose that no τ_k is provable. Then no ϑ_k is provable; so it must not be a tautology. Since ϑ_k is also quantifier-free, there must be a truth-assignment on the set of its atomic subformulas that makes ϑ_k false. We can extend this to a truth-assignment F_k on the set of *all* atomic formulas in variables from V with predicates occurring in σ. Now, for every k in ω, we can understand V as the universe of a structure $\mathfrak{A}_k$ such that, for each n in ω, for each n-ary predicate R occurring in σ,

$$R^{\mathfrak{A}_k} = \{\boldsymbol{u} \in V^n \colon F_k(R\boldsymbol{u}) = 1\}.$$

Here we rely on the assumption that none of the predicates R is $=$. We now have

$$\mathfrak{A}_k \vDash R\boldsymbol{u} \iff F_k(R\boldsymbol{u}) = 1.$$

Then by construction

$$\mathfrak{A}_k \models \neg\vartheta_k.$$

If $k \leqslant \ell$, then, since $\vartheta_k \Rightarrow \vartheta_\ell$ and hence $\neg\vartheta_\ell \Rightarrow \neg\vartheta_k$ are theorems, we have

$$\mathfrak{A}_\ell \models \neg\vartheta_k.$$

Thus for all k in ω,

$$\{j \in \omega \colon \mathfrak{A}_j \models \vartheta_k\} \subseteq k. \tag{8.2}$$

Now let $\mathfrak{C}$ be a non-principal ultraproduct of the structures $\mathfrak{A}_k$. Since each of these structures has the same universe, namely V, each u in V can be interpreted in $\mathfrak{C}$ as its image $(u \colon k \in \omega)$ under the diagonal map. Then for all k in ω,

$$\mathfrak{C} \models \neg\vartheta_k,$$

and so

$$\mathfrak{C} \models \neg\varphi_k.$$

Since we have no operation symbols in our signature, every subset of C is the universe of a substructure of $\mathfrak{C}$. Let B be the image of V under the diagonal map in C. Since φ is quantifier-free, and the interpretations of all of the variables are now in B, we now have

$$\mathfrak{B} \models \neg\varphi_k,$$

and so, treating $\boldsymbol{x}$ and $\boldsymbol{y}$ now as tuples of variables again, we have

$$\mathfrak{B} \models \exists\boldsymbol{y}\ \neg\varphi^{\boldsymbol{x}}_{\boldsymbol{x}_k}.$$

Since every element of B^p is the interpretation of some $\boldsymbol{x}_j$, we conclude

$$\mathfrak{B} \models \forall\boldsymbol{x}\ \exists\boldsymbol{y}\ \neg\varphi,$$

that is, σ is false in $\mathfrak{B}$. □

8.5. Completeness by König's Lemma

In his proof of the Completeness Theorem, Gödel himself does not use an ultraproduct explicitly in his argument, but from the structures $\mathfrak{A}_k$, he can be understood to create the structure $\mathfrak{B}$ as follows.[9] Let $(\alpha_k \colon k \in \omega)$ be a list of all of the atomic formulas appearing in the formulas φ_ℓ. The universe B of $\mathfrak{B}$ will be the set V of variables occurring in these formulas. We define $\mathfrak{B}$ by determining in each case whether α_k, considered as a sentence, is to be true in $\mathfrak{B}$. This determination can be made recursively as follows.

For an arbitrary structure $\mathfrak{A}$ and sentence σ of its signature, the interpretation $\sigma^{\mathfrak{A}}$ of σ in $\mathfrak{A}$ is, formally, a subset of A^0, namely the subset

$$\{x \in A^0 \colon \mathfrak{A} \vDash \sigma\}.$$

But A^0 has a unique element, which is $\varnothing$, also called 0. Thus A^0 itself is $\{0\}$, which is 1, and $\mathscr{P}(A^0) = \{0, 1\}$, which is 2. So $\sigma^{\mathfrak{A}}$ is an element of 2, and

$$\mathfrak{A} \vDash \sigma \iff \sigma^{\mathfrak{A}} = 1.$$

Suppose for some n in ω an element $(e_k \colon k < n)$ of 2^n has been chosen such that the set

$$\left\{ i \in \omega \colon \bigwedge_{k<n} {\alpha_k}^{\mathfrak{A}_i} = e_k \right\} \tag{8.3}$$

is infinite. This set is the union of the two sets of the form

$$\left\{ i \in \omega \colon \bigwedge_{k<n} {\alpha_k}^{\mathfrak{A}_i} = e_k \mathbin{\&} {\alpha_n}^{\mathfrak{A}_i} = e \right\}, \tag{8.4}$$

where $e \in 2$. Hence at least one of these sets is infinite. If it is infinite when $e = 0$, we let $e_n = 0$. Otherwise the set must be infinite when $e = 1$, so we let $e_n = 1$. By recursion, we obtain an element $(e_n \colon n \in \omega)$ of 2^ω. Now we define

$$\mathfrak{B} \vDash \alpha_n \iff e_n = 1.$$

[9] I am guided by Church's version of Gödel's argument here. See below.

It follows by induction that, for each n in ω,

$$\left|\left\{i\colon \bigwedge_{k<n} {\alpha_k}^{\mathfrak{A}_i} = {\alpha_k}^{\mathfrak{B}}\right\}\right| = \omega. \tag{8.5}$$

The construction ensures $\mathfrak{B} \vDash \neg\vartheta_j$ as before. Indeed, suppose $\mathfrak{B} \vDash \vartheta$ for some formula ϑ (interpreted in $\mathfrak{B}$ as a sentence). Then the atomic subformulas of ϑ belong to a finite set $\{\alpha_i\colon i < k\}$, so

$$\left\{i\colon \bigwedge_{k<n} {\alpha_k}^{\mathfrak{A}_i} = {\alpha_k}^{\mathfrak{B}}\right\} \subseteq \{i \in \omega\colon \mathfrak{A}_i \vDash \vartheta\}.$$

In particular, by (8.5), the set $\{i \in \omega\colon \mathfrak{A}_i \vDash \vartheta\}$ must be infinite. However, as in (8.2) we have also $\{i \in \omega\colon \mathfrak{A}_i \vDash \vartheta_j\} \subseteq j$, and in particular the set $\{i \in \omega\colon \mathfrak{A}_i \vDash \vartheta_j\}$ is finite. Thus $\mathfrak{B} \nvDash \vartheta_j$.

There is some arbitrariness in our definition of $\mathfrak{B}$. If both of the sets of the form in (8.4) are infinite, then e_n could be either element of 2; we arbitrarily let it be 1. Alternatively, if we had a nonprincipal ultrafilter D on ω, then we could just define

$$\mathfrak{B} \vDash \alpha_k \iff \{i\colon \mathfrak{A}_i \vDash \alpha_k\} \in D.$$

Thus we would return to the earlier ultraproduct construction. An advantage of our alternative construction is that the Axiom of Choice is not required.

Gödel himself is not explicit about how he obtains $\mathfrak{B}$. His editor van Heijenoort detects an allusion to König's Lemma. There are more than one theorem called by this name, but probably what is meant is the next theorem below [38, Lemma II.5.7, p. 69].

A **tree** is a (partially) ordered set such that, for every a in the set, the subset $\{x\colon x < a\}$ is well-ordered. The ordinal that is isomorphic to this set is then the **height** of a. An element of the underlying set of the tree is a **node** of the tree. If the height of the node a is β, then a **successor** of a is a node b with height $\beta + 1$ such that $a < b$. A **branch** of the tree is a maximal linearly ordered set of nodes. The

height of the tree is the supremum of the heights of its nodes. The tree is an **ω-tree** if its every element has finite height and finitely many successors. Then an ω-tree has height at most ω.

Theorem 189 (König's Lemma). *Every infinite ω-tree has an infinite branch.*

Proof. By the Axiom of Choice, we may assume that the set of successors of every member of the tree is well-ordered. We select an infinite branch recursively by first letting a_0 be a node at height 0 such that $\{x\colon a_0 < x\}$ is infinite; then, assuming $\{x\colon a_k < x\}$ is infinite, we let a_{k+1} be the least successor of a_k such that $\{x\colon a_{k+1} < x\}$ is infinite. □

This theorem applies to the present situation as follows. We start with $2^{<\omega}$, that is, $\bigcup_{n\in\omega} 2^n$, ordered by inclusion, so that $\boldsymbol{a} \leqslant \boldsymbol{b}$ if and only if $\boldsymbol{a}$ is an initial segment of $\boldsymbol{b}$. In this way we obtain the **complete binary tree of height ω**. See Figure 8.1. This has a sub-tree T consisting of those $(e_0, \dots, e_{n-1})$ such that the set $\left\{i \in \omega\colon \bigwedge_{k<n} {\alpha_k}^{\mathfrak{A}_i} = e_k\right\}$ in (8.3) is infinite. This sub-tree T is infinite because, by induction, it has nodes at each finite height. Then König's Lemma applies, giving us an infinite branch of T; the union of this infinite branch is an element $(e_n\colon n \in \omega)$ of 2^ω giving us $\mathfrak{B}$ as before.

The general form of König's Lemma uses the Axiom of Choice; we do not need this here, since the successors of every node $(e_0, \dots, e_{n-1})$ of T are among $(e_0, \dots, e_{n-1}, 0)$ and $(e_0, \dots, e_{n-1}, 1)$, and the former can be understood to precede the latter.

The present situation is simpler in another way too, since every branch of T is infinite.

8.6. Arbitrary formulas

We have to justify the assumptions about σ made at the beginning of §8.4.

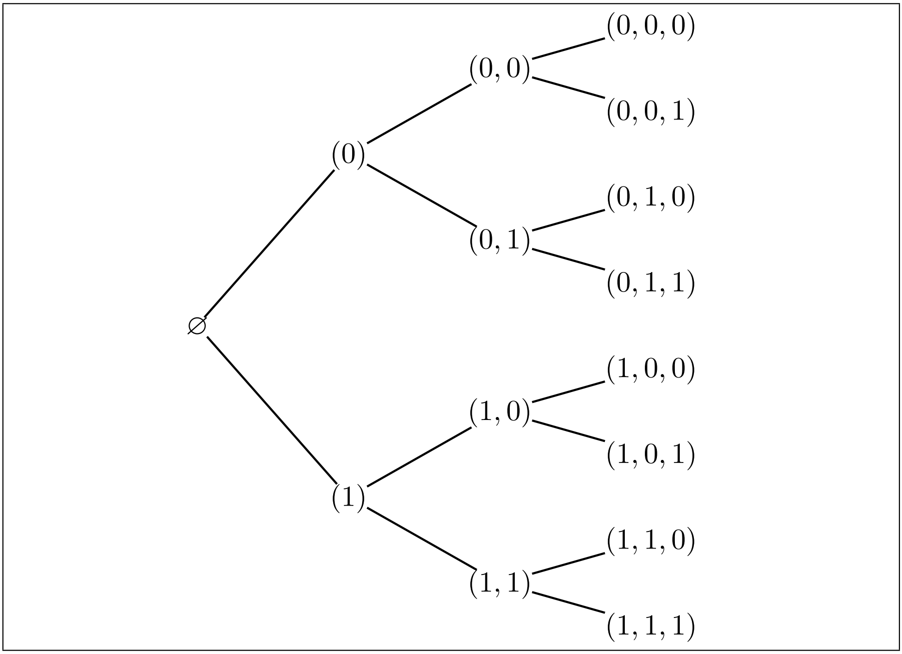

Figure 8.1.: The complete binary tree of height ω

8.6.1. Skolem normal form

Recall from page 18 that a *quantifier* is an expression $\forall x$ or $\exists x$ in a formula. These are *universal* and *existential* quantifiers, respectively. (We currently understand $\exists x$ as an abbreviation of $\neg\forall x\neg$). A formula is in **prenex normal form** if all of its quantifiers are at the front.

Theorem 190. *For every formula φ there is a formula $\hat{\varphi}$ in prenex normal form such that each of φ and $\hat{\varphi}$ is provable from the other.*

Proof. Suppose φ and ψ are formulas, and y is a variable not occurring freely in either of them, but substitutable for x in ψ. Then each of the formulas

$$\varphi \lor \forall x\ \psi, \qquad\qquad \forall y\ (\varphi \lor \psi^x_y)$$

is provable from the other. Indeed,

$$\begin{aligned}
&\vdash \forall x\ \psi \Rightarrow \psi^x_y, && \text{[Ax.]}\\
&\vdash (\forall x\ \psi \Rightarrow \psi^x_y) \Rightarrow \varphi \vee \forall x\ \psi \Rightarrow \varphi \vee \psi^x_y, && \text{[Taut.]}\\
&\vdash \varphi \vee \forall x\ \psi \Rightarrow \varphi \vee \psi^x_y, && \text{[Det.]}\\
\varphi \vee \forall x\ \psi &\vdash \varphi \vee \psi^x_y, && \text{[Det.]}\\
\varphi \vee \forall x\ \psi &\vdash \forall y\ (\varphi \vee \psi^x_y), && \text{[Gen.]}
\end{aligned}$$

and conversely

$$\begin{aligned}
&\vdash \forall y\ (\varphi \vee \psi^x_y) \Rightarrow \varphi \vee \forall y\ \psi^x_y, && \text{[Ax.]}\\
\forall y\ (\varphi \vee \psi^x_y) &\vdash \varphi \vee \forall y\ \psi^x_y, && \text{[Det.]}\\
\forall y\ (\varphi \vee \psi^x_y) &\vdash \varphi \vee \forall x\ \psi. && \text{[Ch. of Var.]}
\end{aligned}$$

Also each of

$$\varphi \wedge \forall x\ \psi, \qquad\qquad \forall y\ (\varphi \wedge \psi^x_y)$$

is provable from the other; for, the same proof that establishes $\forall y\ (\varphi \vee \psi^x_y) \vdash \varphi \vee \forall x\ \psi$ gives us, *mutatis mutandis,* $\forall y\ (\varphi \wedge \psi^x_y) \vdash \varphi \wedge \forall x\ \psi$, while

$$\begin{aligned}
&\vdash \forall y\ (\varphi \wedge \psi^x_y) \Rightarrow \varphi \wedge \psi^x_y, && \text{[Ax.]}\\
\forall y\ (\varphi \wedge \psi^x_y) &\vdash \varphi \wedge \psi^x_y, && \text{[Det.]}\\
&\vdash \varphi \wedge \psi^x_y \Rightarrow \varphi, && \text{[Taut.]}\\
\forall y\ (\varphi \wedge \psi^x_y) &\vdash \varphi, && \text{[Det.]}\\
&\vdash \varphi \wedge \psi^x_y \Rightarrow \psi^x_y, && \text{[Taut.]}\\
\forall y\ (\varphi \wedge \psi^x_y) &\vdash \psi^x_y, && \text{[Det.]}\\
\forall y\ (\varphi \wedge \psi^x_y) &\vdash \forall y\ \psi^x_y, && \text{[Gen.]}\\
\forall y\ (\varphi \wedge \psi^x_y) &\vdash \forall x\ \psi, && \text{[Ch. of Var.]}\\
&\vdash \varphi \Rightarrow \forall x\ \psi \Rightarrow (\varphi \wedge \forall x\ \psi), && \text{[Det.]}\\
\forall y\ (\varphi \wedge \psi^x_y) &\vdash \varphi \wedge \forall x\ \psi. && \text{[Det.]}
\end{aligned}$$

NOW WE NEED SOMETHING LIKE if $\varphi \vdash \psi$ then $\neg\psi \vdash \neg\varphi$, with appropriate restrictions. □

A *sentence* is in **Skolem normal form** if is in prenex normal form, and moreover, no existential quantifier follows a universal quantifier.

Theorem 191. *For every formula φ, there is a sentence σ in Skolem normal form, possibly with new predicates, such that*

- *if σ is valid, then so is φ,*
- *if $\neg\sigma$ has a model, then $\neg\varphi$ will be satisfied in that model (that is, it will define a nonempty subset of that model).*

Proof. A sentence in prenex normal form can be written as

$$\exists \boldsymbol{x}\ \forall y\ \mathsf{Q}\ \vartheta,$$

where Q is a string of quantifiers, and ϑ is quantifier-free. Introduce a new predicate R and form the sentence

$$\exists \boldsymbol{x}\ (\forall y\ (\mathsf{Q}\ \vartheta \Rightarrow R\boldsymbol{x}y) \Rightarrow \forall y\ R\boldsymbol{x}y).$$

This has the desired properties. CHECK!!!!!!!!!!!!!!!! It is also equivalent to

$$\begin{gathered}
\exists \boldsymbol{x}\ (\exists y\ (\mathsf{Q}\ \vartheta \land \neg R\boldsymbol{x}y) \lor \forall y\ R\boldsymbol{x}y),\\
\exists \boldsymbol{x}\ \exists y\ ((\mathsf{Q}\ \vartheta \land \neg R\boldsymbol{x}y) \lor \forall z\ R\boldsymbol{x}z),\\
\exists \boldsymbol{x}\ \exists y\ (\mathsf{Q}\ (\vartheta \land \neg R\boldsymbol{x}y) \lor \forall z\ R\boldsymbol{x}z),\\
\exists \boldsymbol{x}\ \exists y\ \mathsf{Q}\ ((\vartheta \land \neg R\boldsymbol{x}y) \lor \forall z\ R\boldsymbol{x}z),\\
\exists \boldsymbol{x}\ \exists y\ \mathsf{Q}\ \forall z((\vartheta \land \neg R\boldsymbol{x}y) \lor R\boldsymbol{x}z),
\end{gathered}$$

This last sentence is in prenex normal form, though perhaps not in Skolem normal form. Still, the number of universal quantifiers that precede existential quantifiers has decreased. So the process terminates in a sentence that must be in Skolem normal form. □

8.6.2. Operation symbols

What we call relations, Gödel calls functions; but he has no symbols for what we call operations. If we use such symbols, we can deal with them as follows. Suppose, for some n-ary operation symbol F, there is an atomic subformula α of σ featuring a term $Ft_0 \cdots t_{n-1}$. Introducing a new $(n+1)$-ary predicate R_F, we can replace the term $Ft_0 \cdots t_{n-1}$ in α with a new variable x, obtaining an atomic formula α'. We can then replace α in σ with the formula

$$\exists x\ (\alpha' \land R_F t_0 \cdots t_{n-1} x),$$

obtaining the formula σ'. Then σ is valid if and only if the formula

$$\sigma' \land \forall \boldsymbol{x}\ \exists y\ \forall z\ \bigl(R\boldsymbol{x}y \land (R\boldsymbol{x}z \Rightarrow y = z)\bigr)$$

is valid. Now we have to show that σ is provable from this last formula.

8.6.3. Equality

Suppose no operation symbol occurs in σ, but the sign $=$ of equality does occur. We have to deal with the requirement that this sign is interpreted in every structure as equality itself (and not merely an equivalence relation). We introduce a new binary predicate $\equiv$, and we replace each occurrence of $=$ in σ with this new predicate $\equiv$, obtaining a new sentence σ'. Now let $(R_0, \dots, R_m)$ be a list of all predicates (including $\equiv$) occurring in σ', and let σ'' be the sentence

$$\sigma' \land \forall \boldsymbol{x}\ \forall \boldsymbol{y}\ \Bigl(\boldsymbol{x} \equiv \boldsymbol{y} \Rightarrow \bigwedge_{j \leqslant m} (R_j \boldsymbol{x}_j \Rightarrow R_j \boldsymbol{y}_j)\Bigr).$$

(Here $\boldsymbol{x}_j$ and $\boldsymbol{y}_j$ are initial segments, of appropriate length, of $\boldsymbol{x}$ and $\boldsymbol{y}$ respectively; and $\boldsymbol{x}$ and $\boldsymbol{y}$ are long enough to make this possible.) Then σ is valid if and only if σ'' is valid. Also, if $\mathfrak{A} \models \sigma''$, then $\equiv^{\mathfrak{A}}$ is an equivalence relation on A, and the set of equivalence classes is the universe of a model of σ. Now we have to show that σ is provable from σ''.

9. Algebraic geometry

We shall assume the Axiom of Choice throughout this chapter. Also, K will be a field, and L will be a field of which K is a subfield, that is,

$$K \subseteq L.$$

In short, L/K will be a field-extension. For example, K might be $\mathbb{Q}$, and then L might be $\mathbb{C}$. For some n in ω, we shall let $\boldsymbol{X}$ denote an n-tuple $(X^0, \dots, X^{n-1})$ of indeterminates, so that we can form the ring $K[\boldsymbol{X}]$ of polynomials as on page 73. If $n = 1$, we write this field as $K[X]$; if $n = 2$, as $K[X, Y]$.

9.1. The spectrum of a polynomial ring

Given a signature $\mathscr{S}$, we have defined

- the class $\mathbf{Str}_{\mathscr{S}}$ of structures of $\mathscr{S}$ (page 46),
- the set $\operatorname{Sen}(\mathscr{S})$ of sentences of $\mathscr{S}$ (page 138), and
- the relation $\vDash$ between them (page 136).

We shall now consider analogously

- the set L^n of n-tuples of elements of L,
- the set $K[\boldsymbol{X}]$ of polynomials over K, and
- the relation $\{(\boldsymbol{x}, f) \in L^n \times K[\boldsymbol{X}] \colon f(\boldsymbol{x}) = 0\}$ between them.

In particular, we shall be interested in the Galois correspondence induced by this relation as in Theorem 105 (page 110). We shall write the polarities constituting the Galois correspondence as

$$A \mapsto \mathrm{I}_K(A), \qquad\qquad F \mapsto \mathrm{Z}_L(F),$$

respectively. As in the case of model theory, we may use variations of this notation, letting

$$\mathrm{I}_K(\boldsymbol{x}) = \{f \in K[\boldsymbol{X}] \colon f(\boldsymbol{x}) = 0\}, \quad \mathrm{Z}_L(f) = \{\boldsymbol{x} \in L^n \colon f(\boldsymbol{x}) = 0\},$$

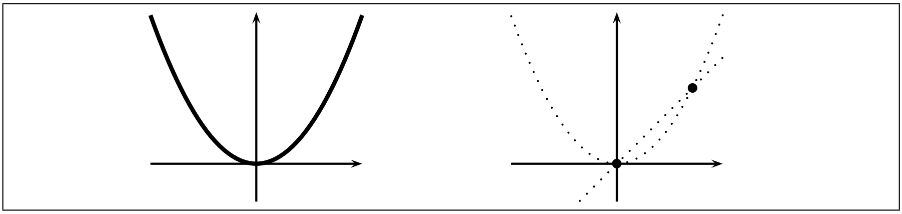

Figure 9.1.: The zero-loci of $Y - X^2$ and $\{Y - X^2, Y - X\}$ in $\mathbb{R}^2$

so that, analogously to (5.2) on page 139,

$$\mathrm{I}_K(A) = \bigcap_{\boldsymbol{x}\in A} \mathrm{I}_K(\boldsymbol{x}), \qquad \mathrm{Z}_L(F) = \bigcap_{f\in F} \mathrm{Z}_L(f).$$

The set $\mathrm{Z}_L(F)$ is the **zero-locus** of F in L^n: see Figure 9.1. The function $A \mapsto \mathrm{Z}_L(A)$ is the **zero-locus map.** A course in so-called analytic geometry is a study of zero-loci in $\mathbb{R}$, in case n is 2 or 3, so that $K[\boldsymbol{X}]$ can be written as $\mathbb{R}[X, Y]$ or $\mathbb{R}[X, Y, Z]$.

The zero-loci of the various subsets of $K[\boldsymbol{X}]$ are also called **algebraic sets.**[1] As the notation is supposed to recall, the definition of $\mathrm{Z}_L(A)$ depends on L. We intend to overcome this dependence.

There is an analogue of logical equivalence, namely the relation

$$\{(f, g) \in K[\boldsymbol{X}] \times K[\boldsymbol{X}]\colon\ \mathrm{Z}_L(f) = \mathrm{Z}_L(g)\}.$$

We shall not be interested in this. The quotient $\mathrm{Sen}(\mathscr{S})/{\sim}$ is a Boolean algebra (by Theorem 148, page 152) and therefore a Boolean ring (by Theorem 145, page 147); but $K[\boldsymbol{X}]$ is already a ring, albeit not a Boolean ring.

The set $\mathrm{I}_K(A)$ is the **ideal of A in $K[\boldsymbol{X}]$.** this terminology is justified by the following, which is a partial analogue of parts of Theorems 138 and 149 (pages 139 and 153):

Theorem 192.

1. *The zero-loci of subsets of $K[\boldsymbol{X}]$ compose a topology on L^n.*

[1] More precisely, *affine algebraic sets.*

2. *The subsets* $\mathrm{I}_K(\boldsymbol{x})$ *of* $K[\boldsymbol{X}]$ *are prime ideals.*
3. *The subsets* $\mathrm{I}_K(A)$ *of* $K[\boldsymbol{X}]$ *are radical ideals.*

Proof. 1. Because

$$\varnothing = \mathrm{Z}_L(1), \qquad \mathrm{Z}_L(f) \cup \mathrm{Z}_L(g) = \mathrm{Z}_L(fg),$$

the sets $\mathrm{Z}_L(f)$ compose a basis of a topology on L^n.

2. The additional observations

$$\mathrm{Z}_L(0) = L^n, \qquad \mathrm{Z}_L(f) \cap \mathrm{Z}_L(g) \subseteq \mathrm{Z}_L(f-g)$$

show that each $\mathrm{I}_K(\boldsymbol{x})$ is a prime ideal. Indeed, we can translate the three equations and one inclusion as

$$\begin{aligned} &1 \notin \mathrm{I}_K(\boldsymbol{x}), && f \in \mathrm{I}_K(\boldsymbol{x}) \text{ OR } g \in \mathrm{I}_K(\boldsymbol{x}) \iff fg \in \mathrm{I}_K(\boldsymbol{x}), \\ &0 \in \mathrm{I}_K(\boldsymbol{x}), && f \in \mathrm{I}_K(\boldsymbol{x}) \;\&\; g \in \mathrm{I}_K(\boldsymbol{x}) \implies f-g \in \mathrm{I}_K(\boldsymbol{x}). \end{aligned}$$

3. Prime ideals are radical, and the intersection of radical ideals is radical by Theorem 115 (page 117). □

Thus in particular the monoid $(K[\boldsymbol{X}], 1, \cdot\ )$ is analogous with the algebra $(\mathrm{Sen}(\mathscr{S}), \bot, \vee)$ and hence with the monoid $(\mathrm{Sen}(\mathscr{S}), \bot, \vee)/{\sim}$. In developing model theory, in place of the relation $\vDash$, we could have used $\nvDash$ (thus replacing pairs $(\mathfrak{A}, \sigma)$ with $(\mathfrak{A}, \neg\sigma)$); then the monoid $(K[\boldsymbol{X}], 1, \cdot)$ would be analogous to the monoid $(\mathrm{Sen}(\mathscr{S}), \top, \wedge)/{\sim}$, and ideals $\mathrm{I}_K(A)$ of $K[\boldsymbol{X}]$ as a ring would be analogous to *ideals* of $\mathrm{Lin}_0(\mathscr{S})$ as a Boolean algebra, rather than to filters as they are now.

The topology given by the theorem is the **Zariski topology,** or more precisely the K-Zariski topology. The closed subsets of $K[\boldsymbol{X}]$, that is, the ideals of subsets of L^n, are radical ideals of $K[\boldsymbol{X}]$. But we do not know whether *every* radical ideal is closed. Equivalently (since every radical ideal is an intersection of prime ideals by Theorem 117, page 117), we do not know whether every prime ideal is closed.

Recall that, as defined on page 105, the *spectrum* $\mathrm{Spec}(R)$ of a commutative ring R is the set of prime ideals of R, and (by Theorem 113, page 114) it is a compact Kolmogorov space with basis consisting of the sets $\{\mathfrak{p} \in \operatorname{Spec}(R) \colon a \in \mathfrak{p}\}$, denoted by $\mathrm{Z}(a)$ (without a subscript), where $a \in R$. We now have the following partial analogue of part of Theorem 143 (page 143):

Theorem 193. *The map $\boldsymbol{x} \mapsto \mathrm{I}_K(\boldsymbol{x})$ from L^n to* $\operatorname{Spec}(K[\boldsymbol{X}])$ *is continuous, and the image of L^n under this map is a Kolmogorov quotient of L^n with respect to the map.*

Proof. We use Theorem 142 (page 142). Let us refer to the map $\boldsymbol{x} \mapsto \mathrm{I}_K(\boldsymbol{x})$ as Φ. If $f \in K[\boldsymbol{X}]$, then

$$\begin{aligned}\Phi^{-1}[\mathrm{Z}(f)] &= \{\boldsymbol{x} \in L^n \colon \mathrm{I}_K(\boldsymbol{x}) \in \mathrm{Z}(f)\}\\ &= \{\boldsymbol{x} \in L^n \colon f \in \mathrm{I}_K(\boldsymbol{x})\}\\ &= \{\boldsymbol{x} \in L^n \colon f(\boldsymbol{x}) = 0\}\\ &= \mathrm{Z}_L(f);\end{aligned}$$

thus Φ is continuous. Also

$$\begin{aligned}\Phi[\mathrm{Z}_L(f)] &= \{\mathrm{I}_K(\boldsymbol{x}) \colon x \in \mathrm{Z}_L(f)\}\\ &= \{\mathrm{I}_K(\boldsymbol{x}) \colon f(x) = 0\}\\ &= \{\mathrm{I}_K(\boldsymbol{x}) \colon x \in L^n \;\&\; f \in \mathrm{I}_K(\boldsymbol{x})\}\\ &= \Phi[L^n] \cap \{\mathfrak{p} \in \operatorname{Spec}(K[\boldsymbol{X}]) \colon f \in \mathfrak{p}\}\\ &= \Phi[L^n] \cap \mathrm{Z}(f);\end{aligned}$$

so Φ is closed onto its image. Finally, $\boldsymbol{x}$ and $\boldsymbol{y}$ in L^n are topologically indistinguishable if and only if $\Phi(\boldsymbol{x}) = \Phi(\boldsymbol{y})$. □

The situation is as in Figure 9.2, a collapsed analogue of Figure 5.6 (page 154). The function $\boldsymbol{x} \mapsto \mathrm{I}_K(\boldsymbol{x})$ injective on K^n, since if $\boldsymbol{a} \in K^n$ then

$$\mathrm{I}_K(\boldsymbol{a}) = (X^0 - a^0, \dots, X^{n-1} - a^{n-1}).$$

The map is not generally injective: if $n = 2$, $K = \mathbb{Q}$, and L is $\mathbb{R}$ or $\mathbb{C}$, then

$$\mathrm{I}_K((\pi, \pi)) = (X - Y) = \mathrm{I}_K((\mathrm{e}, \mathrm{e})).$$

The map is not generally surjective either: If $L = \mathbb{Q}^{\mathrm{alg}}$, then $(X - Y)$ is not in its range, although $\mathrm{I}_K(\{(x, x) \colon x \in \mathbb{Q}^{\mathrm{alg}}\}) = (X - Y)$.

$$\begin{tikzcd}
L^n \arrow[r, leftrightsquigarrow, "f(\boldsymbol{x})=0"] \arrow[d, "\boldsymbol{x}\mapsto\mathrm{I}_K(\boldsymbol{x})"'] & K[\boldsymbol{x}] \arrow[dl, leftrightsquigarrow, "\ni"] \\
\mathrm{Spec}(K[\boldsymbol{X}]) &
\end{tikzcd}$$

Figure 9.2.: The spectrum of a ring of polynomials

Theorem 194. *If $K[\boldsymbol{X}]^{\mathrm{alg}}$ embeds over K in L, then the map $\boldsymbol{x} \mapsto \mathrm{I}_K(\boldsymbol{x})$ on L^n is surjective onto* $\mathrm{Spec}(K[\boldsymbol{X}])$.

Proof. Suppose $\mathfrak{p} \in \mathrm{Spec}(K[\boldsymbol{X}])$. If $K[\boldsymbol{X}]/\mathfrak{p} \subseteq L$, and $\boldsymbol{x}$ is $(X^k + \mathfrak{p} \colon k < n)$, then

$$\mathrm{I}_K(\boldsymbol{x}) = \mathfrak{p}.$$

Then the same is true if $K[\boldsymbol{X}]/\mathfrak{p}$ embeds over K in L, and $\boldsymbol{x}$ is the image in L of $(X^k + \mathfrak{p} \colon k < n)$. Since $K[\boldsymbol{X}]/\mathfrak{p}$ is an integral domain of transcendence degree no greater than n over K, it embeds over K in $K[\boldsymbol{X}]^{\mathrm{alg}}$. □

Thus we have an analogue of the Compactness Theorem (page 171), the spectrum of a polynomial ring being analogous to the Stone space of a Lindenbaum algebra.

9.2. Hilbert Basis Theorem

By Theorem 192, every zero-locus is the zero-locus of a radical ideal:

$$\mathrm{Z}_L(A) = \mathrm{Z}_L((A)) = \mathrm{Z}_L(\sqrt{}(A)).$$

Theorem 195. *If $\mathfrak{a}$ and $\mathfrak{b}$ are two ideals of $K[\boldsymbol{X}]$, then*

$$\mathrm{Z}_L(\mathfrak{a}) \cup \mathrm{Z}_L(\mathfrak{b}) = \mathrm{Z}_L(\mathfrak{a} \cap \mathfrak{b}), \tag{9.1}$$

Proof. Easily $\mathrm{Z}_L(\mathfrak{a}) \cup \mathrm{Z}_L(\mathfrak{b}) \subseteq \mathrm{Z}_L(\mathfrak{a} \cap \mathfrak{b})$. The reverse inclusion holds because

$$\surd(\mathfrak{a} \cap \mathfrak{b}) = \surd(\{fg \colon f \in \mathfrak{a} \ \& \ g \in \mathfrak{b}\}). \qquad \square$$

The union of the zero-loci of an *arbitrary* collection of ideals need not be the zero-locus of the intersection of the ideals. For example, if $K = \mathbb{Q}$ (and L is some larger field) and

$$\mathfrak{a}_k = \left(\prod_{i=1}^{k}(X - i)\right) = \big((X-1)\cdots(X-k)\big),$$

then $\mathrm{Z}_L(\mathfrak{a}_k) = \{1, \dots, k\}$, but $\bigcap_{k\in\mathbb{N}} \mathfrak{a}_k = \{0\}$. Thus

$$\bigcup_{k\in\mathbb{N}} \mathrm{Z}_L(\mathfrak{a}_k) = \mathbb{N} \subset L = \mathrm{Z}_L(\{0\}) = \mathrm{Z}_L\left(\bigcap_{k\in\mathbb{N}} \mathfrak{a}_k\right).$$

AC **Theorem 196** (Hilbert Basis Theorem). *For every n in ω, by the Axiom of Choice, every ideal of the polynomial ring $K[X^0, \dots, X^{n-1}]$ is finitely generated.*

Proof. The claim implies, and is therefore equivalent to, an apparently stronger claim, namely that every ideal (A) of $K[X^0, \dots, X^{n-1}]$ is (A_0) for some finite subset A_0 of A. For, if $(A) = (f_0, \dots, f_{m-1})$, then each f_k is in $(A^{(k)})$ for some finite subset $A^{(k)}$ of A; and then we can let $A_0 = \bigcup_{k<m} A^{(k)}$.

The claim as also equivalent to the claim that every sequence $(\mathfrak{a}_k \colon k \in \omega)$ of ideals of $K[X^0, \dots, X^{n-1}]$ such that

$$\mathfrak{a}_0 \subseteq \mathfrak{a}_1 \subseteq \mathfrak{a}_2 \subseteq \cdots$$

—that is, every increasing chain of ideals (indexed by ω)—is eventually constant. For, the union of such a chain is an ideal $\mathfrak{b}$, and if this ideal is finitely generated, then it has a generating set whose elements all lie in some $\mathfrak{a}_\ell$, and then this ideal is $\mathfrak{b}$. Conversely (or inversely), if $\mathfrak{a}$ were not finitely generated, then for all subsets $\{f_k \colon k < \ell\}$ of $\mathfrak{a}$ we could find f_ℓ in $\mathfrak{a} \smallsetminus (f_k \colon k < \ell)$; thus we could form a strictly increasing chain $((f_k \colon k < \ell) \colon \ell \in \omega)$.

We now have also a fourth form of our claim: every *countably* generated ideal of $K[X^0, \dots, X^{n-1}]$ is finitely generated. We turn to proving the claim, in any convenient form.

The claim is trivially true when $n = 0$, since a field has only two ideals: the trivial ideal and the improper ideal (1).

The claim is still easy when $n = 1$, because $K[X]$ is a **Euclidean domain.** That is, if f and g are in $K[X]$ and are not both 0, we can use the Euclidean algorithm (as on page 75) to find their greatest common divisor—say h; and then $(f, g) = (h)$. Hence if $\mathfrak{a} = (f_k \colon k \in \omega)$, then for each k in ω we can find g_k so that

$$(f_0, \dots, f_k) = (g_k).$$

In particular, g_{k+1} divides g_k. Then $\min\{\deg(g_k) \colon k \in \omega\} = \deg(g_\ell)$ for some ℓ, and consequently $\mathfrak{a} = (g_\ell)$.

When $n \geqslant 2$, we have not got the Euclidean algorithm; but we can come close enough if we use induction. Suppose then that the claim is true when $n = m$. Let $\mathfrak{a}$ be an ideal of $K[X^0, \dots, X^m]$. We shall form a sequence $(f_0, f_1, \dots)$ of elements of $\mathfrak{a}$ by recursion. Given $(f_k \colon k < \ell)$, and using the Axiom of Choice, we let f_ℓ, if it exists, be an element of $\mathfrak{a} \smallsetminus (f_k \colon k < \ell)$ of minimal degree as a polynomial in X^m over $K[X^0, \dots, X^{m-1}]$. Then these degrees form an increasing sequence: **AC**

$$\deg_{X^m}(f_0) \leqslant \deg_{X^m}(f_1) \leqslant \deg_{X^m}(f_2) \leqslant \cdots$$

Let g_k be the leading coefficient of f_k (as a polynomial in X^m over $K[X^0, \dots, X^{m-1}]$; so $g_k \in K[X^0, \dots, X^{m-1}]$). By inductive hypothesis, for some ℓ,

$$(g_k \colon k \in \omega) = (g_k \colon k < \ell).$$

Then in particular $g_\ell \in (g_k \colon k < \ell)$, so by Theorem 77 (page 85), for some b^k in $K[X^0, \dots, X^{m-1}]$,

$$g_\ell = \sum_{k<\ell} b^k \cdot g_k.$$

Now let

$$h = \sum_{k<\ell} b^k \cdot f_k \cdot (X^m)^{r(k)},$$

where $r(k) = \deg_{X^m}(f_\ell) - \deg_{X^m}(f_k)$. Then $h \in (f_k \colon k < \ell)$ and, as a polynomial in X^m over $K[X^0, \dots, X^{m-1}]$, has the leading coefficient and degree of f_ℓ. But then $f_\ell - h$ has lower degree and belongs to $\mathfrak{a} \smallsetminus (f_k \colon k < \ell)$; that is, f_ℓ did not have minimal degree. Thus there *is* no f_ℓ; that is, $\mathfrak{a} = (f_k \colon k < \ell)$. □

A *singly* generated ideal is called **principal.** Then part of our proof of the theorem gives the following:

Porism 196.1. *Every ideal of $K[X]$ is principal.*

Hence, although in the example above $\mathbb{N}$ is the union of zero-loci, it cannot itself be a zero-locus; for, every zero-locus of polynomials in one variable is the zero-locus of a single polynomial, so it is either the whole field L or a finite subset of this.

The Hilbert Basis Theorem itself has the following:

Corollary 196.1. *Every decreasing chain of closed subsets of L^n is eventually constant. In particular, the Zariski topology is compact.*

The corollary would imply the theorem, if we knew that that $\mathfrak{a} \subset \mathfrak{b}$ implied $\mathrm{Z}_L(\mathfrak{a}) \supset \mathrm{Z}_L(\mathfrak{b})$, at least when $\mathfrak{a}$ and $\mathfrak{b}$ were radical ideals. However, this implication can fail. For example, when $L = \mathbb{R}$, then (X^2+1) is a radical ideal whose zero-locus is the same as the zero-locus of (1), namely the empty set.

9.3. Specialization

We denote the fraction-field of $K[\boldsymbol{X}]$ by

$$K(\boldsymbol{X});$$

it is the **field of rational functions** in $\boldsymbol{X}$ over K.

Suppose now $\boldsymbol{a} \in L^n$. Then there is a homomorphism $f \mapsto f(\boldsymbol{a})$ from $K[\boldsymbol{X}]$ to L. (We could write the homomorphism also as $X^i \mapsto a^i$.) The range of this homomorphism is denoted by

$$K[\boldsymbol{a}],$$

and the fraction-field of this ring is denoted by

$$K(\boldsymbol{a});$$

we may consider this field as a subfield of L. Then $K[\boldsymbol{a}]$ is the smallest sub-*ring* of L that includes $K \cup \{a_0, \dots, a_{n-1}\}$, and $K(\boldsymbol{a})$ is the smallest sub*field* of L that includes $K \cup \{a_0, \dots, a_{n-1}\}$.

Let $\mathfrak{p}$ be the kernel of the homomorphism $f \mapsto f(\boldsymbol{a})$ from $K[\boldsymbol{X}]$ to L. Then

$$K[\boldsymbol{a}] \cong K[\boldsymbol{X}]/\mathfrak{p}.$$

Also, $f \mapsto f(\boldsymbol{a})$ is well-defined on the sub-ring $K[\boldsymbol{X}]_{\mathfrak{p}}$ of $K(\boldsymbol{X})$, but not on the complement. This complement is empty, if $\mathfrak{p} = (0)$. If $\mathfrak{p} \neq (0)$, then $\boldsymbol{a}$ is said to be **algebraically dependent** over K, or simply **algebraic** over K in case $n = 1$.

Theorem 197. *If a is algebraic over K, then*

$$K[a] = K(a).$$

Thus nontrivial prime ideals of $K[X]$ are maximal.

Proof. If $a \in K$, then $K[a] = K = K(a)$. If $a \notin K$, but is algebraic over K, then $b_0 + b_1 \cdot a + \dots + b_n \cdot a^m = 0$ for some b_i in K, where $b_0 \neq 0$ (and $m > 0$). Then

$$\frac{1}{a} = -\left(\frac{b_1}{b_0} + \frac{b_2}{b_0} \cdot a + \dots + \frac{b_m}{b_0} \cdot a^{m-1}\right). \qquad \square$$

Easily, $K[X]$ is not a von Neumann regular ring. However, being an integral domain, it is reduced. It is not a counterexample to Theorem 127 (page 127), because the prime ideal (0) is not maximal.

The nontrivial prime ideals of $K[\boldsymbol{X}]$ are not generally maximal. For example $K[X,Y]/(X-Y) \cong K[X]$, which is an integral domain that is not a field; so $(X-Y)$ is a non-maximal prime ideal of $K[X,Y]$.

The field K is **algebraically closed** if every element of a larger field that is algebraic over K is already in K. (The notion was used in Theorem 174, page 191.) An **algebraic closure** of K is an algebraically closed extension of K that has no proper algebraically closed sub-extension.

AC **Theorem 198.** *By the Axiom of Choice, every field K has an algebraic closure. All algebraic closures of K are isomorphic over K.*

We may therefore refer to *the* algebraic closure of K, denoting it by

$$K^{\mathrm{alg}}.$$

9.4. Hilbert Nullstellensatz

The closed subsets of $K[\boldsymbol{X}]$ with respect to the Galois correspondence between $\mathscr{P}(L^n)$ and $\mathscr{P}(K[\boldsymbol{X}])$ defined on page 217—let us refer to these closed subsets more precisely as **L-closed,** because we are going to consider what happens when we change L. Again, by Theorem 192 (page 218), the L-closed subsets of $K[\boldsymbol{X}]$ are radical ideals, and so they have the form of $\mathrm{I}_K(\mathrm{Z}_L(\mathfrak{a}))$ for some radical ideal $\mathfrak{a}$ of $K[\boldsymbol{X}]$; and then

$$\mathfrak{a} \subseteq \mathrm{I}_K(\mathrm{Z}_L(\mathfrak{a})). \tag{9.2}$$

This is an equation if and only if $\mathfrak{a}$ is L-closed. We noted in effect (on page 224) that if $L=\mathbb{R}$ (and K is an arbitrary subfield of this), then the radical ideal (X^2+1) is not L-closed:

$$(X^2+1) \subset (1) = \mathrm{I}_K(\mathrm{Z}_{\mathbb{R}}((X^2+1))).$$

However, as L grows larger, so does $\mathrm{Z}_L(\mathfrak{a})$; but then $\mathrm{I}_K(\mathrm{Z}_L(\mathfrak{a}))$ becomes smaller. In fact

$$(X^2+1) = \mathrm{I}_K(\mathrm{Z}_{\mathbb{C}}((X^2+1))).$$

We now are faced with the following:

Question 1. *For every radical ideal $\mathfrak{a}$ of $K[\boldsymbol{X}]$, is there an extension L of K large enough that*

$$\mathfrak{a} = \mathrm{I}_K(\mathrm{Z}_L(\mathfrak{a}))?$$

Question 2. *Is there an extension L of K large enough that for all ideals $\mathfrak{a}$ of $K[\boldsymbol{X}]$ and all extensions M of K,*

$$\mathrm{I}_K(\mathrm{Z}_L(\mathfrak{a})) \subseteq \mathrm{I}_K(\mathrm{Z}_M(\mathfrak{a}))?$$

Note well that $\mathfrak{a}$ and L are quantified in different orders in the two questions, as $\forall\mathfrak{a}\ \exists L$ and $\exists L\ \forall\mathfrak{a}$ respectively. But the conclusions are different. So it is not immediate that an answer to one question yields the answer to the other question. However, if the answer to Question 1 is indeed yes, then so is the answer to Question 2, if the different fields L corresponding to the different ideals $\mathfrak{a}$ are all included in one large field. They *are* so included, since the class of fields has the **joint embedding property:** If f_0 embeds K in L_0, and f_1 embeds K in L_1, then there is a field M, and there are embeddings g_i of the L_i (respectively) in M, such that $g_0 \circ f_0 = g_1 \circ f_1$. See Figure 9.3.

$$\begin{tikzcd}
 & M & \\
L_0 \arrow[ur, dash, "g_0"] & & L_1 \arrow[ul, dash, "g_1"'] \\
 & K \arrow[ul, dash, "f_0"] \arrow[ur, dash, "f_1"'] &
\end{tikzcd}$$

Figure 9.3.: Joint embedding property of fields

By contrast, even if Question 2 has a positive answer, it is not at all clear that the answer to Question 1 must be positive.

We settle Question 1 first in a special case.

Lemma 21. *For all* maximal *ideals* $\mathfrak{m}$ *of* $K[\boldsymbol{X}]$*, for all extensions* L *of* K *in which* $K[\boldsymbol{X}]/\mathfrak{m}$ *embeds over* K*,*

$$\mathfrak{m} = \mathrm{I}_K(\mathrm{Z}_L(\mathfrak{m})).$$

Proof. As formulated here, the lemma almost proves itself. We just have to show $\mathrm{I}_K(\mathrm{Z}_L(\mathfrak{m}))$ is a proper ideal. But the image of $\boldsymbol{X}$ in $K[\boldsymbol{X}]/\mathfrak{m}$ is in the zero-locus of $\mathfrak{m}$. In particular, if L includes this field, then $\mathrm{Z}_L(\mathfrak{m})$ is not empty, so $\mathrm{I}_K(\mathrm{Z}_L(\mathfrak{m}))$ cannot be all of $K[\boldsymbol{X}]$. □

Since $K[\boldsymbol{X}]/\mathfrak{m}$ is a field by Theorem 78 (page 86), one can show that this field is algebraic over K (as in [41, Ch. IX, Cor. 1.2, p. 379]); but we shall not need this. The lemma yields another another special case of the desired general result:

AC **Theorem 199.** *If* $K[\boldsymbol{X}]^{\mathrm{alg}} \subseteq L$*, then by the Axiom of Choice, for all ideals* $\mathfrak{a}$ *of* $K[\boldsymbol{X}]$ *such that* $\mathrm{I}_K(\mathrm{Z}_L(\mathfrak{a}))$ *is the improper ideal,*

$$\mathfrak{a} = \mathrm{I}_K(\mathrm{Z}_L(\mathfrak{a})).$$

Proof. The claim is

$$\mathrm{I}_K(\mathrm{Z}_L(\mathfrak{a})) = (1) \implies \mathfrak{a} = (1).$$

We prove the contrapositive. If $\mathfrak{a}$ is a proper ideal of $K[\boldsymbol{X}]$, then by
AC the Maximal Ideal Theorem (101), it is included in some maximal ideal $\mathfrak{m}$. The field $K[\boldsymbol{X}]/\mathfrak{m}$ can be understood as an algebraic extension of $K(X^i \colon i \in I)$ for some subset I of n, so it embeds in $K(\boldsymbol{X})^{\mathrm{alg}}$. By the lemma then, since $\mathrm{I}_K(\mathrm{Z}_L(\mathfrak{m}))$ is a proper ideal, so is $\mathrm{I}_K(\mathrm{Z}_L(\mathfrak{a}))$. □

Note that if $\mathrm{I}_K(\mathrm{Z}_L(\mathfrak{a})) \neq (1)$, then $\mathrm{Z}_L(\mathfrak{a}) \neq \varnothing$. Thus every proper ideal has non-empty zero-locus in a large-enough field. *Nullstellensatz* means zero-locus theorem:

Theorem 200 (Nullstellensatz)**.** *If* $K[\boldsymbol{X}, Y]^{\mathrm{alg}} \subseteq L$*, then for all radical ideals* $\mathfrak{a}$ *of* $K[\boldsymbol{X}]$*,*

$$\mathfrak{a} = \mathrm{I}_K(\mathrm{Z}_L(\mathfrak{a})).$$

Proof. Say $f \in \mathrm{I}_K(\mathrm{Z}_L(\mathfrak{a}))$. If $\boldsymbol{x} \in \mathrm{Z}_L(\mathfrak{a})$, then $f(\boldsymbol{x}) = 0$. This shows $\mathrm{Z}_L(\mathfrak{a} \cup \{f - 1\}) = \varnothing$, so

$$\mathrm{I}_K(\mathrm{Z}_L(\mathfrak{a} \cup \{f - 1\})) = (1).$$

By the last theorem, $\mathfrak{a} \cup \{f - 1\}$ too must generate the improper ideal of $K[\boldsymbol{X}]$. We want to be able to conclude $f \in \mathfrak{a}$. To do so, we modify the argument so far. We have $f \cdot Y \in \mathrm{I}_K(\mathrm{Z}_L(\mathfrak{a}))$, if we consider $\mathfrak{a}$ now as a subset of $K[\boldsymbol{X}, Y]$. As before, $\mathfrak{a} \cup \{f \cdot Y - 1\}$ must generate the improper ideal of $K[\boldsymbol{X}, Y]$. Now, by itself, $\mathfrak{a}$ generates the ideal of $K[\boldsymbol{X}, Y]$ whose elements are polynomials in Y with coefficients from $\mathfrak{a}$. Hence there is some such polynomial g, and there is some h in $K[\boldsymbol{X}, Y]$, such that

$$g + h \cdot (f \cdot Y - 1) = 1.$$

Substituting $1/f$ for Y, we get $g(1/f) = 1$; that is,

$$g_0 + g_1 \cdot \frac{1}{f} + \dots g_m \cdot \frac{1}{f^m} = 1$$

for some g_i in $\mathfrak{a}$, and hence

$$g_0 \cdot f^m + g_1 \cdot f^{m-1} + \dots + g_m = f^m.$$

This means $f^m \in \mathfrak{a}$. Assuming $\mathfrak{a}$ is radical, we have $f \in \mathfrak{a}$. Thus $\mathrm{I}_K(\mathrm{Z}_L(\mathfrak{a})) \subseteq \mathfrak{a}$ and therefore $\mathrm{I}_K(\mathrm{Z}_L(\mathfrak{a})) = \mathfrak{a}$. □

We have now settled both Questions 1 and 2. This suggests that understanding algebraic sets can somehow be reduced to understanding radical ideals of $K[\boldsymbol{X}]$. Indeed, there is *some* extension L of K large enough that we have a Galois correspondence between the K-closed subsets of L^n and the radical ideals of $K[\boldsymbol{X}]$. It is not particularly important for what follows that this field L can be chosen as K^{alg}. Nonetheless, it is true: Theorem 199 holds, merely under the hypothesis $K^{\mathrm{alg}} \subseteq L$.

Theorem 201 (Hilbert's Nullstellensatz, weak form)**.** *All proper ideals of $K[\boldsymbol{X}]$ have non-empty zero-loci in all extensions of K^{alg}.*

Proof. In the lemma, by the Hilbert Basis Theorem, $\mathfrak{m}$ has the form $(f_0, \dots, f_\ell)$ for some f_i in $K[\boldsymbol{X}]$. Thus the formula

$$f_0 = 0 \wedge \dots \wedge f_\ell = 0$$

has a solution in $K[\boldsymbol{X}]/\mathfrak{m}$ and *a fortiori* in $(K[\boldsymbol{X}]/\mathfrak{m})^{\mathrm{alg}}$. The latter field is an *elementary* extension of K^{alg}, by the model-completeness of the theory of algebraically closed fields (Theorem 175 on page 192). Therefore the formula has a solution here too. Thus as long as $K^{\mathrm{alg}} \subseteq L$, we have $\mathrm{Z}_L(\mathfrak{m}) \neq 0$. □

As an alternative to using the model-completeness of the theory of algebraically closed fields, one can use the result mentioned above, that $K[\boldsymbol{X}]/\mathfrak{m}$ is algebraic over K. In any case, the proof of Theorem 200 gives:

Corollary 201.1 (Hilbert's Nullstellensatz, strong form). *For all radical ideals $\mathfrak{a}$ of $K[\boldsymbol{X}]$,*

$$\mathrm{I}_K(\mathrm{Z}_{K^{\mathrm{alg}}}(\mathfrak{a})) = \mathfrak{a}.$$

10. Finite fields

10.1. Ultraproducts of finite structures

Suppose a theory T has arbitrarily large finite models. Then there is a sequence $(\mathfrak{A}_m \colon m \in \omega)$ of finite models of T such that $|A_m| > m$ in each case. Consequently, the sentence

$$\exists(x_0, \dots, x_m) \bigwedge_{i<j<m} x_i \neq x_j$$

is true in each $\mathfrak{A}_n$ such that $m \leqslant n$. By Łoś's Theorem then, the sentence is true in every non-principal ultraproduct of the structures $\mathfrak{A}_i$. In particular, this ultraproduct is infinite. Moreover, every sentence that is true in each $\mathfrak{A}_i$ is true in the ultraproduct; that is, the ultraproduct is a model of the theory of the structures $\mathfrak{A}_i$. Thus the ultraproduct is an infinite model of the theory of finite models of T. Such a structure might be called a **pseudo-finite** model of T. We shall consider the case where T is the theory of fields.

10.2. Finite fields

Let us review the basic theorems about finite fields. Suppose K is a field. There is a homomorphism $1 \mapsto 1$ (or $k \mapsto k \cdot 1$) from $\mathbb{Z}$ to K. The kernel of this homomorphism is $n\mathbb{Z}$ for some *positive* n, called the **characteristic** of K, $\operatorname{char}(K)$. Since $\mathbb{Z}/n\mathbb{Z}$ must be an integral domain (by Corollary 121.1, page 123), n is either 0 or prime. If $\operatorname{char}(K) = 0$, we may consider $\mathbb{Q}$ as a subfield of K; if $\operatorname{char}(K)$ is a prime p, we consider $\mathbb{Z}/p\mathbb{Z}$, denoted by $\mathbb{F}_p$, as a subfield of K. Respectively, $\mathbb{Q}$ or $\mathbb{F}_p$ is the **prime field** of K.

Let K be a finite field of characteristic p. Then K is a vector-space over $\mathbb{F}_p$ of some finite dimension m, so K has order p^m. The group

$K^\times$ of units of K has order $p^m - 1$, so its every element is a root of $x^{p^m-1} - 1$. Then *every* element of K is a root of the polynomial

$$x^{p^m} - x.$$

Since the formal derivative of this is -1, it has no repeated roots. Thus its roots (in an algebraic closure ${\mathbb{F}_p}^{\mathrm{alg}}$ of $\mathbb{F}_p$ that includes K) are precisely the elements of K: we have

$$K = \{x \in {\mathbb{F}_p}^{\mathrm{alg}} \colon x^{p^m} = x\}.$$

Conversely, for all m in $\mathbb{N}$, since the map $x \mapsto x^{p^m}$ is an automorphism of ${\mathbb{F}_p}^{\mathrm{alg}}$, the set $\{x \in {\mathbb{F}_p}^{\mathrm{alg}} \colon x^{p^m} = x\}$ (namely the fixed field of the automorphism) is a subfield having order p^m. This then is the *unique* subfield of ${\mathbb{F}_p}^{\mathrm{alg}}$ of this order, and we can denote it by

$$\mathbb{F}_{p^m}.$$

The group ${\mathbb{F}_{p^m}}^\times$ of units of this field is cyclic. For again, it is a finite abelian group of order $p^m - 1$ and is therefore a direct product

$$\prod_{\ell \mid p^m-1} G_\ell,$$

where each G_ℓ is an ℓ-group (a group whose elements have orders that are powers of ℓ; here and elsewhere in this chapter, ℓ is, like p, a prime number). Since G_ℓ is finite, for some positive integer n, every element of G_ℓ is a solution of

$$x^{\ell^n} = 1.$$

But in a field, this equation has no more than ℓ^n solutions. Therefore, if n is minimal, G_ℓ must be cyclic of order ℓ^n. Then the product ${\mathbb{F}_{p^m}}^\times$ is itself cyclic, of order $p^m - 1$.

The collection of finite subfields of ${\mathbb{F}_p}^{\mathrm{alg}}$, ordered by inclusion, is isomorphic, under the map $\mathbb{F}_{p^m} \mapsto m$, to $\mathbb{N}$ as ordered by dividing. That is,

$$\mathbb{F}_{p^m} \subseteq \mathbb{F}_{p^n} \iff m \mid n.$$

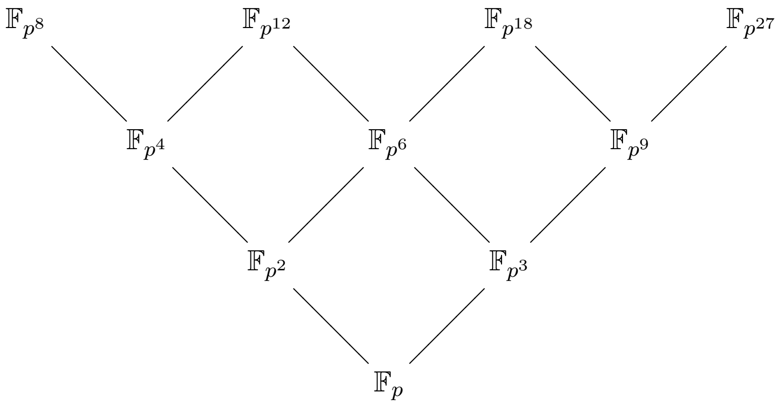

Figure 10.1.: The lattice (in part) of finite fields of characteristic p

See Figure 10.1. Indeed, if $\mathbb{F}_{p^m} \subseteq \mathbb{F}_{p^n}$, then $\mathbb{F}_{p^n}$ is a vector-space over $\mathbb{F}_{p^m}$, so its order is $(p^m)^k$ for some k, and then $n = mk$, so $m \mid n$. Conversely, if $m \mid n$, then

$$p^m - 1 \mid p^n - 1,$$

and therefore

$$x^{p^m-1} - 1 \mid x^{p^n-1} - 1,$$

so $\mathbb{F}_{p^m} \subseteq \mathbb{F}_{p^n}$.

Finally,

$$\mathbb{F}_p{}^{\mathrm{alg}} = \bigcup_{n \in \mathbb{N}} \mathbb{F}_{p^n} \tag{10.1}$$

(since every extension $\mathbb{F}_{p^n}/\mathbb{F}_p$ is certainly algebraic, while every finite algebraic extension of $\mathbb{F}_p$ is a finite field).

10.3. Galois groups

We have shown that for each prime p, for each m in $\mathbb{N}$, there is a subfield $\mathbb{F}_{p^m}$ of $\mathbb{F}_p{}^{\mathrm{alg}}$, and this subfield is generated by (in fact it consists of) the roots of the polynomial $x^{p^m} - x$, which is separable. Therefore the finite field-extension $\mathbb{F}_{p^m}/\mathbb{F}_p$ is normal and separable, that is, Galois. The order of its group of automorphisms is $[\mathbb{F}_{p^m} : \mathbb{F}_p]$, that is, m. But the **Frobenius automorphism** of ${\mathbb{F}_p}^{\mathrm{alg}}$, namely $x \mapsto x^p$ or

$$\mathrm{Frob},$$

restricts to an automorphism of $\mathbb{F}_{p^m}$ of order m, since we have shown in effect

$$\mathrm{Fix}(\mathrm{Frob}^k) = \mathbb{F}_{p^k}.$$

Thus

$$\mathrm{Aut}(\mathbb{F}_{p^m}/\mathbb{F}_p) = \langle \mathrm{Frob} \restriction \mathbb{F}_{p^m} \rangle \cong \mathbb{Z}/m\mathbb{Z}.$$

For any field K, let us write

$$\mathrm{Gal}(K) = \mathrm{Aut}(K^{\mathrm{sep}}/K),$$

the *absolute Galois group* of K. We want to determine $\mathrm{Gal}(\mathbb{F}_p)$. Suppose $\sigma \in \mathrm{Gal}(\mathbb{F}_p)$. For every n in $\mathbb{N}$, we have

$$\sigma \restriction \mathbb{F}_{p^n} \in \mathrm{Aut}(\mathbb{F}_{p^n}/\mathbb{F}_p),$$

and hence for some $\sigma(n)$ in $\mathbb{Z}$

$$\sigma \restriction \mathbb{F}_{p^n} = (\mathrm{Frob} \restriction \mathbb{F}_{p^n})^{\sigma(n)}.$$

All that matters here is the congruence-class of $\sigma(n)$ *modulo* n. Thus we have an injective map

$$\sigma \mapsto (\sigma(n) \colon n \in \mathbb{N})$$

from $\mathrm{Gal}(\mathbb{F}_p)$ to $\prod_{n\in\mathbb{N}} \mathbb{Z}/n\mathbb{Z}$. The map is not surjective, but if $m \mid n$, then since $\mathbb{F}_{p^m} \subseteq \mathbb{F}_{p^n}$ we must have

$$\sigma(n) \equiv \sigma(m) \pmod{m}.$$

However, suppose an element $(\sigma(n) \colon n \in \mathbb{N})$ of $\prod_{n\in\mathbb{N}} \mathbb{Z}/n\mathbb{Z}$ meets this condition. For any x in ${\mathbb{F}_p}^{\mathrm{alg}}$ we can define an element σ of $\mathrm{Gal}(\mathbb{F}_p)$ by

$$x^\sigma = x^{p^{\sigma(m)}},$$

where $x \in \mathbb{F}_{p^m}$. (Here x^σ is of course the image of x under σ.) This definition of x^σ is independent of the choice of m, since if also $x \in \mathbb{F}_{p^n}$, then

$$x \in \mathbb{F}_{p^{\gcd(m,n)}},$$

so

$$\sigma(m) \equiv \sigma(\gcd(m,n)) \equiv \sigma(n) \pmod{\gcd(m,n)}$$

and therefore

$$x^{p^{\sigma(m)}} = x^{p^{\sigma(\gcd(m,n))}} = x^{p^{\sigma(n)}}.$$

Thus

$$\operatorname{Gal}(\mathbb{F}_p) \cong \{(\sigma(n) \colon n \in \mathbb{N}) \in \prod_{i \in \mathbb{N}} \mathbb{Z}/n\mathbb{Z} \colon \bigwedge_{m \mid n} \pi^n_m(\sigma(n)) = \sigma(m)\}$$

where π^n_m is the quotient-map $x + n\mathbb{Z} \mapsto x + m\mathbb{Z}$ from $\mathbb{Z}/n\mathbb{Z}$ to $\mathbb{Z}/m\mathbb{Z}$.

In particular, $\operatorname{Gal}(\mathbb{F}_p)$ has a certain 'universal property' with respect to the system of groups $\mathbb{Z}/n\mathbb{Z}$ and homomorphisms π^n_m:

1. $\operatorname{Gal}(\mathbb{F}_p)$ is a group G from which there is a homomorphism h^G_n to $\mathbb{Z}/n\mathbb{Z}$ for every n in $\mathbb{N}$ such that, if $m \mid n$, then

$$\pi^n_m \circ h^G_n = h^G_m.$$

2. For every such group G, there is a unique homomorphism h from G to $\operatorname{Gal}(\mathbb{F}_p)$ such that, for each n in $\mathbb{N}$,

$$h^G_n = h^{\operatorname{Gal}(\mathbb{F}_p)}_n \circ h.$$

See Figure 10.2. Therefore $\operatorname{Gal}(\mathbb{F}_p)$ is called a **limit** of the given system of groups and homomorphisms. This is the category-theoretic sense of *limit* as given in, say, [16, p. 705] or [3]. Every set of groups, equipped with some homomorphisms, has a limit in this sense, though the limit might be empty.

The group $\operatorname{Gal}(\mathbb{F}_p)$ is called more precisely a **projective limit** or an **inverse limit** of the system of groups $\mathbb{Z}/n\mathbb{Z}$ with the quotient-maps, because any two of these groups are quotients of a third. This condition is not required for the existence of the limit.

$$\begin{tikzcd}
& \mathrm{Gal}(\mathbb{F}_p) \arrow[ddl, "h_n^{\mathrm{Gal}(\mathbb{F}_p)}"'] \arrow[ddr, "h_m^{\mathrm{Gal}(\mathbb{F}_p)}"] & \\
& G \arrow[u, "h"'] \arrow[dl, "h_n^G"] \arrow[dr, "h_m^G"'] & \\
\mathbb{Z}/n\mathbb{Z} \arrow[rr, "\pi_m^n"'] & & \mathbb{Z}/m\mathbb{Z}
\end{tikzcd}$$

Figure 10.2.: The universal property of $\mathrm{Gal}(\mathbb{F}_p)$

We give the finite groups $\mathbb{Z}/n\mathbb{Z}$ the discrete topology, and their product the product topology. This product is compact by the Tychonoff Theorem (page 129). The image of $\mathrm{Gal}(\mathbb{F}_p)$ in this group is closed, so it too is compact: it is called a **pro-finite completion** of the system of finite cyclic groups.[1]

10.4. Pseudo-finite fields

Two examples of infinite models of the theory of finite fields are:

$$\prod_{p \text{ prime}} \mathbb{F}_p/M, \qquad \prod_{n \in \mathbb{N}} \mathbb{F}_{p^n}/M, \tag{10.2}$$

where in each case M is some non-principal maximal ideal. The first example has characteristic 0; the second, characteristic p.

[1] Perhaps one should talk about convergent sequences here…

By the 'Riemann Hypothesis for curves' as proved by Weil,[2] for every prime power q, for every curve C of genus g over $\mathbb{F}_q$, the number of $\mathbb{F}_q$-rational points of C is at least

$$1+q-2g\sqrt{q}.$$

In particular, if q is large enough, then C does have an $\mathbb{F}_q$-rational point.

A field K is called **pseudo-algebraically-closed** or **PAC** if every plane curve defined over K has a K-rational point. This condition entails that every absolutely irreducible variety over K has a K-rational point.[3]

The following are now true of every infinite model of the theory of finite fields:

1. It is perfect.
2. It has exactly one extension of each degree (in some algebraic closure).
3. It is pseudo-algebraically-closed.

This is not obvious, even given the results stated above; one must show that these conditions are *first-order,* that is, the structures that satisfy them make up an elementary class. By the definition of Ax [2], a field with the first two of these properties is **quasi-finite;** with all three of these properties, **pseudo-finite.** So every infinite model of the theory of finite fields is (quasi-finite and) pseudo-finite. Ax proves the converse. In particular, Ax proves that every pseudo-finite field is elementarily equivalent to a non-principal ultraproduct of finite fields, and indeed to one of the ultraproducts given above in (10.2). The method is as follows; here I use Ax [2] and also Chatzidakis [9].

For every field K, the field $\mathrm{Abs}(K)$ of **absolute numbers** of K consists of the algebraic elements of K (here algebraic means algebraic over the prime field). The following is [2, Prop. 7′, §10, p. 261].

Lemma 22. *For every field K of prime characteristic p, there is a*

[2] See for example [26, Ex. V.1.10, p. 368] or [21, Thm 3.14, p. 35].

[3] See [21, ch. 10, pp. 129–131].

maximal ideal M of $\prod_{n\in\mathbb{N}}\mathbb{F}_{p^n}$ such that

$$\operatorname{Abs}(K)\cong\operatorname{Abs}(\prod_{n\in\mathbb{N}}\mathbb{F}_{p^n}/M).$$

Proof. Because ${\mathbb{F}_p}^{\mathrm{alg}}=\bigcup_{n\in\mathbb{N}}\mathbb{F}_{p^n}$ as in (10.1) on page 233, we need only choose M so that, for all m in $\mathbb{N}$,

$$\mathbb{F}_{p^m}\subseteq K\iff\mathbb{F}_{p^m}\subseteq\prod_{n\in\mathbb{N}}\mathbb{F}_{p^n}/M.$$

For each m in $\mathbb{N}$, let f_m be an irreducible element of $\mathbb{F}_p[X]$ of degree m. Then each zero of f_m generates $\mathbb{F}_{p^m}$ over $\mathbb{F}_p$. So we want M to be such that

$$\mathbb{F}_{p^m}\subseteq K\iff f_m\text{ has a zero in }\prod_{n\in\mathbb{N}}\mathbb{F}_{p^n}/M.$$

Let F be the ultrafilter on $\mathbb{N}$ corresponding to M, that is,

$$F=\{\mathbb{N}\smallsetminus\operatorname{supp}(f)\colon f\in M\}=\big\{\{n\colon f_n=0\}\colon f\in M\big\}.$$

Then

$$f_m\text{ has a zero in }\prod_{n\in\mathbb{N}}\mathbb{F}_{p^n}/M\iff\big\{n\colon f_m\text{ has a zero in }\mathbb{F}_{p^n}\big\}\in F.$$

Moreover,

$$f_m\text{ has a zero in }\mathbb{F}_{p^n}\iff m\mid n.$$

So, combining all of our equivalences, we want to choose F on $\mathbb{N}$ such that

$$\mathbb{F}_{p^m}\subseteq K\iff\{n\colon m\mid n\}\in F.$$

For each m in $\mathbb{N}$, the subset

$$\{k\colon k\mid m\ \&\ \mathbb{F}_{p^k}\subseteq K\}$$

of $\mathbb{N}$ is a sublattice of the lattice of factors of m with respect to divisibility: in particular, it contains the least common multiple of any two members. It also contains 1.[4] Therefore it has a maximum element, say $g(m)$. The arithmetic function g is multiplicative:

$$\gcd(m,n)=1 \implies g(mn)=g(m)\cdot g(n).$$

Now let

$$b_m=\{x\colon \gcd(m,x)=g(m)\}.$$

Then the function $m\mapsto b_m$ is also multiplicative, in the sense that

$$\gcd(m,n)=1 \implies b_{mn}=b_m\cap b_n. \tag{10.3}$$

Indeed, suppose $\gcd(m,n)=1$. Then or all x in $\mathbb{N}$,

$$\gcd(mn,x)=\gcd(m,x)\cdot\gcd(n,x),$$

and these factors are co-prime, being respectively factors of m and n. But also $g(mn)=g(m)\cdot g(n)$, and these factors are co-prime, being respectively factors of m and n. Therefore

$$\gcd(mn,x)=g(mn) \iff \gcd(m,x)=g(m)\ \&\ \gcd(n,x)=g(n).$$

So we have (10.3). Moreover, we have also

$$m\leqslant n \implies b_{\ell^n}\subseteq b_{\ell^m}. \tag{10.4}$$

For, we have

$$b_{\ell^n}=\begin{cases}\{g(\ell^n)\cdot y\colon \ell\nmid y\}, & \text{if } g(\ell^n)<\ell^n,\\ \{\ell^n y\colon y\in\mathbb{N}\}, & \text{if } g(\ell^n)=\ell^n,\end{cases}$$

and also

$$m\leqslant n \implies g(\ell^m)=\min\bigl(\ell^m,g(\ell^n)\bigr).$$

Now we can just check that (10.4) holds in each of the three cases

$$g(\ell^n)=\ell^n,\qquad \ell^m<g(\ell^n)<\ell^n,\qquad g(\ell^n)<\ell^m.$$

[4] Thus it contains the least common multiple of every (finite) set of members, including the empty set.

So we have finally

$$b_m \cap b_n = b_{\operatorname{lcm}(m,n)}.$$

Thus, since each b_m is nonempty, the set of these generates a proper filter on $\mathbb{N}$. Let F be an ultrafilter on $\mathbb{N}$ that contains all of the sets b_m. We claim that this F is as desired. Indeed,

- if $\mathbb{F}_{p^m} \subseteq K$, so $g(m) = m$, then $b_m = \{mx \colon x \in \mathbb{N}\}$;
- if $\mathbb{F}_{p^m} \nsubseteq K$, so $g(m) < m$, then $b_m \cap \{mx \colon x \in \mathbb{N}\} = \varnothing$.

Consequently the following are equivalent:

$$\begin{gathered}
\mathbb{F}_{p^m} \subseteq K, \\
\{mx \colon x \in \mathbb{N}\} \in F, \\
f_m \text{ has a root in } \prod_{n\in\mathbb{N}} \mathbb{F}_{p^n}/M, \\
\mathbb{F}_{p^m} \subseteq \prod_{n\in\mathbb{N}} \mathbb{F}_{p^n}/M.
\end{gathered}$$

□

The lemma has a companion [2, Prop. 7], namely that for every quasi-finite field K of characteristic 0, there is a maximal ideal M of $\prod_p \mathbb{F}_p$ such that

$$\operatorname{Abs}(K) = \operatorname{Abs}(\prod_p \mathbb{F}_p/M),$$

but the proof is more difficult. Since all fields of characteristic 0 are perfect, quasi-finiteness in this case just means having exactly one extension of each degree. In this case the field of absolute numbers has *at most* one extension of each degree. This is because, if α is algebraic over $\operatorname{Abs}(K)$, then α has the same degree over K that it has over $\operatorname{Abs}(K)$. For, the minimal polynomial of α over $\operatorname{Abs}(K)$ is a product

$$\prod_{i<n}(X - \alpha_i),$$

the α_i being the conjugates of α over $\operatorname{Abs}(K)$. The minimal polynomial over K is a factor of this; so its coefficients are polynomial functions of (some of) the conjugates of α over $\operatorname{Abs}(K)$. So the coefficients are algebraic (over $\mathrm{Abs}(K)$); therefore the already belong to $\mathrm{Abs}(K)$, by its definition.

We now want to prove [2, Thm 4, §8, p. 255], that if F and F' are pseudo-finite fields, then

$$\mathrm{Abs}(F) \cong \mathrm{Abs}(F') \implies F \equiv F'. \tag{10.5}$$

With this and the foregoing lemma, we shall have that every pseudo-finite field (at least in positive characteristic) is elementarily equivalent to an ultraproduct of finite fields.

To establish (10.5), since $\mathrm{Abs}(F)$ is determined by $\mathrm{Th}(F)$, we can replace F and F' (respectively) by elementarily equivalent fields. In particular, we can replace them with ultrapowers with exponent ω; these ultrapowers are ω_1-saturated by Theorem 178 on page 194. Now take a countable elementary substructure F_0 of F; this exists by the downward Löwenheim–Skolem–Tarski Theorem, Theorem 153. One shows [9, 5.10, Lemme de plongement] that this embeds in F' under a monomorphism φ_0. Then F' has an elementary substructure F_0' that includes the image of F_0; and F_0' embeds in F under a monomorphism φ_0' that extends ${\varphi_0}^{-1}$. Continuing, we obtain isomorphic elementary substructures F_ω and F_ω' of F and F' respectively. See Figure 10.3. This establishes (10.5).

Throughout the chapter, K will be a field, and L will be a field of which K is a subfield, that is,

$$K \subseteq L.$$

$$\begin{tikzcd}
F \arrow[d, dash] & F' \arrow[d, dash] \\
F_\omega \arrow[r, "\bigcup_{n\in\omega}\varphi_n"] \arrow[d, dotted, dash] & F'_\omega \arrow[d, dotted, dash] \\
F_1 \arrow[r, "\varphi_1"] \arrow[d, dash] & \varphi_1[F_1] \arrow[d, dash] \\
\varphi'_0[F'_0] \arrow[d, dash] & F'_0 \arrow[l, "\varphi'_0"'] \arrow[d, dash] \\
F_0 \arrow[r, "\varphi_0"] & \varphi_0[F_0]
\end{tikzcd}$$

Figure 10.3.: Isomorphisms of pseudo-finite fields

11. Schemes

Throughout this chapter, as in Chapter 9 (page 217), K will be a field, and L will be a field of which K is a subfield, that is,

$$K \subseteq L.$$

Sources for the algebraic geometry of this chapter include Coombes [12] and Hartshorne [26]. The main point is to look at the *ultraproduct scheme* at the end; this work is based on the first of the three MSRI/Evans Hall Lectures, given at the University of California at Berkeley in the spring of 1998 by Angus Macintyre.[1]

11.1. Zero-loci

Throughout this section, let $R = K[\boldsymbol{X}]$. In §9.4 (page 226), letting f range over R, and letting $\boldsymbol{x}$ range over some L^n, where $K \subseteq L$, we used the equation $f(\boldsymbol{x}) = 0$ to establish a one-to-one correspondence between the K-closed subsets of L^n and certain radical ideals of R. By Hilbert's Nullstellensatz (page 229), if L includes K^{alg}, then the correspondence is between the K-closed subsets of L^n and (all of) the radical ideals of R. The correspondence is inclusion-reversing. Thus the set of radical ideals of R encodes the topological structure of L^n for L that include K^{alg}.

Suppose indeed $K^{\mathrm{alg}} \subseteq L$, we are given a particular f in R. We are interested in its zero-locus, the K-closed set $\mathrm{Z}_L(f)$; and this now corresponds to the prime ideal $\mathrm{I}_K(\mathrm{Z}_L(f))$, which is $\sqrt{(f)}$. We should should like to have a way of picking out this ideal among all of the radical ideals of $K[\boldsymbol{X}]$, without having to refer to L^n. One way of

[1] These lectures used to be preserved on the MSRI website; but I could not find them there, the last time I looked.

doing this is simply to observe that $\sqrt{}(f)$ is the intersection of all radical ideals of $K[\boldsymbol{X}]$ that contain f. More is true, by Theorem 117 (page 117):

$$\begin{aligned}\sqrt{}(f) &= \bigcap\{\mathfrak{p} \in \operatorname{Spec}(R)\colon f \in \mathfrak{p}\} \\ &= \bigcap \mathrm{Z}(f).\end{aligned} \tag{11.1}$$

We can also give a new proof of this, using the Nullstellensatz. Given an ideal $\mathfrak{a}$ of R, we have

$$\boldsymbol{x} \in \mathrm{Z}_L(\mathfrak{a}) \iff \mathfrak{a} \subseteq \mathrm{I}_K(\boldsymbol{x}),$$

and so

$$\begin{aligned}\mathfrak{a} &\subseteq \bigcap\{\mathfrak{p} \in \operatorname{Spec}(R)\colon \mathfrak{a} \subseteq \mathfrak{p}\} \\ &\subseteq \bigcap\{\mathrm{I}_K(\boldsymbol{x})\colon \boldsymbol{x} \in L^n \ \ \& \ \ \mathfrak{a} \subseteq \mathrm{I}_K(\boldsymbol{x})\} \\ &= \bigcap\{\mathrm{I}_K(\boldsymbol{x})\colon \boldsymbol{x} \in \mathrm{Z}_L(\mathfrak{a})\} \\ &= \mathrm{I}_K(\mathrm{Z}_L(\mathfrak{a})).\end{aligned}$$

The Nullstellensatz then makes the inclusions equalities, if $\mathfrak{a}$ is radical; in general,

$$\sqrt{\mathfrak{a}} = \bigcap\{\mathfrak{p} \in \operatorname{Spec}(R)\colon \mathfrak{a} \subseteq \mathfrak{p}\}.$$

We may use the obvious notation

$$\mathrm{Z}(\mathfrak{a}) = \{\mathfrak{p} \in \operatorname{Spec}(R)\colon \mathfrak{a} \subseteq \mathfrak{p}\} = \bigcap_{f \in \mathfrak{a}} \mathrm{Z}(f),$$

so that

$$\sqrt{\mathfrak{a}} = \bigcap \mathrm{Z}(\mathfrak{a}).$$

So if L is large enough in the sense of including K^{alg}, then we have a one-to-one correspondence between:

- closed subsets $\mathrm{Z}_L(\mathfrak{a})$ of L^n;
- radical ideals $\sqrt{\mathfrak{a}}$ of R;
- closed subsets $\mathrm{Z}(\mathfrak{a})$ of $\operatorname{Spec}(R)$.

We want to understand the sets $\mathrm{Z}(\mathfrak{a})$ as being zero-loci like $\mathrm{Z}_L(\mathfrak{a})$. In (11.1), the condition that $f \in \mathfrak{p}$ is equivalent to the condition that $f + \mathfrak{p} = 0$ in $R/\mathfrak{p}$. Suppose we write $f + \mathfrak{p}$ as $f_\mathfrak{p}$. As in (4.8) on page 128, we have an embedding

$$f \mapsto (f_\mathfrak{p} \colon \mathfrak{p} \in \operatorname{Spec}(R))$$

of R in the product

$$\prod_{\mathfrak{p} \in \operatorname{Spec}(R)} R/\mathfrak{p}.$$

Also

$$\mathrm{Z}(f) = \{\mathfrak{p} \in \operatorname{Spec}(R) \colon f_\mathfrak{p} = 0\},$$

a zero-locus. To establish

$$\mathrm{Z}(\mathfrak{a}) \cup \mathrm{Z}(\mathfrak{b}) = \mathrm{Z}(\mathfrak{a} \cap \mathfrak{b})$$

corresponding to (9.1) on page 221, we need that the functions $\mathfrak{p} \mapsto f_\mathfrak{p}$ on $\operatorname{Spec}(R)$ take values in integral domains; and this is the case, since $f_\mathfrak{p} \in R/\mathfrak{p}$.

It will be useful to have a notation for the *open* subsets of $\operatorname{Spec}(R)$. If $f \in R$, let us write

$$\mathrm{U}_f = \mathrm{Z}(f)^\mathrm{c} = \{\mathfrak{p} \in \operatorname{Spec}(R) \colon f \notin \mathfrak{p}\}.$$

If $A \subseteq R$, we let

$$\mathrm{U}_A = \mathrm{Z}(A)^\mathrm{c} = \bigcup_{f \in A} \mathrm{U}_f = \{\mathfrak{p} \in \operatorname{Spec}(R) \colon A \nsubseteq \mathfrak{p}\}.$$

These are the open subsets of $\operatorname{Spec}(R)$, and each of them is $\mathrm{U}_\mathfrak{a}$ for some radical ideal $\mathfrak{a}$ of R.

11.2. Regular functions

At the beginning of the last section, we considered the equation $f(\boldsymbol{x}) = 0$, where $f \in K[\boldsymbol{X}]$ and $\boldsymbol{x} \in L^n$. We have generally $f(\boldsymbol{x}) \in L$, that is,

f is a function from L^n to L. There can be other such functions. An arbitrary function h from a subset S of L^n to L is **regular** (or more precisely *K-regular*) *at* a point $\boldsymbol{a}$ of S if there is a neighborhood U of $\boldsymbol{a}$ (in the Zariski topology over K, restricted to S) and there are elements f and g of $K[\boldsymbol{X}]$ such that, for all $\boldsymbol{x}$ in U,

$$h(\boldsymbol{x}) = \frac{f(\boldsymbol{x})}{g(\boldsymbol{x})}.$$

The function is **regular,** simply, if it is regular at all points of its domain. The only regular functions on L^n itself are the elements of $K[\boldsymbol{X}]$. However, let

$$S_0 = \mathrm{Z}_L(Y^2 - X^3) \smallsetminus \mathrm{Z}_L(X), \qquad S_1 = \mathrm{Z}_L(Y^2 - X^3) \smallsetminus \mathrm{Z}_L(Y).$$

These are open subsets of their union. On S_0 and S_1 respectively there are regular functions h_0 and h_1 given by

$$h_0(x,y) = \frac{y}{x^2}, \qquad\qquad h_1(x,y) = \frac{x}{y}.$$

These two functions agree on $S_0 \cap S_1$, since $y^2 = x^3$ for all (x,y) in that set (and even in $S_0 \cup S_1$). Thus $h_0 \cup h_1$ is a regular function h on $S_0 \cup S_1$. However, there are no f and g in $K[X,Y]$ such that, for all (x,y) in $S_0 \cup S_1$, $h(x,y) = f(x,y)/g(x,y)$.

In the example, $S_0 \cup S_1$ is an open subset of the closed subset $\mathrm{Z}_L(Y^2 - X^3)$ of L^2. For now, we shall look just at open subsets of the powers L^n themselves.

If $\mathfrak{p}$ is a prime ideal of $K[\boldsymbol{X}]$, and f and g in $K[\boldsymbol{X}]$ are such that $\boldsymbol{x} \mapsto f(\boldsymbol{x})/g(\boldsymbol{x})$ is well-defined (and therefore regular) on $L^n \smallsetminus \mathrm{Z}_L(\mathfrak{p})$, this means f/g is a well-defined element of the local ring $K[\boldsymbol{X}]_{\mathfrak{p}}$.

Now write $R = K[\boldsymbol{X}]$ as before, and let $\mathfrak{a}$ be an arbitrary radical ideal of R, so that $\mathrm{U}_{\mathfrak{a}}$ is an open subset of $\operatorname{Spec}(R)$. We define shall define a sub-ring, to be denoted by

$$\mathscr{O}(\mathrm{U}_{\mathfrak{a}}),$$

of the product[2]

$$\prod_{\mathfrak{p}\in\mathrm{U}_{\mathfrak{a}}} R_{\mathfrak{p}}.$$

See Figure 11.1. Elements of this product are functions on $\mathrm{U}_{\mathfrak{a}}$; so as before we have a notion of being *regular*: An element h of the product is **regular** at a point $\mathfrak{p}$ of $\mathrm{U}_{\mathfrak{a}}$ if, for some open subset V of $\mathrm{U}_{\mathfrak{a}}$ that contains $\mathfrak{p}$, there are f and g in R such that, for all $\mathfrak{q}$ in V,

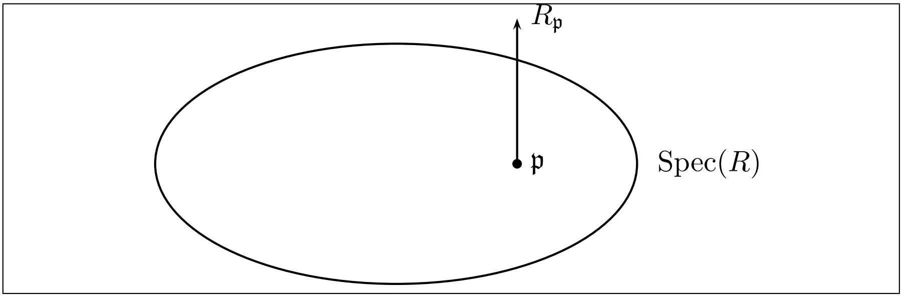

Figure 11.1.: A stalk of a sheaf (see p. 249)

$$h_{\mathfrak{q}} = \frac{f}{g}.$$

Note that this requires $g \notin \mathfrak{q}$. The ring $\mathscr{O}(\mathrm{U}_{\mathfrak{a}})$ consists of the elements of $\prod_{\mathfrak{p}\in\mathrm{U}_{\mathfrak{a}}} R_{\mathfrak{p}}$ that are regular at all points of $\mathrm{U}_{\mathfrak{a}}$.

There is a simpler definition when $\mathfrak{a}$ is a principal ideal (g). In this case, one shows

$$\mathscr{O}(\mathrm{U}_{(g)}) \cong \{g^k : k \in \omega\}^{-1}R,$$

because the map $x/g^n \mapsto (x/g^n : \mathfrak{p} \in \mathrm{U}_{(g)})$ from this ring to $\mathscr{O}(\mathrm{U}_{(g)})$ is injective and surjective. See Hartshorne [26, Prop. II.2.2, p. 71].

[2] Note well that the factors of the product are the localizations $R_{\mathfrak{p}}$, rather than, say, the quotient-fields of the quotients $R/\mathfrak{p}$. However, in the other case that we shall be interested in, where R is itself a product of fields, then the integral domains $R/\mathfrak{p}$ will already be fields, which are isomorphic to the localizations $R_{\mathfrak{p}}$. See §4.6.

If U and V are open subsets of R such that $U \supseteq V$, then the restriction-map from $\prod_{\mathfrak{p}\in U} R_{\mathfrak{p}}$ to $\prod_{\mathfrak{p}\in V} R_{\mathfrak{p}}$ itself restricts to a map ρ^U_V from $\mathscr{O}(U)$ to $\mathscr{O}(V)$. If $h \in \mathscr{O}(U)$, we then write

$$\rho^U_V(h) = h \restriction V.$$

The function $U \mapsto \mathscr{O}(U)$ on the collection of open subsets of R, together with these homomorphisms ρ_{UV}, is called a **pre-sheaf** of rings on Spec(R) because:

$$\mathscr{O}(\varnothing) = \{0\}, \qquad \rho^U_U = \mathrm{id}_U, \qquad \rho^U_W = \rho^V_W \circ \rho^U_V.$$

(The notation ρ^U_V implies $U \supseteq V$; so for the last equation we have $U \supseteq V \supseteq W$.) We now have a situation that is 'dual' (because the arrows are reversed) to that of the Galois group Gal$(\mathbb{F}_p)$: see page 235. For all $\mathfrak{p}$ in Spec(R), $R_{\mathfrak{p}}$ has a certain 'universal property' with respect to the system of rings $\mathscr{O}(U)$ such that $\mathfrak{p} \in U$:

1. $R_{\mathfrak{p}}$ is a ring A to which there is a homomorphism h^U_A from $\mathscr{O}(U)$ for such that, if $U \supseteq V$, then
$$h^V_A \circ \rho^U_V = h^U_A.$$
2. For every such ring A, there is a unique homomorphism h to A from $R_{\mathfrak{p}}$ such that
$$h^U_A = h \circ h^U_{R_{\mathfrak{p}}}.$$

See Figure 11.2. Therefore $R_{\mathfrak{p}}$ is called a **co-limit** or **direct limit** of the given system of rings. This limit can be obtained as a quotient of the sum $\sum_{\mathfrak{p}\in U} \mathscr{O}(U)$ by the smallest ideal that contains, for each pair U and V such that $U \supset V$, every element x such that $x_V = \rho^U_V(x_U)$, and $x_W = 0$ if W is not U or V.

The pre-sheaf $U \mapsto \mathscr{O}(U)$ is further a **sheaf** of rings because it has two additional properties:

1. If $h \in \mathscr{O}(U)$, and $h \restriction V = 0$ for all V in an open covering of U, then $h = 0$.

$$\begin{tikzcd}
 & R_{\mathfrak{p}} \arrow[d, "h"] & \\
 & A & \\
\mathscr{O}(V) \arrow[uur, "h^V_{R_{\mathfrak{p}}}"] \arrow[ur, "h^V_A"] & & \mathscr{O}(U) \arrow[uul, "h^U_{R_{\mathfrak{p}}}"'] \arrow[ul, "h^U_A"'] \arrow[ll, "\rho^U_V"]
\end{tikzcd}$$

Figure 11.2.: The universal property of $R_{\mathfrak{p}}$

2. If there is h_V in $\mathscr{O}(V)$ for every V in an open covering of U, and

$$h_V \restriction (V \cap W) = h_W \restriction (V \cap W)$$

for all V and W in this open covering, then for some h in $\mathscr{O}(U)$, for each V in the open covering,

$$h_V = h \restriction V.$$

The local ring $R_{\mathfrak{p}}$ is the **stalk** of the sheaf at $\mathfrak{p}$. In the fullest sense, the **spectrum** of R is Spec(R) as a topological space equipped with this sheaf. The sheaf is then the **structure sheaf** of the spectrum of R.

11.3. Generic points and irreducibility

This section is here for completeness, but will not be used later. Every point $\boldsymbol{a}$ of L^n is called a **generic point** of $\mathrm{Z}_L(\mathrm{I}_K(\boldsymbol{a}))$; more precisely, $\boldsymbol{a}$ is a generic point *over* K of $\mathrm{Z}_L(\mathrm{I}_K(\boldsymbol{a}))$. In the example on page 220, (π, π) and (e, e) are generic points of $\mathrm{Z}_L(X - Y)$ over $\mathbb{Q}$.

In any case, if for some radical ideal $\mathfrak{a}$, the algebraic set $\mathrm{Z}_L(\mathfrak{a})$ has a generic point, then $\mathfrak{a}$ must be prime. The converse may fail. For example, $\mathrm{Z}_L((X - Y))$ has no generic point if $L \subseteq \mathbb{Q}^{\mathrm{alg}}$. However, to Theorem 193, we have the following

Corollary 201.2. *If $K(\boldsymbol{X})^{\mathrm{alg}} \subseteq L$, then the zero-locus in L of every prime ideal has a generic point.*

A closed subset of L^n is called **irreducible** if it cannot be written as the union of two closed subsets, neither of which includes the other.

Theorem 202. *For all radical ideals $\mathfrak{a}$ of $K[\boldsymbol{X}]$, if $K^{\mathrm{alg}} \subseteq L$,*

$$\mathfrak{a} \textit{ is prime} \iff \mathrm{Z}_L(\mathfrak{a}) \textit{ is irreducible.}$$

Proof. If $\mathfrak{p}$ is prime, and $\mathrm{Z}_L(\mathfrak{p}) = \mathrm{Z}_L(\mathfrak{a}) \cup \mathrm{Z}_L(\mathfrak{b})$ for some radical ideals $\mathfrak{a}$ and $\mathfrak{b}$, then (by Hilbert's Nullstellensatz)

$$\mathfrak{p} = \mathfrak{a} \cap \mathfrak{b},$$

so we may assume $\mathfrak{p} = \mathfrak{a}$ and therefore $\mathrm{Z}_L(\mathfrak{a}) \supseteq \mathrm{Z}_L(\mathfrak{b})$.

Suppose conversely $\mathrm{Z}_L(\mathfrak{a})$ is irreducible, and $fg \in \mathfrak{a}$. Then

$$\mathrm{Z}_L(\mathfrak{a}) = \mathrm{Z}_L(\mathfrak{a} \cup \{f\}) \cup \mathrm{Z}_L(\mathfrak{a} \cup \{g\}),$$

so we may assume $\mathrm{Z}_L(\mathfrak{a}) = \mathrm{Z}_L(\mathfrak{a} \cup \{f\})$ and therefore (again by Hilbert's Nullstellensatz) $f \in \mathfrak{a}$. □

For example, L^n itself is irreducible, since the zero-ideal of $K[\boldsymbol{X}]$ is prime. Therefore the closure of every open subset is the whole space L^n. In any case, every closed set is the union of only finitely many irreducible closed sets: this is by the corollary to the Hilbert Basis Theorem (Theorem 196 on page 222). Hence every radical ideal of $K[\boldsymbol{X}]$ is the intersection of just finitely many elements of $\mathrm{Spec}(K[\boldsymbol{X}])$.

11.4. Affine schemes

For an arbitrary commutative ring R, every element f of R determines a function $\mathfrak{p} \mapsto f_{\mathfrak{p}}$ on $\operatorname{Spec}(R)$, where $f_{\mathfrak{p}}$ is the element $f + \mathfrak{p}$ of $R/\mathfrak{p}$. However, the corresponding map

$$f \mapsto (f_{\mathfrak{p}} \colon \mathfrak{p} \in \operatorname{Spec}(R)) \tag{11.2}$$

from R to $\prod_{\mathfrak{p} \in \operatorname{Spec}(R)} R/\mathfrak{p}$ is injective if and only if R is reduced (Theorem 117, page 117). For example, the kernel of this map contains $X + (X^2)$ when $R = K[X]/(X^2)$. In general, the kernel is $\sqrt{\{0\}}$.

We may refer to the topology on $\operatorname{Spec}(R)$ as the **Zariski topology.** Just as before, we obtain the sheaf $U \mapsto \mathscr{O}(U)$ of rings on $\operatorname{Spec}(R)$, with stalks $R_{\mathfrak{p}}$. Continuing the example on page 246, we may let

$$R = K[X, Y]/(Y^2 - X^3).$$

Let x and y be the images of X and Y respectively in R. Then

$$\mathrm{U}_{(x,y)} = \mathrm{U}_x \cup \mathrm{U}_y,$$

and

$$\mathfrak{p} \in \mathrm{U}_x \implies \frac{y}{x^2} \in R_{\mathfrak{p}}, \qquad \mathfrak{p} \in \mathrm{U}_y \implies \frac{x}{y} \in R_{\mathfrak{p}},$$

and if $\mathfrak{p} \in \mathrm{U}_x \cap \mathrm{U}_y$, then y/x^2 and x/y are the same element of $R_{\mathfrak{p}}$. Thus we obtain an element of $\mathscr{O}(\mathrm{U}_{(x,y)})$.

The spectrum of a ring is called an **affine scheme.** One point of introducing this terminology is that a *scheme,* simply, is a topological space with a sheaf of rings such that such that every point of the space has a neighborhood that, with the restriction of the sheaf to it, is an affine scheme. However, we shall not look at schemes in general. In fact we shall look at just one affine scheme whose underlying ring is not a polynomial ring.

11.5. The ultraproduct scheme

Now let R be the product $\prod_{i\in\Omega} K_i$ of fields as above. As $\mathfrak{p}$ ranges over $\operatorname{Spec}(R)$, the quotients $R/\mathfrak{p}$ are just the possible ultraproducts of the fields K_i. We want to investigate how these arise from the structure sheaf of the spectrum of R. So, letting $\mathfrak{a}$ be an ideal of R, we want to understand $\mathscr{O}(\mathrm{U}_{\mathfrak{a}})$.

We can identify $\operatorname{Spec}(R)$ with $\operatorname{Spec}(\mathscr{P}(\Omega))$, and more generally, we can identify ideals of R with ideals of $\mathscr{P}(\Omega)$. Because $R_{\mathfrak{p}} \cong R/\mathfrak{p}$, we may assume

$$\mathscr{O}(\mathrm{U}_{\mathfrak{a}}) \subseteq \prod_{\mathfrak{p}\in\mathrm{U}_{\mathfrak{a}}} R/\mathfrak{p}.$$

Here we may treat $\mathfrak{a}$ as an ideal of $\mathscr{P}(\Omega)$, so $\mathrm{U}_{\mathfrak{a}}$ can be thought of as an open subset of $\operatorname{Spec}(\mathscr{P}(\Omega))$. Then, in the product $\prod_{\mathfrak{p}\in\mathrm{U}_{\mathfrak{a}}} R/\mathfrak{p}$, the index $\mathfrak{p}$ ranges over this open subset, but in the quotient $R/\mathfrak{p}$, the index returns to being the corresponding ideal of R.

Let $s \in \prod_{\mathfrak{p}\in\mathrm{U}_{\mathfrak{a}}} R/\mathfrak{p}$. Every *principal* ideal in $\mathrm{U}_{\mathfrak{a}}$ is $(\Omega\smallsetminus\{i\})$ for some i in Ω. In this case we have $\mathfrak{a} \nsubseteq (\Omega\smallsetminus\{i\})$, that is,

$$i \in \bigcup \mathfrak{a}.$$

Let us denote $(\Omega\smallsetminus\{i\})$ by $\mathfrak{p}(i)$. There is only one prime ideal of $\mathscr{P}(\Omega)$ that does not contain $\{i\}$, namely $\mathfrak{p}(i)$. Thus

$$\mathrm{U}_{\{i\}} = \{\mathfrak{p}(i)\}.$$

In particular, s is automatically regular at $\mathfrak{p}(i)$. We want to understand when s is regular at not-principal ideals.

Still considering also the principal ideals, we have

$$R/\mathfrak{p}(i) \cong K_i.$$

Let $s_{\mathfrak{p}(i)}$ be sent to s_i under this isomorphism, so whenever x in R is such that $x_i = s_i$, we have

$$s_{\mathfrak{p}(i)} = x + \mathfrak{p}(i).$$

By definition, we have $s \in \mathscr{O}(\mathrm{U}_{\mathfrak{a}})$ if and only if, for all $\mathfrak{p}$ in $\mathrm{U}_{\mathfrak{a}}$, for some subset $\mathrm{U}_{\mathfrak{b}}$ of $\mathfrak{a}$ such that $\mathfrak{b} \nsubseteq \mathfrak{p}$, for some x in R, for all $\mathfrak{q}$ in $\mathrm{U}_{\mathfrak{b}}$,

$$s_{\mathfrak{q}} = x + \mathfrak{q}.$$

We may assume $\mathfrak{b}$ is a principal ideal (A), where $A \in \mathfrak{a} \smallsetminus \mathfrak{p}$. If $\mathfrak{q}$ in U_A here is the principal ideal $\mathfrak{p}(j)$, so that $j \in A$, we must have $x_j = s_j$. More generally, $\mathfrak{q} \in \mathrm{U}_A$ means $A \notin \mathfrak{q}$, so A is $\mathfrak{q}$-large, and hence for all x in R, $x + \mathfrak{q}$ is determined by $(x_i \colon i \in A)$. Thus we may assume

$$x = (s_i \colon i \in \Omega).$$

This establishes that $\mathscr{O}(\mathrm{U}_{\mathfrak{a}})$ is the image of R in $\prod_{\mathfrak{p} \in \mathrm{U}_{\mathfrak{a}}} R/\mathfrak{p}$:

$$\mathscr{O}(\mathrm{U}_{\mathfrak{a}}) = \{(x + \mathfrak{p} \colon \mathfrak{p} \in \mathrm{U}_{\mathfrak{a}}) \colon x \in R\}.$$

In particular, $\mathscr{O}(\mathrm{U}_{\mathfrak{a}})$ is a quotient of R, that is, a reduced product of the K_i. More precisely,

$$\mathscr{O}(\mathrm{U}_{\mathfrak{a}}) \cong R/\mathfrak{b},$$

where

$$\mathfrak{b} = \bigcap_{\mathfrak{p} \in \mathrm{U}_{\mathfrak{a}}} \mathfrak{p} = \bigcap_{\mathfrak{a} \nsubseteq \mathfrak{p}} \mathfrak{p}.$$

It follows that

$$\mathscr{O}(\mathrm{U}_{\mathfrak{a}}) \cong \prod_{i \in \bigcup \mathfrak{a}} K_i. \tag{11.3}$$

We can see this in two ways. For example, if $\mathfrak{p} \in \mathrm{U}_{\mathfrak{a}}$, so that $\mathfrak{a} \nsubseteq \mathfrak{p}$, then $\bigcup \mathfrak{a} \notin \mathfrak{p}$, that is, $\bigcup \mathfrak{a}$ is $\mathfrak{p}$-large. Therefore the image of x in $\mathscr{O}(\mathrm{U}_{\mathfrak{a}})$ depends only on $(x_i \colon i \in \bigcup \mathfrak{a})$. This shows that $\mathscr{O}(\mathrm{U}_{\mathfrak{a}})$ is a quotient of $\prod_{i \in \bigcup \mathfrak{a}} K_i$.

It is moreover the quotient by the trivial ideal. For, if $i \in \bigcup \mathfrak{a}$, then $\mathfrak{p}(i) \in \mathrm{U}_{\mathfrak{a}}$, so that $x + \mathfrak{p}(i)$ depends only on x_i, that is,

$$x + \mathfrak{p}(i) = 0 \iff x_i = 0.$$

This gives us (11.3).

Note that possibly $\bigcup\mathfrak{a}\notin\mathfrak{p}$, although $\mathfrak{a}\subseteq\mathfrak{p}$. Such is the case when $\mathfrak{p}$ is non-principal, but $\mathfrak{a}$ is the ideal of finite sets. However, we always have

$$\bigcup\mathfrak{a}\notin\mathfrak{p}\implies(\bigcup\mathfrak{a})\nsubseteq\mathfrak{p}.$$

Another way to establish (11.3) is to show

$$\bigcap_{\mathfrak{a}\nsubseteq\mathfrak{p}}\mathfrak{p}=(\Omega\smallsetminus\bigcup\mathfrak{a}).$$

If $X\subseteq\Omega\smallsetminus\bigcup\mathfrak{a}$, and $\mathfrak{a}\nsubseteq\mathfrak{p}$, then $\bigcup\mathfrak{a}\notin\mathfrak{p}$, so $\Omega\smallsetminus\bigcup\mathfrak{a}\in\mathfrak{p}$, and therefore $X\in\mathfrak{p}$. Inversely, if $X\nsubseteq\Omega\smallsetminus\bigcup\mathfrak{a}$, then $X\cap\bigcup\mathfrak{a}$ has an element i, so that $X\notin\mathfrak{p}(i)$ and $\mathfrak{a}\nsubseteq\mathfrak{p}(i)$.

Because the stalk $R_{\mathfrak{p}}$ is always a direct limit of those $\mathscr{O}(U)$ such that $\mathfrak{p}\in U$, we have in the present situation that the ultraproduct $\prod_{i\in\Omega}K_i/\mathfrak{p}$ is a direct limit of those products $\prod_{i\in A}K_i$ such that $A\notin\mathfrak{p}$. Symbolically,

$$\prod_{i\in\Omega}K_i/\mathfrak{p}=\varinjlim\Bigl\{\prod_{i\in A}K_i\colon A\notin\mathfrak{p}\Bigr\}.$$

Equivalently, the ultraproduct is the direct limit of those $R/\mathfrak{a}$ such that $\mathfrak{a}$ is a principal ideal included in $\mathfrak{p}$:

$$\prod_{i\in\Omega}K_i/\mathfrak{p}=\varinjlim\Bigl\{\prod_{i\in\Omega}K_i/(B)\colon B\in\mathfrak{p}\Bigr\}.$$

A. The German script

In his *Model Theory* of 1993, Wilfrid Hodges observes [31, Ch. 1, p. 21]:

> Until about a dozen years ago, most model theorists named structures in horrible Fraktur lettering. Recent writers sometimes adopt a notation according to which all structures are named M, M', M^*, $\bar{M}$, M_0, M_i or occasionally N. I hope I cause no offence by using a more freewheeling notation.

For Hodges, *structures* (such as we define in §2.6 on page 45 above) are denoted by the letters A, B, C, and so forth; Hodges refers to their universes as **domains** and denotes these by $\operatorname{dom}(A)$ and so forth. In his *Model Theory: An Introduction* of 2002, David Marker [46] uses "calligraphic" letters to denote structures, as distinct from their universes: so M is the universe of $\mathcal{M}$, and N of $\mathcal{N}$. I still prefer the older practice of using capital Fraktur letters for structures:

$$\begin{array}{ccccccccccccc} \mathfrak{A} & \mathfrak{B} & \mathfrak{C} & \mathfrak{D} & \mathfrak{E} & \mathfrak{F} & \mathfrak{G} & \mathfrak{H} & \mathfrak{I} & \mathfrak{J} & \mathfrak{K} & \mathfrak{L} & \mathfrak{M} \\ \mathfrak{N} & \mathfrak{O} & \mathfrak{P} & \mathfrak{Q} & \mathfrak{R} & \mathfrak{S} & \mathfrak{T} & \mathfrak{U} & \mathfrak{V} & \mathfrak{W} & \mathfrak{X} & \mathfrak{Y} & \mathfrak{Z} \end{array}$$

For the record, here are the minuscule Fraktur letters, which are sometimes used in this text for denoting ideals:

$$\begin{array}{ccccccccccccc} \mathfrak{a} & \mathfrak{b} & \mathfrak{c} & \mathfrak{d} & \mathfrak{e} & \mathfrak{f} & \mathfrak{g} & \mathfrak{h} & \mathfrak{i} & \mathfrak{j} & \mathfrak{k} & \mathfrak{l} & \mathfrak{m} \\ \mathfrak{n} & \mathfrak{o} & \mathfrak{p} & \mathfrak{q} & \mathfrak{r} & \mathfrak{s} & \mathfrak{t} & \mathfrak{u} & \mathfrak{v} & \mathfrak{w} & \mathfrak{x} & \mathfrak{y} & \mathfrak{z} \end{array}$$

A way to write these letters by hand is seen on the page reproduced below from a 1931 textbook [27] on the German language:

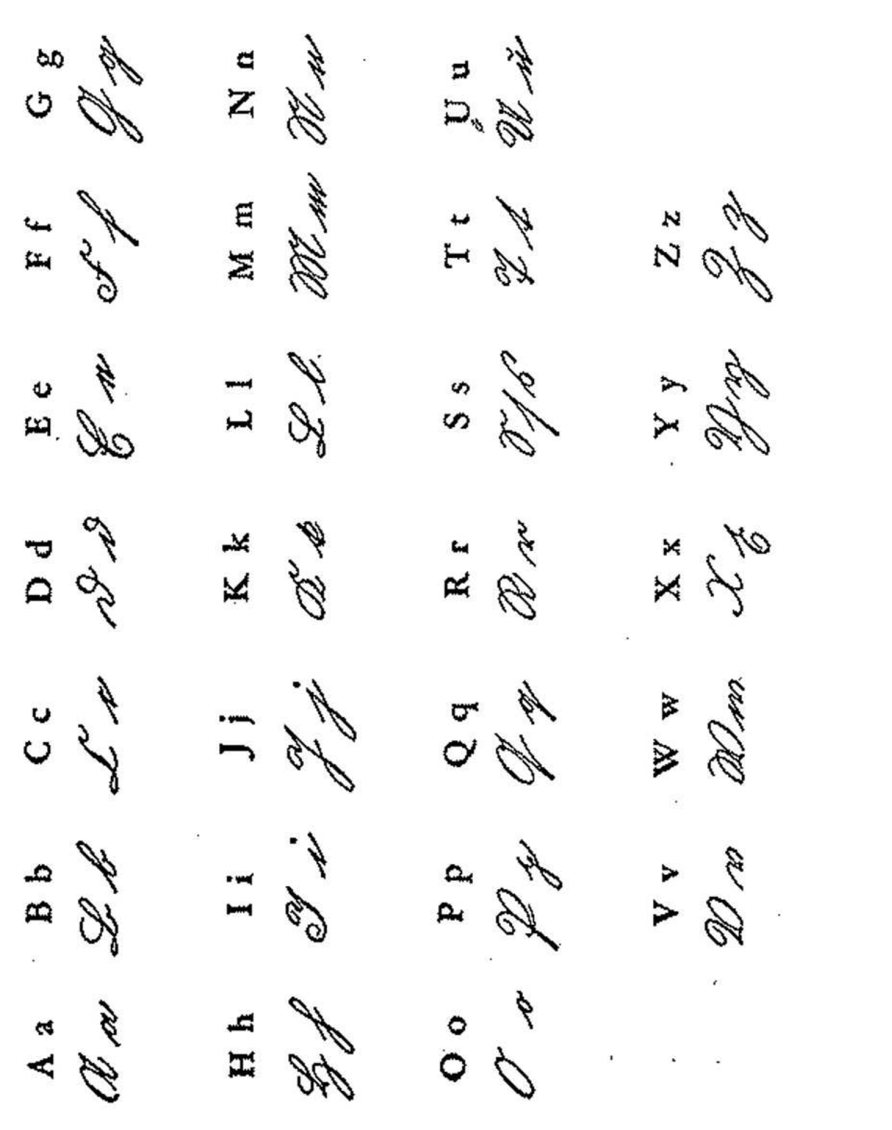

Figure A.1.: The German alphabet

Bibliography

[1] Emil Artin. *Galois theory.* Dover Publications, Inc., Mineola, NY, second edition, 1998. Edited and with a supplemental chapter by Arthur N. Milgram, "Unabridged and unaltered republication of the last corrected printing of the 1944 second, revised edition of the work first published by The University of Notre Dame Press in 1942 as Number 2 in the series, *Notre Dame Mathematical Lectures.*".

[2] James Ax. The elementary theory of finite fields. *Ann. of Math. (2)*, 88:239–271, 1968.

[3] Michael Barr and Charles Wells. *Category theory for computing science.* Prentice Hall International Series in Computer Science. Prentice Hall International, New York, 1990.

[4] J. L. Bell and A. B. Slomson. *Models and ultraproducts: An introduction.* North-Holland Publishing Co., Amsterdam, 1969. reissued by Dover, 2006.

[5] Garrett Birkhoff. *Lattice theory.* Third edition. American Mathematical Society Colloquium Publications, Vol. XXV. American Mathematical Society, Providence, R.I., 1967.

[6] Alexandre Borovik and Mikhael Katz. Inevitability of infinitesimals. `http://manchester.academia.edu/AlexandreBorovik/Papers/305871/Inevitability_of_infinitesimals`. accessed July 18, 2012.

[7] Cesare Burali-Forti. A question on transfinite numbers. In van Heijenoort [58], pages 104–12. First published 1897.

[8] C. C. Chang and H. J. Keisler. *Model theory*, volume 73 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, third edition, 1990.

[9] Zoé Chatzidakis. *Théorie de modèles des corps finis et pseudo-finis*. Prépublications de l'Equipe de Logique. Université Paris VII, Octobre 1996. `http://www.logique.jussieu.fr/~zoe/`.

[10] Alonzo Church. *Introduction to mathematical logic. Vol. I*. Princeton University Press, Princeton, N. J., 1956.

[11] Harvey Cohn. *Advanced Number Theory*. Dover, New York, 1980. Corrected republication of 1962 edition.

[12] Kevin R. Coombes. Agathos: Algebraic geometry: A total hypertext online system. `http://www.silicovore.com/agathos/contents.html`. accessed July 9, 2014.

[13] Richard Dedekind. *Essays on the theory of numbers. I: Continuity and irrational numbers. II: The nature and meaning of numbers*. authorized translation by Wooster Woodruff Beman. Dover Publications Inc., New York, 1963.

[14] René Descartes. *The Geometry of René Descartes*. Dover Publications, Inc., New York, 1954. Translated from the French and Latin by David Eugene Smith and Marcia L. Latham, with a facsimile of the first edition of 1637.

[15] Apostolos Doxiadis and Christos H. Papadimitriou. *Logicomix*. Bloomsbury, London, 2009.

[16] David Eisenbud. *Commutative algebra*, volume 150 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. With a view toward algebraic geometry.

[17] Euclid. *The thirteen books of Euclid's Elements translated from the text of Heiberg. Vol. I: Introduction and Books I, II. Vol. II:*

Books III–IX. Vol. III: Books X–XIII and Appendix. Dover Publications Inc., New York, 1956. Translated with introduction and commentary by Thomas L. Heath, 2nd ed.

[18] Euclid. *Euclid's Elements.* Green Lion Press, Santa Fe, NM, 2002. All thirteen books complete in one volume. The Thomas L. Heath translation, edited by Dana Densmore.

[19] T. Frayne, A. C. Morel, and D. S. Scott. Reduced direct products. *Fund. Math.*, 51:195–228, 1962/1963.

[20] T. Frayne, A. C. Morel, and D. S. Scott. Correction to the paper "Reduced direct products". *Fund. Math.*, 53:117, 1963.

[21] Michael D. Fried and Moshe Jarden. *Field arithmetic*, volume 11 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)].* Springer-Verlag, Berlin, 1986.

[22] Carl Friedrich Gauss. *Disquisitiones Arithmeticae.* Springer-Verlag, New York, 1986. Translated into English by Arthur A. Clarke, revised by William C. Waterhouse.

[23] Carolo Friderico Gauss. *Disquisitiones Arithmeticae.* Gerh. Fleischer Jun., Lipsiae, 1801. Electronic version of the original Latin text from Goettingen State and University Library.

[24] Kurt Gödel. The completeness of the axioms of the functional calculus of logic. In van Heijenoort [58], pages 582–91. First published 1930.

[25] K. R. Goodearl. *von Neumann regular rings*, volume 4 of *Monographs and Studies in Mathematics.* Pitman (Advanced Publishing Program), Boston, Mass., 1979.

[26] Robin Hartshorne. *Algebraic geometry.* Springer-Verlag, New York, 1977. Graduate Texts in Mathematics, No. 52.

[27] Roe-Merrill S. Heffner. *Brief German Grammar.* D. C. Heath and Company, Boston, 1931.

[28] Leon Henkin. The completeness of the first-order functional calculus. *J. Symbolic Logic*, 14:159–166, 1949.

[29] Leon Henkin. On mathematical induction. *Amer. Math. Monthly*, 67:323–338, 1960.

[30] Leon Henkin. The discovery of my completeness proofs. *Bull. Symbolic Logic*, 2(2):127–158, 1996.

[31] Wilfrid Hodges. *Model theory*, volume 42 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 1993.

[32] Wilfrid Hodges. *Building models by games.* Dover Publications, Mineola, New York, 2006. Original publication, 1985.

[33] Paul E. Howard. Łoś' theorem and the Boolean prime ideal theorem imply the axiom of choice. *Proc. Amer. Math. Soc.*, 49:426–428, 1975.

[34] Thomas W. Hungerford. *Algebra*, volume 73 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1980. Reprint of the 1974 original.

[35] Edward V. Huntington. Errata: "Sets of independent postulates for the algebra of logic" [Trans. Amer. Math. Soc. **5** (1904), no. 3, 288–309; 1500675]. *Trans. Amer. Math. Soc.*, 5(4):552, 1904.

[36] Edward V. Huntington. Sets of independent postulates for the algebra of logic. *Trans. Amer. Math. Soc.*, 5(3):288–309, 1904.

[37] J. L. Kelley. The Tychonoff product theorem implies the axiom of choice. *Fund. Math.*, 37:75–76, 1950.

[38] Kenneth Kunen. *Set theory*, volume 102 of *Studies in Logic and the Foundations of Mathematics.* North-Holland Publishing

Co., Amsterdam, 1983. An introduction to independence proofs, Reprint of the 1980 original.

[39] Casimir Kuratowski. Une méthode d'élimination des nombres transfinis des raisonnements mathématiques. *Fundamenta Mathematicae*, 3(1):76–108, 1922.

[40] Edmund Landau. *Foundations of Analysis. The Arithmetic of Whole, Rational, Irrational and Complex Numbers.* Chelsea Publishing Company, New York, N.Y., third edition, 1966. Translated by F. Steinhardt; first edition 1951; first German publication, 1929.

[41] Serge Lang. *Algebra.* Addison-Wesley, Reading, Massachusetts, third edition, 1993. Reprinted with corrections, 1997.

[42] J. Łoś and C. Ryll-Nardzewski. On the application of Tychonoff's theorem in mathematical proofs. *Fund. Math.*, 38:233–237, 1951.

[43] J. Łoś and C. Ryll-Nardzewski. Effectiveness of the representation theory for Boolean algebras. *Fund. Math.*, 41:49–56, 1954.

[44] Jerzy Łoś. Quelques remarques, théorèmes et problèmes sur les classes définissables d'algèbres. In *Mathematical interpretation of formal systems*, pages 98–113. North-Holland Publishing Co., Amsterdam, 1955.

[45] Leopold Löwenheim. On possibilities in the calculus of relatives. In van Heijenoort [58], pages 228–251. First published 1915.

[46] David Marker. *Model theory: an introduction*, volume 217 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2002.

[47] Oystein Ore. Galois connexions. *Trans. Amer. Math. Soc.*, 55:493–513, 1944.

[48] Giuseppe Peano. The principles of arithmetic, presented by a new method. In van Heijenoort [58], pages 83–97. First published 1889.

[49] Emil L. Post. Introduction to a general theory of elementary propositions. *Amer. J. Math.*, 43(3):163–185, July 1921.

[50] Herman Rubin and Jean E. Rubin. *Equivalents of the axiom of choice. II*, volume 116 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, 1985.

[51] Bertrand Russell. Letter to Frege. In van Heijenoort [58], pages 124–5. First published 1902.

[52] Eric Schechter. Kelley's specialization of Tychonoff's theorem is equivalent to the Boolean prime ideal theorem. *Fund. Math.*, 189(3):285–288, 2006.

[53] Dana Scott. Prime ideal theorems for rings, lattices, and boolean algebras. *Bull. Amer. Math. Soc.*, 60(4):390, July 1954. Preliminary report.

[54] Joseph R. Shoenfield. *Mathematical logic*. Association for Symbolic Logic, Urbana, IL, 2001. reprint of the 1973 second printing.

[55] Thoralf Skolem. Some remarks on axiomatized set theory. In van Heijenoort [58], pages 290–301. First published 1922.

[56] Thoralf Skolem. Logico-combinatorial investigations in the satisfiability or provability of mathematical propositions: A simplified proof of a theorem by L. Löwenheim and generalizations of the theorem. In van Heijenoort [58], pages 252–63. First published 1920.

[57] M. H. Stone. The theory of representations for Boolean algebras. *Trans. Amer. Math. Soc.*, 40(1):37–111, 1936.

[58] Jean van Heijenoort, editor. *From Frege to Gödel: A source book in mathematical logic, 1879–1931*. Harvard University Press, Cambridge, MA, 2002.

[59] John von Neumann. An axiomatization of set theory. In van Heijenoort [58], pages 393–413. First published 1925.

[60] John von Neumann. On the introduction of transfinite numbers. In van Heijenoort [58], pages 346–354. First published 1923.

[61] Alfred North Whitehead and Bertrand Russell. *Principia Mathematica*, volume I. University Press, Cambridge, 1910.

[62] Stephen Willard. *General topology*. Addison-Wesley Publishing Co., Reading, Mass.–London–Don Mills, Ont., 1970.

[63] Ernst Zermelo. Investigations in the foundations of set theory I. In van Heijenoort [58], pages 199–215. First published 1908.

[64] Max Zorn. A remark on method in transfinite algebra. *Bull. Amer. Math. Soc.*, 41(10):667–670, 1935.