Lorenz Halbeisen
Regula Krapf

Gödel's Theorems and Zermelo's Axioms

A Firm Foundation of Mathematics

Second Edition

Birkhäuser

Birkhäuser

Lorenz Halbeisen • Regula Krapf

Gödel's Theorems and Zermelo's Axioms

A Firm Foundation of Mathematics

Second Edition

Birkhäuser

Lorenz Halbeisen
Department of Mathematics
ETH Zurich
Zurich, Switzerland

Regula Krapf
Mathematics Institute
Universität Bonn
Bonn, Nordrhein-Westfalen, Germany

ISBN 978-3-031-85105-6 ISBN 978-3-031-85106-3 (eBook)
https://doi.org/10.1007/978-3-031-85106-3

Mathematics Subject Classification (2020): 03Exx, 03C35, 03C62, 11B25, 03B30, 03-xx, 03E30, 03C10, 03H15

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2020, 2025

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This book is published under the imprint Birkhäuser, www.birkhauser-science.com by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

If disposing of this product, please recycle the paper.

Preface

This book provides a self-contained introduction to the foundations of mathematics, where self-contained means that we assume as little prerequisites as possible. One such assumption is the notion of *finiteness*, which *cannot* be defined in mathematics.

The firm foundation of mathematics we provide is based on logic and models. In particular, it is based on Hilbert's axiomatisation of formal logic (including the notion of formal proofs), and on the notion of models of mathematical theories. On this basis, we first prove Gödel's Completeness Theorem and Gödel's Incompleteness Theorems, and then we introduce Zermelo's Axioms of Set Theory. On the one hand, Gödel's Theorems set the framework within which mathematics takes place. On the other hand, using the example of Analysis, we shall see how mathematics can be developed in a model of Set Theory. So, Gödel's Theorems and Zermelo's Axioms are indeed a firm foundation of mathematics.

The book consists of four parts. The first part is an introduction to First-Order Logic from scratch. Starting with a set of symbols, the basic concepts of formal proofs and models are developed, where special care is given to the notion of *finiteness*.

The second part is concerned with Gödel's Completeness Theorem. Our proof follows Henkin's construction [23]. However, we modified Henkin's construction in order to work just with *potentially infinite* sets and to avoid the use of *actually infinite* sets. Even though Henkin's construction works also for uncountable signatures, we prove the general Completeness Theorem with an ultraproduct construction, using Łoš's Theorem.

After a preliminary chapter on countable models of Peano Arithmetic, the third part is mainly concerned with Gödel's Incompleteness Theorems, which will be proved from scratch (i.e., purely from the axioms of Peano Arithmetic) without any use of Recursion Theory. In Chapter 10 and 12 some weaker theories of arithmetic are investigated. In particular, in Chapter 10 it is shown that Gödel's First Incompleteness Theorem also applies for

Robinson Arithmetic and in Chapter 12 it is shown that *Presburger Arithmetic* is complete.

In the last part we present first Zermelo's axioms of Set Theory, including the Axiom of Choice. Then we discuss the consistency of this axiomatic system and provide standard and non-standard models of Set Theory (including Gödel's model **L**). After introducing the construction of models with ultraproducts, we prove the Completeness Theorem for uncountable signatures as well the the Löwenheim-Skolem Theorems. In the last two chapters, we construct several standard and non-standard models of Peano Arithmetic and of the real numbers, and give a brief introduction to Non-Standard Analysis.

Zürich and Koblenz, 28 April 2020 *L. Halbeisen and R. Krapf*

To the second edition.

Besides many corrections and additions, this second edition now contains detailed solutions to all exercises.

Zürich and Bonn, 27 July 2024 *L. Halbeisen and R. Krapf*

Acknowledgement to the First Edition

We would like to thank all our colleagues and students for many fruitful and inspiring discussions. In particular, we would like to thank Johann Birnick, Marius Furter, Adony Ghebressilasie, Marc Lischka, Daniel Paunovic, Philipp Provenzano, Michele Reho, Matthias Roshardt, Joel Schmitz, and especially Michael Yan for their careful reading and all their remarks and hints.

Acknowledgement to the Second Edition

First of all, we would like to thank Lukas Keller and Quirin Reding for writing the solutions to essentially all exercises of Part I–III. Furthermore, we would like to thank Joseph Federico Arangath, Jeremy Feusi, Joscha Gillessen, Jannik Kochert, Robert Schweizer, and especially Ivan Baburin and Mikhail Zaytsev for their numerous corrections and comments, and Norbert Hungerbühler for the solution to the very last exercise.

Contents

Chapter 0
A Framework for Metamathematics

In the late 19th and early 20th century, several unsuccessful attempts were made to develop the natural numbers from logic. The most promising approaches were the ones due to Frege and Russell, but also their approaches failed at the end. Even though it seems impossible to develop the natural numbers just from logic, it is still necessary to formalise them.

In fact, the problem with the natural numbers is, that we need the notion of finiteness in order to define them. This presupposes the existence of a kind of infinite list of objects, and it is not clear whether these objects are — in some sense — not already the natural numbers which we would like to define.

However, in our opinion there is a subtle distinction between the infinite set of natural numbers and an arbitrarily long list of objects, since the set of natural numbers is an *actually infinite* set, whereas an arbitrarily long list is just *potentially infinite*. The difference between these two types of infinity is, that the actual infinity is something which is completed and definite and consists of infinitely many elements. On the other hand, the potential infinity — introduced by Aristotle — is something that is always finite, even though more and more elements can be added to make it arbitrarily large. For example, the set of prime numbers can be considered as an actually infinite set (as Cantor did), or just as a potentially infinite list of numbers without last element which is never completed (as Euclid did).

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_1

As mentioned above, it seems that there is no way to define the natural numbers just from logic. Hence, if we would like to define them, we have to make some assumptions which cannot be formalised within logic or mathematics in general. In other words, in order to define the natural numbers we have to presuppose some *metamathematical* notions like, for example, the notion of FINITENESS. To emphasise this fact, we shall use a wider letter spacing for the metamathematical notions we suppose.

The combination of all metamathematical assumptions we take, forms the so-called *metatheory*, that is then implicitly used to carry out logical arguments. Metatheories form an essential part of logic, because describing the machinery behind any syntactic rule necessarily requires semantic explanation. While there are many canonical ways of choosing a certain metatheory to work with, there are two properties which all of them have in common:

1. A metatheory has to be informal by design — otherwise, if we decide to formalise a syntactic metatheory, we will in turn require a "metametatheory" to achieve that, which in term has to be formalised, requiring this process to be repeated ad infinitum. To prevent this from happening, formal logic always has to rely on some inherently non-formal foundation.
2. A metatheory has to capture some inherently *physical* properties about the world we exist in. Indeed, the standard procedure for creating mathematical proofs requires that we can physically write them down in finite time. While this approach seems very natural, it forces us to define FINITE objects as those that satisfy a certain material property inside this universe. This way FINITENESS becomes something inherently physical and needs to be included inside the metatheory per se, without a formal mathematical definition.

We remark that the ideal metatheory ought to be inherently uncontroversial, i.e., it should only contain the bare necessities needed to carry out logical arguments, with the remaining load being offset to formal axiomatic systems. This way we do not have to question the "validity" of mathematical results, since those become nothing more than logical consequences derived from axiomatic systems.

So, let us assume that we all have a notion of FINITENESS. Let us further assume that we have two characters, say 0 and **s**. With these characters, we build now the following FINITE strings:

$$0 \quad \mathbf{s}0 \quad \mathbf{ss}0 \quad \mathbf{sss}0 \quad \mathbf{ssss}0 \quad \mathbf{sssss}0 \quad \ldots$$

The three dots on the right of the above expression mean that we always build the next string by appending on the left the character **s** to the string we just built. Proceeding this way, we get in fact a potentially infinite "list" $\mathbb{N}$ of different strings which is never completed.

More formally, we build this potentially infinite list step by step as follows: We start with the empty list, denoted [ ]. Then we append the list [0], which contains just the string 0, to the list [ ], denoted [ ] + [0], and obtain the list [0]. Similarly, we append the list [**s**0] to the list [0], denoted [0] + [**s**0], and obtain the list [0, **s**0]. In this way we obtain arbitrary long FINITE lists of FINITE strings of symbols. This never-ending process leads to the potentially infinite list of so-called **natural numbers**, i.e.,

$$\mathbb{N} = [0,\, \mathbf{s}0\,,\, \mathbf{ss}0\,,\, \mathbf{sss}0\,,\, \mathbf{ssss}0\,,\, \mathbf{sssss}0\,, \ldots$$

For the sake of simplicity, we consider $\mathbb{N}$ as a list, i.e.,

$$\mathbb{N} = [0,\, \mathbf{s}0\,,\, \mathbf{ss}0\,,\, \mathbf{sss}0\,,\, \mathbf{ssss}0\,,\, \mathbf{sssss}0\,, \ldots]$$

even though $\mathbb{N}$ itself, without assuming *actual infinity*, is not a proper list.

For each natural number n in the list $\mathbb{N}$ we have:

$$\text{either} \quad n \equiv 0 \quad \text{or} \quad n \equiv \sigma 0\,,$$

where the symbol $\equiv$ means "identical to" and σ is a non-empty FINITE string of the form $\mathbf{s} \cdots \mathbf{s}$ and hence $\sigma 0$ has the form

$$\underbrace{\mathbf{s} \cdots \mathbf{s}}_{\substack{\text{non-empty}\\ \text{finite string}}} 0\,.$$

If σ and π are both (possibly empty) FINITE strings of the form $\mathbf{s} \cdots \mathbf{s}$, then we write $\sigma\pi$ for the concatenation of σ and π, i.e., for the string obtained by writing first the sequence σ followed by the sequence π.

Remark. For any (possibly empty) strings σ, π, ϱ of the form $\mathbf{s} \cdots \mathbf{s}$ we get

$$\sigma\pi 0 \;\equiv\; \pi\sigma 0\,, \qquad \mathbf{s}\sigma\pi 0 \;\equiv\; \mathbf{s}\pi\sigma 0\,, \qquad \mathbf{s}\sigma\pi 0 \;\equiv\; \sigma\mathbf{s}\pi 0\,,$$

and further we get:

$$\sigma 0 \equiv \pi 0 \;\Longleftrightarrow\; \mathbf{s}\sigma 0 \equiv \mathbf{s}\pi 0$$
$$\sigma 0 \equiv \pi 0 \;\Longleftrightarrow\; \sigma\varrho 0 \equiv \pi\varrho 0$$

If we order FINITE strings of the form

$$\underbrace{\mathbf{s} \cdots \mathbf{s}}_{\substack{\text{possibly empty}\\ \text{finite string}}} 0$$

by their length, we obtain that two strings are identical if and only if they have the same length. From this rather geometric point of view, the facts

above can be deduced from Euclid's *Elements*, where he writes the following statements (see [9, p. 155]):

1. *Things which are equal to the same thing are also equal to one another.*
2. *If equals be added to equals, the wholes are equal.*
3. *If equals be subtracted from equals, the remainders are equal.*
4. *Things which coincide with one another are equal to one another.*

It is convenient to use Hindu-Arabic numerals to denote explicitly given natural numbers (e.g., we write the symbol 1 for $\mathbf{s}0$) and Latin letters like $n, m, \ldots$ for non-specified natural numbers. If n and m denote different natural numbers, where n appears earlier than m in the list $\mathbb{N}$, then we write $n < m$ and the expression $n, \ldots, m$ means the natural numbers which belong to the sublist $[n, \ldots, m]$ of $\mathbb{N}$; if n appears later than m in $\mathbb{N}$, then we write $n > m$ and the expression $n, \ldots, m$ denotes the empty set.

We shall use natural numbers frequently as subscripts for FINITE lists of objects like $t_1, \ldots, t_n$. In this context we mean that for each natural number k in the list $[1, \ldots, n]$, there is an object t_k, where in the case when $n = 0$, the set of objects is empty.

If n is a natural number, then $n + 1$ denotes the natural number $\mathbf{s}n$ (i.e., the number which appears immediately after n in the list $\mathbb{N}$); and if $n \neq 0$, then $n - 1$ denotes the natural number which appears immediately before n in the list $\mathbb{N}$. Furthermore, for $\sigma 0, \pi 0$ in the list $\mathbb{N}$, we define

$$\sigma 0 + 0 :\equiv \sigma 0 \quad \text{and} \quad 0 + \pi 0 :\equiv \pi 0\,,$$

and in general, we define:

$$\sigma 0 + \pi 0 :\equiv \sigma \pi 0$$

Finally, by our construction of natural numbers we get the following fact:

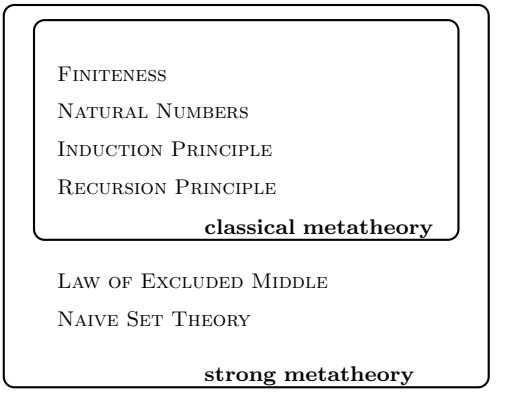
> *If* a statement A holds for 0 and *if* whenever A holds for a natural number n in $\mathbb{N}$ then it also holds for $n + 1$, *then* the statement A holds for *all* natural numbers n in $\mathbb{N}$.

This fact is known as INDUCTION PRICIPLE, which is an important tool in proving statements about natural numbers.

A second principle, which also uses that each natural number is either 0 or the successor of some natural number n in $\mathbb{N}$, is the RECURSION PRINCIPLE:

> *If* we define X_0, and *if* whenever X_n is defined then we can define X_{n+1}, *then* X_n can be defined for *all* natural numbers n in $\mathbb{N}$.

All together, Finiteness, Natural Numbers, the Induction Principle and the Recursion Principle define the *classical metatheory* that we will almost exclusive use to derive metamathematical results. There will be only a handful of exceptions where we strengthen our metatheory by adding the Law of Excluded Middle and some elements from Naive Set Theory, primarily to argue about models and their construction. In this case we are using a *strong metatheory*, as summarised in the drawing below.



Part I
Introduction to First-Order Logic

First-Order Logic is the system of Symbolic Logic concerned not only to represent the logical relations between sentences or propositions as wholes (like *Propositional Logic*), but also to consider their internal structure in terms of subject and predicate. First-Order Logic can be considered as a kind of language which is distinguished from higher-order languages in that it does not allow quantification over sets of elements of the domain or other objects of higher type (like statements of infinite length or statements about formulas). Nevertheless, First-Order Logic is strong enough to formalise all of Set Theory and thereby virtually all of Mathematics.
The goal of this brief introduction to First-Order Logic is to introduce the basic concepts of formal proofs and models, which will be investigated further in Part II and Part III.

Chapter 1
Syntax: The Grammar of Symbols

The goal of this chapter is to develop the formal language of First-Order Logic from scratch. At the same time, we introduce some terminology of the so-called metalanguage, which is the language we use when we speak *about* the formal language (e.g., when we want to express that two strings of symbols are equal). In the metalanguage, we shall use some notions of NAIVE SET THEORY like *sets* (which will always be FINITE), the *membership relation* $\in$, the *empty set* $\emptyset$, or the *subset relation* $\subseteq$. We would like to emphasise that these notions are not part of the language of formal logic and that they are just used in an informal way.

Alphabet

Like any other written language, First-Order Logic is based on an *alphabet*, which consists of the following *symbols*:

(a) **Variables** such as $x, y, v_0, v_1, \ldots$, which are place holders for objects of the *domain* under consideration (which can, e.g., be the elements of a group, natural numbers, or sets). We mainly use lower case Latin letters (with or without subscripts) for variables.

(b) **Logical operators** which are $\neg$ (*not*), $\wedge$ (*and*), $\vee$ (*or*), and $\rightarrow$ (*implies*).

(c) **Logical quantifiers** which are the *existential quantifier* $\exists$ (*there is* or *there exists*) and the *universal quantifier* $\forall$ (*for all* or *for each*), where quantification is restricted to objects only and not to formulae or sets of objects (but the objects themselves may be sets).

(d) **Equality symbol** $=$ which is a special binary *relation symbol* (see below).

(e) **Constant symbols** like the number `0` in Peano Arithmetic, or the neutral element `e` in Group Theory. Constant symbols stand for fixed individual objects in the domain.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_2

(f) **Function symbols** such as $\circ$ (the operation in Group Theory), or $+$, $\cdot$, $\mathtt{s}$ (the operations in Peano Arithmetic). Function symbols stand for fixed functions taking objects as arguments and returning objects as values. With each function symbol we associate a positive natural number, its co-called *arity* (e.g., $\circ$ is a 2-ary or binary function, and the successor operation $\mathtt{s}$ is a 1-ary or unary function). More formally, to each function symbol F we adjoin a fixed FINITE string of place holders $\times\cdots\times$ and write $F\times\cdots\times$.

(g) **Relation symbols** or **predicate constants** (such as $\in$ in Set Theory) stand for fixed relations between (or properties of) objects in the domain. Again, we associate an arity with each relation symbol (e.g., $\in$ is a binary relation). More formally, to each relation symbol R we adjoin a fixed FINITE string of place holders $\times\cdots\times$ and write $R\times\cdots\times$.

The symbols in (a)–(d) form the core of the alphabet and are called **logical symbols**. The symbols in (e)–(g) depend on the specific topic we are investigating and are called **non-logical symbols**. The set of non-logical symbols which are used in order to formalise a certain mathematical theory is called the **signature** (or **language**) of this theory, and *formulae* which are formulated in a language $\mathscr{L}$ are usually called $\mathscr{L}$-formulae. For example, if we investigate groups, then the only non-logical symbols we use are $\mathtt{e}$ and $\circ$, thus $\mathscr{L} = \{\mathtt{e}, \circ\}$ is the language of Group Theory.

Terms & Formulae

With the symbols of our alphabet, we can now start to compose names. In the language of First-Order Logic, these names are called *terms*. Suppose that $\mathscr{L}$ is a signature.

Terms. A string of symbols is an **$\mathscr{L}$-term**, if it results from applying FINITELY many times the following rules:

(T0) Each variable is an $\mathscr{L}$-term.
(T1) Each constant symbol in $\mathscr{L}$ is an $\mathscr{L}$-term.
(T2) If $\tau_1, \ldots, \tau_n$ are any $\mathscr{L}$-terms which we have already built and $F\times\cdots\times$ is an n-ary function symbol in $\mathscr{L}$, then $F\tau_1\cdots\tau_n$ is an $\mathscr{L}$-term (each place holder $\times$ is replaced by an $\mathscr{L}$-term).

When we write general statements which are independent of the signature $\mathscr{L}$, we omit the prefix $\mathscr{L}$ and simply write **term** rather than $\mathscr{L}$-term. Terms of the form (T0) or (T1) are the most basic terms we have, and since every term is built up from such terms, they are called **atomic terms**. In order to define the rule (T2) we had to use variables for terms, but since the variables of our alphabet stand just for objects of the domain and not for terms or other objects of the formal language, we had to introduce new symbols. For these

new symbols, which do not belong to the alphabet of the formal language, we have chosen Greek letters. In fact, we shall mainly use Greek letters for variables which stand for objects of the formal language, also to emphasise the distinction between the formal language and the metalanguage. However, we shall use the Latin letters F and R as variables for function and relation symbols, respectively.

Note that this recursive definition of terms allows us to use the following principle: If we want to prove that all terms satisfy some property Φ, then one has to prove that

- all variables satisfy Φ;
- each constant symbol satisfies Φ;
- if some terms $\tau_1, \ldots, \tau_n$ satisfy Φ, then so does $F\tau_1, \cdots, \tau_n$ for every n-ary function symbol F.

We call this principle **induction on term construction**.

In order to make terms, relations, and other expressions in the formal language easier to read, it is convenient to introduce some more symbols, like brackets and commas, to our alphabet. For example, we usually write $F(\tau_1, \ldots, \tau_n)$ rather than $F\tau_1 \cdots \tau_n$.

To some extent, terms correspond to names, since they denote objects of the domain under consideration. Like real names, they are not statements and cannot express or describe possible relations between objects. So, the next step is to build sentences, or more precisely *formulae*, with these terms.

Formulae. A string of symbols is called an $\mathscr{L}$-**formula**, if it results from applying FINITELY many times the following rules:

(F0) If τ_1 and τ_2 are $\mathscr{L}$-terms, then $= \tau_1\tau_2$ is an $\mathscr{L}$-formula.
(F1) If $\tau_1, \ldots, \tau_n$ are any $\mathscr{L}$-terms and $R\times\cdots\times$ is any non-logical n-ary relation symbol in $\mathscr{L}$, then $R\tau_1 \cdots \tau_n$ is an $\mathscr{L}$-formula.
(F2) If φ is any $\mathscr{L}$-formula which we have already built, then $\neg\varphi$ is an $\mathscr{L}$-formula.
(F3) If φ and ψ are $\mathscr{L}$-formulae which we have already built, then $\wedge\varphi\psi$, $\vee\varphi\psi$, and $\to\varphi\psi$ are $\mathscr{L}$-formulae.
(F4) If φ is an $\mathscr{L}$-formula which we have already built, and ν is an arbitrary variable, then $\exists\nu\varphi$ and $\forall\nu\varphi$ are $\mathscr{L}$-formulae.

As in the case of terms, we usually write simply **formula** rather than $\mathscr{L}$-formula unless the statement in question refers to a specific language. Formulae of the form (F0) or (F1) are the most basic formulae we have, and since every formula is a logical connection or a quantification of these formulae, they are called **atomic formulae**.

In order to make formulae easier to read, we usually use *infix notation* instead of *Polish notation*, and use brackets if necessary. For example, we usually write $\varphi \wedge \psi$ instead of $\wedge\varphi\psi$, $\varphi \to (\psi \to \varphi)$ instead of $\to \varphi \to \psi\varphi$, and $(\varphi \to \psi) \to \varphi$ instead of $\to\to \varphi\psi\varphi$.

In the same way as for terms, a property Φ is satisfied by all formulae if we check the following:

- All atomic formulae satisfy Φ.
- If φ and ψ satisfy Φ and ν is a variable, then so do $\neg\varphi, \varphi \wedge \psi$, $\varphi \vee \psi$, $\varphi \to \psi$, $\exists\nu\varphi$ and $\forall\nu\varphi$.

In accordance with the corresponding principle for terms, we denote this as **induction on formula construction**.

For binary relation symbols $R\times\times$ and binary function symbols $F\times\times$, it is convenient to write xRy and xFy instead of $R(x, y)$ and $F(x, y)$, respectively. For example, we usually write $x = y$ instead of $= xy$.

If a formula φ is of the form $\exists\nu\psi$ or $\forall\nu\psi$ (for some variable ν and some formula ψ) and the variable ν occurs in ψ, but not immediately after a quantifier, then we say that ν is in the *range* of a logical quantifier. Every occurrence of a variable ν in a formula φ, where ν occurs not immediately after a quantifier, is said to be **bound** by the innermost quantifier in whose range it is. If an occurrence of the variable ν at a particular place — not after a quantifier — is not in the range of a quantifier, it is said to be **free** at that particular place. Notice that it is possible that a variable occurs in a given formula at a certain place at bound and at another place at free. For example, in the formula $\exists z(x = z) \wedge \forall x(x = y)$, the variable x occurs bound and free, whereas z occurs just bound and y occurs just free. However, one can always rename the bound variables occurring in a given formula φ such that each variable in φ is either bound or free — the rules for this procedure are given later. For a formula φ, the set of variables occurring free in φ is denoted by $\text{free}(\varphi)$. A formula φ is a **sentence** (or a **closed formula**) if it contains no free variables (i.e., $\text{free}(\varphi) = \emptyset$). For example, $\forall x(x = x)$ is a sentence but $x = x$ is just a formula.

In analogy to this definition, we say that a term is a **closed term** if it contains no variables. Obviously, the only terms which are closed are the constant symbols and the function symbols followed by closed terms.

Sometimes it is useful to indicate explicitly which variables occur free in a given formula φ, and for this we usually write $\varphi(x_1, \ldots, x_n)$ to indicate that $\{x_1, \ldots, x_n\} \subseteq \text{free}(\varphi)$.

If τ and τ_0 are terms and ν is a variable, then $\tau(\nu/\tau_0)$ is the term that we obtain from τ after replacing all instances of ν bei τ_0. In the case of a formula, this is more complicated since variables are either free or bound. Hence, if φ is a formula, ν a variable, and τ a term, then $\varphi(\nu/\tau)$ is the formula we get after replacing all *free* instances of the variable ν by τ. The process by which we obtain the formula $\varphi(\nu/\tau)$ is called **substitution**. Now, a substitution is **admissible** if and only if no free occurrence of ν in φ is in the range of a quantifier that binds any variable which appears in τ (i.e., for each variable $\tilde{\nu}$ appearing in τ, no place where ν occurs free in φ is in the range of $\exists\tilde{\nu}$ or $\forall\tilde{\nu}$). For example, if $x \notin \text{free}(\varphi)$, then $\varphi(x/\tau)$ is admissible for any term τ. In this case, the formulae φ and $\varphi(x/\tau)$ are identical, which we express by

$\varphi \equiv \varphi(x/\tau)$. In general, we use the symbol $\equiv$ in the metalanguage to denote an equality of strings of symbols of the formal language. Furthermore, if φ is a formula and the substitution $\varphi(x/\tau)$ is admissible, then we write just $\varphi(\tau)$ instead of $\varphi(x/\tau)$. In order to express this, we write $\varphi(\tau) :\equiv \varphi(x/\tau)$, where we use the symbol $:\equiv$ in the metalanguage to define symbols (or strings of symbols) of the formal language.

So far, we have letters, and we can build names and sentences. However, these sentences are just strings of symbols without any inherent meaning. At a later stage, we shall interpret formulae in the intuitively natural way by giving the symbols their intended meaning (e.g., $\wedge$ meaning "and", $\forall x$ meaning "for all x", et cetera). But before we shall do so, let us stay a little bit longer on the syntactical side — nevertheless, one should consider the formulae from a semantical point of view as well.

Axioms

In what follows, we shall label certain formulae or types of formulae as **axioms**, which are used in connection with *inference rules* in order to derive further formulae. From a semantical point of view we can think of axioms as "true" statements from which we deduce or prove further results. We distinguish two types of axiom, namely *logical axioms* and *non-logical axioms* (which will be discussed later). A **logical axiom** is a sentence or formula φ which is universally valid (i.e., φ is true in any possible universe, no matter how the variables, constants, et cetera, occurring in φ are interpreted). Usually, one takes as logical axioms some minimal set of formulae that is sufficient for deriving all universally valid formulae — such a set is given below.

If a symbol, involved in an axiom, stands for an arbitrary relation, function, or even for a first-order formula, then we usually consider the statement as an **axiom schema** rather than a single axiom, since each instance of the symbol represents a single axiom. The following list of axiom schemata is a system of logical axioms.

Let φ, φ_1, φ_2, φ_3, and ψ, be arbitrary formulae:

L_0: $\varphi \vee \neg\varphi$
L_1: $\varphi \to (\psi \to \varphi)$
L_2: $(\psi \to (\varphi_1 \to \varphi_2)) \to ((\psi \to \varphi_1) \to (\psi \to \varphi_2))$
L_3: $(\varphi \wedge \psi) \to \varphi$
L_4: $(\varphi \wedge \psi) \to \psi$
L_5: $\varphi \to (\psi \to (\psi \wedge \varphi))$
L_6: $\varphi \to (\varphi \vee \psi)$
L_7: $\psi \to (\varphi \vee \psi)$
L_8: $(\varphi_1 \to \varphi_3) \to ((\varphi_2 \to \varphi_3) \to ((\varphi_1 \vee \varphi_2) \to \varphi_3))$
L_9: $\neg\varphi \to (\varphi \to \psi)$

Let τ be a term, ν a variable, and assume that the substitution which leads to $\varphi(\nu/\tau)$ is admissible:

L_{10}: $\forall\nu\varphi(\nu) \to \varphi(\tau)$
L_{11}: $\varphi(\tau) \to \exists\nu\varphi(\nu)$

Let ψ be a formula and let ν a variable such that $\nu \notin \mathrm{free}(\psi)$:

L_{12}: $\forall\nu(\psi \to \varphi(\nu)) \to (\psi \to \forall\nu\varphi(\nu))$
L_{13}: $\forall\nu(\varphi(\nu) \to \psi) \to (\exists\nu\varphi(\nu) \to \psi)$

What is not yet covered is the symbol $=$, so let us now have a closer look at the binary equality relation. The defining properties of equality can already be found in Book VII, Chapter 1 of Aristotle's *Topics* [2], where one of the rules to decide whether two things are the same is as follows: ... *you should look at every possible predicate of each of the two terms and at the things of which they are predicated and see whether there is any discrepancy anywhere. For anything which is predicated of the one ought also to be predicated of the other, and of anything of which the one is a predicate the other also ought to be a predicate.*

In our formal system, the binary equality relation is defined by the following three axioms. Let $\tau, \tau_1, \ldots, \tau_n, \tau'_1, \ldots, \tau'_n$ be arbitrary terms, let R be an n-ary relation symbol (e.g., the binary relation symbol $=$), and let F be an n-ary function symbol:

L_{14}: $\tau = \tau$
L_{15}: $\big(\tau_1 = \tau'_1 \wedge \cdots \wedge \tau_n = \tau'_n\big) \to \big(R(\tau_1, \ldots, \tau_n) \to R(\tau'_1, \ldots, \tau'_n)\big)$
L_{16}: $\big(\tau_1 = \tau'_1 \wedge \cdots \wedge \tau_n = \tau'_n\big) \to \big(F(\tau_1, \ldots, \tau_n) = F(\tau'_1, \ldots, \tau'_n)\big)$

where the ambiguous formula $\big(\tau_1 = \tau'_1 \wedge \cdots \wedge \tau_n = \tau'_n\big)$ written in Polish notation reads as:

$$\wedge{=}\,\tau_1\tau'_1\,\wedge{=}\,\tau_2\tau'_2\,\wedge{=}\,\tau_3\tau'_3\,\wedge\,\cdots\,\wedge{=}\,\tau_{n-1}\tau'_{n-1}\,{=}\,\tau_n\tau'_n$$

Finally, we define the logical operator $\leftrightarrow$, the quantifier $\exists!$ and the binary relation symbol $\neq$ by stipulating:

$$\begin{aligned} \varphi \leftrightarrow \psi \quad &:\Longleftrightarrow \quad (\varphi \to \psi) \wedge (\psi \to \varphi)\,, \\ \exists!\nu\varphi \quad &:\Longleftrightarrow \quad \exists\nu(\varphi(\nu) \wedge \forall\mu(\varphi(\mu) \to \mu = \nu)) \\ \tau \neq \tau' \quad &:\Longleftrightarrow \quad \neg(\tau = \tau')\,, \end{aligned}$$

where we use the symbol $:\Longleftrightarrow$ in the metalanguage to define relations between symbols (or strings of symbols) of the formal language (i.e., $\leftrightarrow$, $\exists!\nu\varphi$ and $\neq$ are just abbreviations).

This completes the list of our logical axioms. In addition to these axioms, we are allowed to state arbitrarily many formulae. In logic, such a (possibly empty) set of formulae is also called a **theory**, or, when the signature $\mathscr{L}$

is specified, an $\mathscr{L}$**-theory**. The formulae of a theory are the axioms of the theory, which are called **non-logical axioms**. So, a non-logical axiom is in fact just a formula which is not a logical axiom. Furthermore, the axioms of mathematical theories are always *sentences*, i.e., formulae without free variables. However, in order to develop the notion of formal proofs, we will also consider theories consisting of arbitrary sets of formulae — such theories are relevant only in logic.

Examples of mathematical theories (i.e., of sets of non-logical axioms) which will be discussed in this book are the axioms of Set Theory (see Part IV), the axioms of Peano Arithmetic PA (also known as *Number Theory*), and the axioms of Group Theory GT, which we discuss first.

GT: The language of **Group Theory** is $\mathscr{L}_{\mathsf{GT}} = \{\mathsf{e}, \circ\}$, where e is a constant symbol and $\circ$ is a binary function symbol.

GT_0: $\forall x \forall y \forall z (x \circ (y \circ z) = (x \circ y) \circ z)$ (i.e., $\circ$ is *associative*)

GT_1: $\forall x (\mathsf{e} \circ x = x)$ (i.e., e is a *left-neutral* element)

GT_2: $\forall x \exists y (y \circ x = \mathsf{e})$ (i.e., every element has a *left-inverse*)

PA: The language of **Peano Arithmetic** is $\mathscr{L}_{\mathsf{PA}} = \{0, \mathsf{s}, +, \cdot\}$, where 0 is a constant symbol, s is a unary function symbol, and $+$, $\cdot$ are binary function symbols.

PA_0: $\neg \exists x (\mathsf{s}x = 0)$

PA_1: $\forall x \forall y (\mathsf{s}x = \mathsf{s}y \to x = y)$

PA_2: $\forall x (x + 0 = x)$

PA_3: $\forall x \forall y (x + \mathsf{s}y = \mathsf{s}(x + y))$

PA_4: $\forall x (x \cdot 0 = 0)$

PA_5: $\forall x \forall y (x \cdot \mathsf{s}y = (x \cdot y) + x)$

Let φ be an arbitrary $\mathscr{L}_{\mathsf{PA}}$-formula with $\mathrm{free}(\varphi) = \{x\}$:

PA_6: $\big(\varphi(0) \wedge \forall x (\varphi(x) \to \varphi(\mathsf{s}x))\big) \to \forall x \varphi(x)$

Notice that PA_6 is an axiom schema, known as the Induction Schema, and not just a single axiom like PA_0–PA_5.

It is often convenient to add certain *defined symbols* to a given language so that the expressions get shorter or are at least easier to read. For example, in Peano Arithmetic — which is an axiomatic system for the natural numbers — we usually replace the expression $\mathsf{s}0$ with 1 and $\mathsf{ss}0$ with 2. More formally, we define:

$$1 :\equiv \mathsf{s}0 \qquad \text{and} \qquad 2 :\equiv \mathsf{ss}0$$

where we use the symbol $:\equiv$ in the metalanguage to define new constant symbols or certain formulae. Obviously, all that can be expressed in the language $\mathscr{L}_{\mathsf{PA}} \cup \{1, 2\}$ can also be expressed in $\mathscr{L}_{\mathsf{PA}}$.

Formal Proofs

So far, in a certain language we have a set of logical axioms and can assume some non-logical axioms. Furthermore, we can define, if we wish, as many new constants, functions, and relations as we like. However, we are still not able to deduce anything from the given axioms, since until now, we do not have *inference rules* which allow us, for example, to infer a certain sentence from a given set of axioms.

Surprisingly, just two **inference rules** are sufficient, namely:

$$\text{Modus Ponens (MP):}\quad \frac{\varphi \to \psi,\ \varphi}{\psi} \qquad \text{and} \qquad \text{Generalisation } (\forall):\quad \frac{\varphi}{\forall \nu \varphi}$$

In the former case we say that the formula ψ is obtained from $\varphi \to \psi$ and φ by Modus Ponens, abbreviated (MP), and in the latter case we say that $\forall \nu \varphi$ (where ν can be any variable) is obtained from φ by Generalisation, abbreviated $(\forall)$.

Using these two inference rules, we are now able to define the notion of a formal proof: Let $\mathscr{L}$ be a signature (i.e., a possibly empty set of non-logical symbols) and let $\mathbf{\Phi}$ be a possibly empty set of $\mathscr{L}$-formulae (e.g., a set of non-logical axioms). An $\mathscr{L}$-formula ψ is **provable** from $\mathbf{\Phi}$ (or provable in $\mathbf{\Phi}$), denoted $\mathbf{\Phi} \vdash \psi$, if there is a FINITE sequence $\varphi_0, \ldots, \varphi_n$ of $\mathscr{L}$-formulae such that $\varphi_n \equiv \psi$ (i.e., the formulae φ_n and ψ are identical), and for all i with $i \leq n$ we are in at least one of the following cases:

- φ_i is a logical axiom, or
- $\varphi_i \in \mathbf{\Phi}$, or
- there are $j, k < i$ such that $\varphi_j \equiv \varphi_k \to \varphi_i$, or
- there is a $j < i$ such that $\varphi_i \equiv \forall \nu \, \varphi_j$, where ν is a variable which does not occur free in any formula of $\mathbf{\Phi}$.

The sequence $\varphi_0, \ldots, \varphi_n$ is then called a **formal proof** of ψ.

In the case when $\mathbf{\Phi}$ is the empty set, we simply write $\vdash \psi$. If a formula ψ is not provable from $\mathbf{\Phi}$, i.e., if there is no formal proof for ψ which uses just formulae from $\mathbf{\Phi}$, then we write $\mathbf{\Phi} \nvdash \psi$.

Formal proofs, even of very simple statements, can get quite long and tricky. Nevertheless, we shall give a few examples:

EXAMPLE 1.0. To warm up, let us formally prove that the equality relation is reflexive, which is expressed by the sentence $\forall x(x = x)$:

φ_0:	$x = x$	instance of L_{14}
φ_1:	$\forall x(x = x)$	from φ_0 by $(\forall)$

Now, let us give a formal proof of $\exists x(x = x)$:

φ_0:	$x = x$	instance of L_{14}
φ_1:	$x = x \to \exists x(x = x)$	instance of L_{11} (notice that for $\varphi \equiv x = x$, the substitution $\varphi(x/x)$ is admissible)
φ_2:	$\exists x(x = x)$	from φ_1 and φ_0 by (MP)

Example 1.1. For every formula φ we have:

$$\vdash \varphi \to \varphi$$

A formal proof of $\varphi \to \varphi$ is given by:

φ_0: $(\varphi \to ((\varphi \to \varphi) \to \varphi)) \to ((\varphi \to (\varphi \to \varphi)) \to (\varphi \to \varphi))$	instance of L_2
φ_1: $\varphi \to ((\varphi \to \varphi) \to \varphi)$	instance of L_1
φ_2: $(\varphi \to (\varphi \to \varphi)) \to (\varphi \to \varphi)$	from φ_0 and φ_1 by (MP)
φ_3: $\varphi \to (\varphi \to \varphi)$	instance of L_1
φ_4: $\varphi \to \varphi$	from φ_2 and φ_3 by (MP)

Example 1.2. We give a formal proof of $\mathsf{PA} \vdash 1 + 1 = 2$. Recall that we have defined $1 :\equiv \mathsf{s0}$ and $2 :\equiv \mathsf{ss0}$, so we need to prove $\mathsf{PA} \vdash \mathsf{s0} + \mathsf{s0} = \mathsf{ss0}$.

φ_0:	$\forall x \forall y(x + \mathsf{s}y = \mathsf{s}(x + y))$	PA_3
φ_1:	$\forall x \forall y(x + \mathsf{s}y = \mathsf{s}(x + y)) \to \forall y(\mathsf{s0} + \mathsf{s}y = \mathsf{s}(\mathsf{s0} + y))$	instance of L_{10}
φ_2:	$\forall y(\mathsf{s0} + \mathsf{s}y = \mathsf{s}(\mathsf{s0} + y))$	from φ_1 and φ_0 by (MP)
φ_3:	$\forall y(\mathsf{s0} + \mathsf{s}y = \mathsf{s}(\mathsf{s0} + y)) \to \mathsf{s0} + \mathsf{s0} = \mathsf{s}(\mathsf{s0} + 0)$	instance of L_{10}
φ_4:	$\mathsf{s0} + \mathsf{s0} = \mathsf{s}(\mathsf{s0} + 0)$	from φ_3 and φ_2 by (MP)
φ_5:	$\forall x(x + 0 = x)$	PA_2
φ_6:	$\forall x(x + 0 = x) \to \mathsf{s0} + 0 = \mathsf{s0}$	instance of L_{10}
φ_7:	$\mathsf{s0} + 0 = \mathsf{s0}$	from φ_6 and φ_5 by (MP)
φ_8:	$\mathsf{s0} + 0 = \mathsf{s0} \to \mathsf{s}(\mathsf{s0} + 0) = \mathsf{ss0}$	instance of L_{16}
φ_9:	$\mathsf{s}(\mathsf{s0} + 0) = \mathsf{ss0}$	from φ_8 and φ_7 by (MP)
φ_{10}:	$\mathsf{s0} + \mathsf{s0} = \mathsf{s0} + \mathsf{s0}$	instance of L_{14}
φ_{11}:	$\varphi_9 \to (\varphi_{10} \to (\varphi_{10} \wedge \varphi_9))$	instance of L_5
φ_{12}:	$\varphi_{10} \to (\varphi_{10} \wedge \varphi_9)$	from φ_{11} and φ_9 by (MP)
φ_{13}:	$\varphi_{10} \wedge \varphi_9$	from φ_{12} and φ_{10} by (MP)
φ_{14}:	$(\varphi_{10} \wedge \varphi_9) \to (\mathsf{s0} + \mathsf{s0} = \mathsf{s}(\mathsf{s0} + 0) \to \mathsf{s0} + \mathsf{s0} = \mathsf{ss0})$	instance of L_{15}
φ_{15}:	$\mathsf{s0} + \mathsf{s0} = \mathsf{s}(\mathsf{s0} + 0) \to \mathsf{s0} + \mathsf{s0} = \mathsf{ss0}$	from φ_{14} and φ_{13} by (MP)
φ_{16}:	$\mathsf{s0} + \mathsf{s0} = \mathsf{ss0}$	from φ_{15} and φ_4 by (MP)

In Chapter 2, we will introduce some techniques which allow us to simplify formal proofs such as the one presented above.

Tautologies & Logical Equivalence

We say that two formulae φ and ψ are **logically equivalent** (or just **equivalent**), denoted $\varphi \Leftrightarrow \psi$, if $\vdash \varphi \leftrightarrow \psi$. More formally:

$$\varphi \Leftrightarrow \psi \quad :\Longleftrightarrow \quad \vdash \varphi \leftrightarrow \psi$$

In other words, if $\varphi \Leftrightarrow \psi$, then — from a logical point of view — φ and ψ state exactly the same, and therefore we could call $\varphi \leftrightarrow \psi$ a tautology, which means *saying the same thing twice*. Indeed, in logic, a formula φ is a **tautology** if $\vdash \varphi$. Thus, the formulae φ and ψ are equivalent if and only if $\varphi \leftrightarrow \psi$ is a tautology. More generally, if $\mathbf{\Phi}$ is a set of formulae, we write $\varphi \Leftrightarrow_{\mathbf{\Phi}} \psi$ to denote $\mathbf{\Phi} \vdash \varphi \leftrightarrow \psi$.

Example 1.3. For every formula φ we have:

$$\varphi \Leftrightarrow \varphi$$

This follows directly from Example 1.1, since $\varphi \leftrightarrow \varphi$ is simply an abbreviation for $(\varphi \to \varphi) \wedge (\varphi \to \varphi)$:

φ_0:	$\varphi \to \varphi$	Example 1.1
φ_1:	$(\varphi \to \varphi) \to ((\varphi \to \varphi) \to (\varphi \leftrightarrow \varphi))$	instance of L_5
φ_2:	$(\varphi \to \varphi) \to (\varphi \leftrightarrow \varphi)$	from φ_0 and φ_1 by (MP)
φ_3:	$\varphi \leftrightarrow \varphi$	from φ_0 and φ_2 by (MP)

Example 1.4. For every formula φ we have:

$$\varphi \Leftrightarrow \neg\neg\varphi$$

which corresponds to Tautology (F) given at the end of the book. By applying L_5 as in Example 1.3, one can easily check that it suffices to prove separately that the formulae $\varphi \to \neg\neg\varphi$ and $\neg\neg\varphi \to \varphi$ are tautologies. We only prove the former statement, the latter one is proved in Example 2.2.

φ_0:	$(\neg\varphi \to (\varphi \to \neg\neg\varphi)) \to ((\neg\neg\varphi \to (\varphi \to \neg\neg\varphi)) \to ((\neg\varphi \vee \neg\neg\varphi) \to (\varphi \to \neg\neg\varphi)))$	instance of L_8
φ_1:	$\neg\varphi \to (\varphi \to \neg\neg\varphi)$	instance of L_9
φ_2:	$(\neg\neg\varphi \to (\varphi \to \neg\neg\varphi)) \to ((\neg\varphi \vee \neg\neg\varphi) \to (\varphi \to \neg\neg\varphi))$	from φ_0 and φ_1 by (MP)
φ_3:	$\neg\neg\varphi \to (\varphi \to \neg\neg\varphi)$	instance of L_1
φ_4:	$(\neg\varphi \vee \neg\neg\varphi) \to (\varphi \to \neg\neg\varphi)$	from φ_2 and φ_2 by (MP)
φ_5:	$\neg\varphi \vee \neg\neg\varphi$	instance of L_0
φ_6:	$\varphi \to \neg\neg\varphi$	from φ_4 and φ_5 by (MP)

Example 1.5. Commutativity and associativity of $\wedge$ and $\vee$ are tautological, i.e., for all formulae φ, ψ and χ we have $\varphi \wedge \psi \Leftrightarrow \psi \wedge \varphi$ and $\varphi \wedge (\psi \wedge \chi) \Leftrightarrow (\varphi \wedge \psi) \wedge \chi$; and similarly for $\vee$. Again, we omit the proof since it will be trivial once we have proved the Deduction Theorem 2.0 (see also Exercises 1.1 and 1.2). This result legitimises the notations $\varphi_0 \wedge \ldots \wedge \varphi_n$ and $\varphi_0 \vee \ldots \vee \varphi_n$, respectively, for $\varphi_0 \wedge (\varphi_1 \wedge (\ldots \wedge \varphi_n) \ldots)$ and $\varphi_0 \vee (\varphi_1 \vee (\ldots \vee \varphi_n) \ldots)$, respectively.

At the end of the book, there is a list of tautologies which will be frequently used in formal proofs. Note that it follows from Exercise 2.1 that $\Leftrightarrow$ defines an equivalence relation on all $\mathscr{L}$-formulae for some given signature $\mathscr{L}$. Moreover, it even defines a congruence relation (i.e., equivalence is closed under all logical operations). More precisely, if $\varphi \Leftrightarrow \varphi'$ and $\psi \Leftrightarrow \psi'$, then:

$$\begin{gathered} \neg\varphi \Leftrightarrow \neg\varphi' \\ \varphi \circ \psi \Leftrightarrow \varphi' \circ \psi' \\ \nexists \nu\varphi \Leftrightarrow \nexists \nu\varphi' \end{gathered}$$

where $\circ$ stands for either $\wedge, \vee$, or $\rightarrow$, and $\nexists$ stands for either $\exists$ or $\forall$. A proof of these statements will be easier once we have proved the the Deduction Theorem 2.0.

The above observation enables us to replace *subformulae* (i.e., proper formulae which are part of the formula φ) of a given formula φ by equivalent formulae so that the resulting formula is equivalent to φ.

Theorem 1.6 (Substitution Theorem). *Let φ be a formula and let α be a subformula of φ. Let ψ be the formula obtained from φ by replacing one or multiple occurrences of α by some formula β. Then we have:*

$$\alpha \Leftrightarrow \beta \quad \Longrightarrow \quad \varphi \Leftrightarrow \psi$$

Proof. We prove the theorem by induction on the recursive construction of the formula φ. If φ is an atomic formula or if α is φ, then the statement is trivial. If φ is a composite formula, then we use the observation that $\Leftrightarrow$ is a congruence relation: For example, if the formula φ is of the form $\neg\varphi'$, and ψ' is the formula obtained from φ' by replacing one or multiple occurrences of α by β, then by induction we may assume that $\varphi' \Leftrightarrow \psi'$. Consequently, we have $\neg\varphi' \Leftrightarrow \neg\psi'$ as desired. The other cases can be checked in a similar way. ⊣

Theorem 1.7 (Three-Symbols). *For every formula φ there is an equivalent formula ψ which contains only the symbols $\neg$ and $\wedge$ as logical operators and $\exists$ as quantifier.*

Proof. By THEOREM 1.6, it suffices to prove the following equivalences:

$$\begin{aligned} \varphi \vee \psi &\Leftrightarrow \neg(\neg\varphi \wedge \neg\psi) \\ \varphi \to \psi &\Leftrightarrow \neg\varphi \vee \psi \\ \forall\nu\varphi &\Leftrightarrow \neg\exists\nu\neg\varphi \end{aligned}$$

The proof of these equivalences is left as an exercise (see EXERCISE 2.5). Note that the methods of proof introduced in Chapter 2 will simplify such proofs to a great extent. ⊣

As a consequence of THEOREM 1.7, one could simplify both the alphabet and the logical axioms. Nevertheless, we do not wish to do so, since this would also decrease the readability of formulae.

NOTES

The logical axioms are essentially those given by Hilbert (see, e.g., [27]). However, Hilbert also introduced the axiom schemata $\big((\varphi \to \psi) \wedge (\neg\varphi \to \psi)\big) \to \psi$ and $\neg\neg\varphi \to \varphi$. On the other hand, he did not introduce the LAW OF EXCLUDED MIDDLE L_0, because it is not needed any more in this setting (see also EXERCISES 2.9 and 2.10.(c)).

EXERCISES

1.0 (a) $\{\varphi, \psi\} \vdash \varphi \wedge \psi$
(b) $\{\varphi \wedge \psi\} \vdash \psi \wedge \varphi$

1.1 (a) $\vdash \varphi \vee \psi \to \psi \vee \varphi$
(b) $\vdash \varphi \wedge \psi \to \psi \wedge \varphi$

1.2 (a) $\{\psi_1 \wedge (\psi_2 \wedge \psi_3)\} \vdash (\psi_1 \wedge \psi_2) \wedge \psi_3$
(b) $\{(\psi_1 \wedge \psi_2) \wedge \psi_3\} \vdash \psi_1 \wedge (\psi_2 \wedge \psi_3)$

1.3 $\{\varphi \to \psi\} \vdash \neg\psi \to \neg\varphi$

1.4 (a) $\{\psi_0 \to \psi_1, \psi_1 \to \psi_2\} \vdash \psi_0 \to \psi_2$
(b) $\{\psi_0 \to \varphi, \psi_1 \to \varphi\} \vdash (\psi_0 \vee \psi_1) \to \varphi$
(c) $\{\varphi \to \psi_0, \varphi \to \psi_1\} \vdash \varphi \to (\psi_0 \wedge \psi_1)$

1.5 $\vdash \big((\psi_1 \wedge \psi_2) \vee \psi_3\big) \to \big((\psi_1 \vee \psi_3) \wedge (\psi_2 \vee \psi_3)\big)$

1.6 $\vdash \forall x \forall y (x = y \to y = x)$

Chapter 2
The Art of Proof

In Example 1.2 we gave a proof of $1+1=2$ in seventeen proof steps. At that point you may have asked yourself: If it takes that much effort to prove such a simple statement, how can one ever prove any non-trivial mathematical result using formal proofs? This objection is of course justified; however, we will show in this chapter how one can simplify formal proofs using some methods of proof such as proofs by cases or by contradiction. It is crucial to note that the following results are not theorems of a formal theory, but theorems about formal proofs. In particular, they show how — under certain conditions — a formal proof can be transformed into another.

The Deduction Theorem

In common mathematics, one usually proves implications of the form

$$\text{If } \Phi \text{ then } \Psi$$

by simply assuming the truth of Φ and deriving from this the truth of Ψ. When writing formal proofs, the so-called Deduction Theorem enables us to use a similar trick: Rather than proving $\mathbf{\Phi} \vdash \varphi \to \psi$ we simply add φ to our set of formulae $\mathbf{\Phi}$ and prove $\mathbf{\Phi} \cup \{\varphi\} \vdash \psi$.

If $\mathbf{\Phi}$ is a set of formulae and $\mathbf{\Phi}'$ is another set of formulae in the same language as $\mathbf{\Phi}$, then we write $\mathbf{\Phi} + \mathbf{\Phi}'$ for $\mathbf{\Phi} \cup \mathbf{\Phi}'$. In the case when the set $\mathbf{\Phi}'$ consists of a single formula φ, we write $\mathbf{\Phi} + \varphi$ instead of $\mathbf{\Phi} \cup \{\varphi\}$.

Theorem 2.0 (Deduction Theorem). *If $\mathbf{\Phi}$ is a set of formulae and $\mathbf{\Phi} + \psi \vdash \varphi$, then $\mathbf{\Phi} \vdash \psi \to \varphi$; and vice versa, if $\mathbf{\Phi} \vdash \psi \to \varphi$, then $\mathbf{\Phi} + \psi \vdash \varphi$, i.e., we have:*

$$\mathbf{\Phi} + \psi \vdash \varphi \quad \Longleftrightarrow \quad \mathbf{\Phi} \vdash \psi \to \varphi \tag{DT}$$

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_3

Proof. It is clear that $\mathbf{\Phi} \vdash \psi \to \varphi$ implies $\mathbf{\Phi} + \psi \vdash \varphi$. Conversely, suppose that $\mathbf{\Phi} + \psi \vdash \varphi$ holds and let the sequence $\varphi_0, \ldots, \varphi_n$ with $\varphi_n \equiv \varphi$ be a formal proof for φ from $\mathbf{\Phi} + \psi$. For each $i \leq n$ we will replace the formula φ_i by a sequence of formulae which ends with $\psi \to \varphi_i$. Let $i \leq n$ and assume $\mathbf{\Phi} \vdash \psi \to \varphi_j$ for every $j < i$.

- If φ_i is a logical axiom or $\varphi_i \in \mathbf{\Phi}$, we have

$\varphi_{i,0}$:	φ_i	$\varphi_i \in \mathbf{\Phi}$ or φ_i is a logical axiom
$\varphi_{i,1}$:	$\varphi_i \to (\psi \to \varphi_i)$	instance of L_1
$\varphi_{i,2}$:	$\psi \to \varphi_i$	from $\varphi_{i,1}$ and $\varphi_{i,0}$ by (MP)

- The case $\varphi_i \equiv \psi$ follows directly from Example 1.1.

- If φ_i is obtained from φ_j and $\varphi_k \equiv (\varphi_j \to \varphi_i)$ by Modus Ponens, where $j, k < i$, we have:

$\varphi_{i,0}$:	$\psi \to \varphi_j$	since $j < i$
$\varphi_{i,1}$:	$\psi \to (\varphi_j \to \varphi_i)$	since $k < i$
$\varphi_{i,2}$:	$\varphi_{i,1} \to ((\psi \to \varphi_j) \to (\psi \to \varphi_i))$	instance of L_2
$\varphi_{i,3}$:	$(\psi \to \varphi_j) \to (\psi \to \varphi_i)$	from $\varphi_{i,2}$ and $\varphi_{i,1}$ by (MP)
$\varphi_{i,4}$:	$\psi \to \varphi_i$	from $\varphi_{i,3}$ and $\varphi_{i,0}$ by (MP)

- If φ_i is obtained from φ_j by Generalisation, where $j < i$, i.e., $\varphi_i \equiv \forall \nu \varphi_j$ for some variable ν, then, by the rules of Generalisation, the variable ν does not occur free in ψ. In particular, $\nu \notin \text{free}(\psi)$, and the claim follows from:

$\varphi_{i,0}$:	$\psi \to \varphi_j$	since $j < i$
$\varphi_{i,1}$:	$\forall \nu(\psi \to \varphi_j)$	from $\varphi_{i,0}$ by $(\forall)$
$\varphi_{i,2}$:	$\forall \nu(\psi \to \varphi_j) \to (\psi \to \varphi_i)$	instance of L_{12}
$\varphi_{i,3}$:	$\psi \to \varphi_i$	from $\varphi_{i,2}$ and $\varphi_{i,1}$ by (MP)

Hence, we have $\mathbf{\Phi} \vdash \psi \to \varphi$. ⊣

Notice that the Deduction Theorem allows us under certain conditions to transform a formal proof into another. So, the Deduction Theorem is a theorem about formal proofs (i.e., about sequences of formulae) and not a theorem of a theory. Notice also that in order to prove the Deduction Theorem restricted to Propositional Logic (i.e., quantifier-free formulae), we need only the logical axioms L_1 and L_2.

Note that $\vdash \varphi \to \varphi$ is a trivial consequence of the Deduction Theorem, whereas its formal proof in Example 1.1 has five steps. Let us now consider a few application of the Deduction Theorem.

Example 2.1. As a first application, we show that the equality relation is symmetric and transitive, i.e., we show

$$\vdash \forall x\, \forall y\, (x = y \to y = x) \quad \text{and} \quad \vdash \forall x\, \forall y\, \forall z\, \big((x = y \wedge y = z) \to y = x\big)\,.$$

In order to show that the equality relation is symmetric, we first show that $\{x = y\} \vdash y = x$, i.e., we assume the non-logical axiom $x = y$ and show that from this axiom we can prove $y = x$:

φ_0:	$(x = y \wedge x = x) \to (x = x \to y = x)$	instance of L_{15}
φ_1:	$x = x$	instance of L_{14}
φ_2:	$x = y$	$x = y$ belongs to $\{x = y\}$
φ_3:	$x = x \to (x = y \to (x = y \wedge x = x))$	instance of L_5
φ_4:	$x = y \to (x = y \wedge x = x)$	from φ_3 and φ_1 by (MP)
φ_5:	$x = y \wedge x = x$	from φ_4 and φ_2 by (MP)
φ_6:	$x = x \to y = x$	from φ_0 and φ_5 by (MP)
φ_7:	$y = x$	from φ_6 and φ_1 by (MP)

Thus, we have $\{x = y\} \vdash y = x$, and by the Deduction Theorem we obtain $\vdash x = y \to y = x$, i.e., from the empty set of non-logical axioms we can prove $x = y \to y = x$, and by applying twice Generalisation we finally get

$$\vdash \forall x \forall y (x = y \to y = x)\,.$$

For a proof that does not use the Deduction Theorem see the Solution to Exercise 1.6.

To show that the equality relation is transitive, we set $\mathbf{\Phi} = \{x = y \wedge y = z\}$ and first show that $\mathbf{\Phi} \vdash x = z$, where in the formal proof we make use of the fact that $\vdash x = y \to y = x$ has already been proven:

φ_0:	$x = y \to y = x$	already proven
φ_1:	$x = y \wedge y = z$	$x = y \wedge y = z$ belongs to $\mathbf{\Phi}$
φ_2:	$(x = y \wedge y = z) \to x = y$	instance of L_3
φ_3:	$x = y$	from φ_2 and φ_1 by (MP)
φ_4:	$y = x$	from φ_0 and φ_3 by (MP)
φ_5:	$(x = y \wedge y = z) \to y = z$	instance of L_4
φ_6:	$y = z$	from φ_5 and φ_1 by (MP)
φ_7:	$y = z \to \big(y = x \to (y = x \wedge y = z)\big)$	instance of L_5
φ_8:	$y = x \to (y = x \wedge y = z)$	from φ_7 and φ_6 by (MP)
φ_9:	$y = x \wedge y = z$	from φ_8 and φ_4 by (MP)
φ_{10}:	$(y = x \wedge y = z) \to (y = y \to x = z)$	instance of L_{15}
φ_{11}:	$y = y \to x = z$	from φ_{10} and φ_9 by (MP)
φ_{12}:	$y = y$	instance of L_{14}
φ_{13}:	$x = z$	from φ_{11} and φ_{12} by (MP)

Thus, we have $\{x = y \wedge y = z\} \vdash x = z$, and by the Deduction Theorem we obtain $\vdash (x = y \wedge y = z) \to x = z$, and after applying three times Generalisation we finally get

$$\vdash \forall x\, \forall y\, \forall z\, \big((x = y \wedge y = z) \to y = x\big)\,.$$

Example 2.2. As a second application, we prove $\neg\neg\varphi \to \varphi$, which is one direction of Tautology (F) given at the end of the book. For the other direction see Example 1.4 and for the relationship between $\neg\neg\varphi \to \varphi$ and the axioms of Propositional Logic, see the Solutions to Exercises 2.9.(b) and 2.10. By the Deduction Theorem it suffices to prove $\{\neg\neg\varphi\} \vdash \varphi$:

φ_0:	$\neg\neg\varphi \to (\neg\varphi \to \varphi)$	instance of L_9
φ_1:	$\neg\neg\varphi$	$\neg\neg\varphi \in \{\neg\neg\varphi\}$
φ_2:	$\neg\varphi \to \varphi$	from φ_0 and φ_1 by (MP)
φ_3:	$(\varphi \to \varphi) \to ((\neg\varphi \to \varphi) \to ((\varphi \vee \neg\varphi) \to \varphi))$	instance of L_8
φ_4:	$\varphi \to \varphi$	by Example 1.1
φ_5:	$(\neg\varphi \to \varphi) \to ((\varphi \vee \neg\varphi) \to \varphi)$	from φ_3 and φ_4 by (MP)
φ_6:	$(\varphi \vee \neg\varphi) \to \varphi$	from φ_5 and φ_2 by (MP)
φ_7:	$\varphi \vee \neg\varphi$	instance of L_0
φ_8:	φ	from φ_6 and φ_7 by (MP)

Example 2.3. As a third application, we prove Tautology (K), which is the statement $(\varphi \to \psi) \leftrightarrow (\neg\varphi \vee \psi)$. We first show $\vdash (\neg\varphi \vee \psi) \to (\varphi \to \psi)$:

φ_0:	$(\neg\varphi \to (\varphi \to \psi)) \to$ $((\psi \to (\varphi \to \psi)) \to ((\neg\varphi \vee \psi) \to (\varphi \to \psi)))$	instance of L_8
φ_1:	$(\neg\varphi \to (\varphi \to \psi))$	instance of L_9
φ_2:	$(\psi \to (\varphi \to \psi)) \to ((\neg\varphi \vee \psi) \to (\varphi \to \psi))$	from φ_0 and φ_1 by (MP)
φ_3:	$\psi \to (\varphi \to \psi)$	instance of L_1
φ_4:	$(\neg\varphi \vee \psi) \to (\varphi \to \psi)$	from φ_2 and φ_3 by (MP)

Now, we show $\{\varphi \to \psi\} \vdash (\neg\varphi \vee \psi)$:

φ_0:	$\psi \to (\neg\varphi \vee \psi)$	instance of L_7
φ_1:	$(\psi \to (\neg\varphi \vee \psi)) \to (\varphi \to (\psi \to (\neg\varphi \vee \psi)))$	instance of L_1
φ_2:	$(\varphi \to (\psi \to (\neg\varphi \vee \psi))$	from φ_1 and φ_0 by (MP)
φ_3:	$(\varphi \to (\psi \to (\neg\varphi \vee \psi))) \to$ $((\varphi \to \psi) \to (\varphi \to (\neg\varphi \vee \psi)))$	instance of L_2
φ_4:	$(\varphi \to \psi) \to (\varphi \to (\neg\varphi \vee \psi))$	from φ_3 and φ_2 by (MP)
φ_5:	$\varphi \to \psi$	assumption
φ_6:	$\varphi \to (\neg\varphi \vee \psi)$	from φ_4 and φ_5 by (MP)
φ_7:	$(\varphi \to (\neg\varphi \vee \psi)) \to$ $((\neg\varphi \to (\neg\varphi \vee \psi)) \to ((\varphi \vee \neg\varphi) \to (\neg\varphi \vee \psi)))$	instance of L_8
φ_8:	$(\neg\varphi \to (\neg\varphi \vee \psi)) \to ((\varphi \vee \neg\varphi) \to (\neg\varphi \vee \psi))$	from φ_7 and φ_6 by (MP)
φ_9:	$\neg\varphi \to (\neg\varphi \vee \psi)$	instance of L_6
φ_{10}:	$(\varphi \vee \neg\varphi) \to (\neg\varphi \vee \psi)$	from φ_8 and φ_9 by (MP)
φ_{11}:	$\varphi \vee \neg\varphi$	instance of L_0
φ_{12}:	$\neg\varphi \vee \psi$	from φ_{10} and φ_{11} by (MP)

By applying (DT) we find $\vdash (\varphi \to \psi) \to (\neg\varphi \vee \psi)$, and after combining this statement with the previous statement using L_5, we finally obtain Tautology (K).

Remark. The tautologies listed at the end of the book can all be proven from the logical axioms (for some examples see the Solution to Exercise 2.4). So, we can consider tautologies as if they were additional logical axioms.

Natural Deduction

We have introduced predicate logic so that there are many logical axioms and only two inference rules. However, it is also possible to introduce calculi with an opposite approach: few axioms and many inference rules. In the calculus of **natural deduction** there are, in fact, no axioms at all. Its inference rules essentially state how to transform a given formal proof to another one. We write $\mathbf{\Phi} \vdash \varphi$ to state that there is a formal proof of φ in the calculus of natural deduction with the non-logical axioms given by $\mathbf{\Phi}$.

Let $\mathbf{\Phi}$ be a set of formulae and let φ, ψ, χ be any formulae. The first rule states how formal proofs can be initialized.

$$\text{Initial Rule (IR):} \quad \frac{}{\mathbf{\Phi} \vdash \varphi} \qquad \text{for } \varphi \in \mathbf{\Phi}.$$

In the calculus of natural deduction there are so-called **introduction rules** and **elimination rules** for each logical symbol.

$$(\mathsf{I}\wedge): \quad \frac{\mathbf{\Phi} \vdash \varphi,\ \mathbf{\Phi} \vdash \psi}{\mathbf{\Phi} \vdash \varphi \wedge \psi} \qquad (\mathsf{E}\wedge): \quad \frac{\mathbf{\Phi} \vdash \varphi \wedge \psi}{\mathbf{\Phi} \vdash \varphi} \quad \frac{\mathbf{\Phi} \vdash \varphi \wedge \psi}{\mathbf{\Phi} \vdash \psi}$$

$$(\mathsf{I}\vee): \quad \frac{\mathbf{\Phi} \vdash \varphi}{\mathbf{\Phi} \vdash \varphi \vee \psi} \quad \frac{\mathbf{\Phi} \vdash \psi}{\mathbf{\Phi} \vdash \varphi \vee \psi} \qquad (\mathsf{E}\vee): \quad \frac{\mathbf{\Phi} \vdash \varphi \vee \psi,\ \mathbf{\Phi} + \varphi \vdash \chi,\ \mathbf{\Phi} + \psi \vdash \chi}{\mathbf{\Phi} \vdash \chi}$$

$$(\mathsf{I}\to): \quad \frac{\mathbf{\Phi} + \varphi \vdash \psi}{\mathbf{\Phi} \vdash \varphi \to \psi} \qquad (\mathsf{E}\to): \quad \frac{\mathbf{\Phi} \vdash \varphi \to \psi,\ \mathbf{\Phi} \vdash \varphi}{\mathbf{\Phi} \vdash \psi}$$

$$(\mathsf{I}\neg): \quad \frac{\mathbf{\Phi} + \varphi \vdash \psi \wedge \neg\psi}{\mathbf{\Phi} \vdash \neg\varphi} \qquad (\mathsf{E}\neg): \quad \frac{\mathbf{\Phi} \vdash \neg\neg\varphi}{\mathbf{\Phi} \vdash \varphi}$$

Let τ be a term and ν be a variable such that the substitution $\varphi(\nu/\tau)$ is admissible and $\nu \notin \text{free}(\chi)$ for any formula $\chi \in \mathbf{\Phi}$ and — in the case of $(\mathsf{E}\exists)$ — $\nu \notin \text{free}(\psi)$. Now we can state the corresponding introduction and elimination rules for quantifiers:

$$(\mathsf{I}\exists): \quad \frac{\mathbf{\Phi} \vdash \varphi(\tau)}{\mathbf{\Phi} \vdash \exists\nu\varphi(\nu)} \qquad (\mathsf{E}\exists): \quad \frac{\mathbf{\Phi} \vdash \exists\nu\varphi(\nu),\ \mathbf{\Phi} + \varphi(\nu) \vdash \psi}{\mathbf{\Phi} \vdash \psi}$$

$$(\mathsf{I}\forall): \quad \frac{\mathbf{\Phi} \vdash \varphi(\nu)}{\mathbf{\Phi} \vdash \forall\nu\varphi(\nu)} \qquad (\mathsf{E}\forall): \quad \frac{\mathbf{\Phi} \vdash \forall\nu\varphi(\nu)}{\mathbf{\Phi} \vdash \varphi(\tau)}$$

Finally, we need to deal with equality and atomic formulae. Let τ, τ_1 and τ_2 be terms and φ an atomic formula. The following introduction and elimination

rules for equality are closely related to the logical axioms L_{14}–L_{16}:

$$(\mathsf{I}{=}): \quad \frac{}{\tau = \tau} \qquad\qquad (\mathsf{E}{=}): \quad \frac{\mathbf{\Phi} \vdash \tau_1 = \tau_2,\ \mathbf{\Phi} \vdash \varphi(\nu/\tau_1)}{\mathbf{\Phi} \vdash \varphi(\nu/\tau_2)}$$

Formal proofs in the calculus of natural deduction are defined in a similar way as in our usual calculus: There is a formal proof of a formula φ from a set of formulae $\mathbf{\Phi}$, denoted by $\mathbf{\Phi} \vdash \varphi$, if there is a FINITE sequence of of pairs $(\mathbf{\Phi}_0, \varphi_0), \ldots, (\mathbf{\Phi}_n, \varphi_n)$ such that $\mathbf{\Phi}_n \equiv \mathbf{\Phi}$, $\varphi_n \equiv \varphi$ and for each $i \leq n$, $\mathbf{\Phi}_i \vdash \varphi_i$ is obtained by the application of an inference rule

$$\frac{\mathbf{\Phi}_{j_0} \vdash \varphi_{j_0},\ \ldots,\ \mathbf{\Phi}_{j_k} \vdash \varphi_{j_k}}{\mathbf{\Phi}_i \vdash \varphi_i}$$

with $k \leq 3$ and $j_0, \ldots, j_k < i$. Note that the the case $k = 0$ is permitted, which corresponds to an application of the INITIAL RULE. In the case when $\mathbf{\Phi}$ is the empty set, we simply write $\vdash \varphi$.

We have now described two ways of introducing formal proofs. It is therefore natural to ask whether the two systems prove the same theorems. Fortunately, this question turns out to have a positive answer.

THEOREM 2.4. *Let $\mathbf{\Phi}$ be a set of formulae and let φ be a formula. Then we have*

$$\mathbf{\Phi} \vdash \varphi \iff \mathbf{\Phi} \vdash \varphi.$$

Proof. We need to verify that every formal proof in the usual sense can be turned into a formal proof in the calculus of natural deduction and vice versa. In order to prove that $\mathbf{\Phi} \vdash \varphi$ implies $\mathbf{\Phi} \vdash \varphi$ for every formula φ, we need to derive all introduction and elimination rules from our logical axioms and (MP) and ($\forall$). We focus only on some of the rules and leave the others as an exercise.

Formal proofs of the form $\mathbf{\Phi} \vdash \varphi$ with $\varphi \in \mathbf{\Phi}$ using only (IR) obviously correspond to trivial formal proofs of the form $\mathbf{\Phi} \vdash \varphi$. We consider the more interesting elimination rule (E$\vee$). Suppose that $\mathbf{\Phi} \vdash \varphi \vee \psi, \mathbf{\Phi} + \varphi \vdash \chi$ and $\mathbf{\Phi} + \psi \vdash \chi$. We verify that $\mathbf{\Phi} \vdash \chi$.

φ_0:	$\varphi \to \chi$	from $\mathbf{\Phi} + \varphi \vdash \chi$ by (DT)
φ_1:	$\psi \to \chi$	from $\mathbf{\Phi} + \psi \vdash \chi$ by (DT)
φ_2:	$(\varphi \to \chi) \to ((\psi \to \chi) \to ((\varphi \vee \psi) \to \chi))$	instance of L_8
φ_3:	$(\psi \to \chi) \to ((\varphi \vee \psi) \to \chi)$	from φ_2 and φ_0 by (MP)
φ_4:	$(\varphi \vee \psi) \to \chi$	from φ_3 and φ_1 by (MP)
φ_5:	$\varphi \vee \psi$	by assumption
φ_6:	χ	from φ_4 and φ_5 by (MP)

The corresponding introduction rule (I$\vee$) follows directly from L_6 and L_7 using (DT). Note that (I$\to$) follows directly from (DT) and (E$\to$) from (MP).

We further prove the rules for negation. For (I¬) suppose that $\mathbf{\Phi} + \varphi \vdash \psi \wedge \neg\psi$. It follows from (E∧) that $\mathbf{\Phi} + \varphi \vdash \psi$ and $\mathbf{\Phi} + \varphi \vdash \neg\psi$. We prove that $\mathbf{\Phi} + \varphi \vdash \neg\varphi$, since then $\mathbf{\Phi} \vdash \neg\varphi$ by (E∨) and L_0. We have:

φ_0:	$\neg\psi \to (\psi \to \neg\varphi)$	instance of L_9
φ_1:	$\neg\psi$	by assumption
φ_2:	$\psi \to \neg\varphi$	from φ_0 and φ_1 by (MP)
φ_3:	ψ	by assumption
φ_4:	$\neg\varphi$	from φ_2 and φ_3 by (MP)

The corresponding elimination rule (E¬) follows from Example 2.2. Finally, we prove (I∃) and (E∃). Note that (I∃) follows directly from L_{11} using (DT) and (MP). For (E∃), let ν be a variable such that $\nu \notin \mathrm{free}(\chi)$ for any $\chi \in \mathbf{\Phi}$ and suppose that $\mathbf{\Phi} \vdash \exists\nu\varphi(\nu)$ and $\mathbf{\Phi} + \varphi(\nu) \vdash \psi$. An application of (DT) then yields $\mathbf{\Phi} \vdash \varphi(\nu) \to \psi$. Then we obtain $\mathbf{\Phi} \vdash \psi$ by the following formal proof:

φ_0:	$\forall\nu(\varphi(\nu) \to \psi) \to (\exists\nu\varphi(\nu) \to \psi)$	instance of L_{13}
φ_1:	$\varphi(\nu) \to \psi$	by assumption
φ_2:	$\forall\nu(\varphi(\nu) \to \psi)$	from φ_1 by (∀)
φ_3:	$\exists\nu\varphi(\nu) \to \psi$	from φ_0 and φ_2 by (MP)
φ_4:	$\exists\nu\varphi(\nu)$	by assumption
φ_5:	ψ	from φ_3 and φ_4 by (MP)

This completes the proof of (E∃). The verification of the other rules of the calculus of natural deduction are left to the reader (see Exercise 2.2).

Conversely, we need to check that the calculus of natural deduction proves the logical axioms L_0–L_{16} as well as the inference rules (MP) and (∀). Observe that (MP) corresponds to (E→) and (∀) corresponds to (I∀). As before, we only present the proof for some axioms and leave the others to the reader. We consider first L_9. We need to check that $\vdash \neg\varphi \to (\varphi \to \psi)$.

$\{\neg\varphi, \varphi, \neg\psi\}$	$\vdash \varphi$	by (IR)
$\{\neg\varphi, \varphi, \neg\psi\}$	$\vdash \neg\varphi$	by (IR)
$\{\neg\varphi, \varphi, \neg\psi\}$	$\vdash \varphi \wedge \neg\varphi$	by (I∧)
$\{\neg\varphi, \varphi\}$	$\vdash \neg\neg\psi$	by (I¬)
$\{\neg\varphi, \varphi\}$	$\vdash \psi$	by (E¬)
$\{\neg\varphi\}$	$\vdash \varphi \to \psi$	by (I→)
	$\vdash \neg\varphi \to (\varphi \to \psi)$	by (I→)

Secondly, we derive Axiom L_{13} using the calculus of natural deduction, i.e., we show $\vdash \forall\nu(\varphi(\nu) \to \psi) \to (\exists\nu\varphi(\nu) \to \psi)$:

$\{\forall\nu(\varphi(\nu) \to \psi), \exists\nu\varphi(\nu), \varphi(\nu)\}$	$\vdash \varphi(\nu)$	by (IR)
$\{\forall\nu(\varphi(\nu) \to \psi), \exists\nu\varphi(\nu), \varphi(\nu)\}$	$\vdash \exists\nu\varphi(\nu)$	by (IR)
$\{\forall\nu(\varphi(\nu) \to \psi), \exists\nu\varphi(\nu), \varphi(\nu)\}$	$\vdash \forall\nu(\varphi(\nu) \to \psi)$	by (IR)
$\{\forall\nu(\varphi(\nu) \to \psi), \exists\nu\varphi(\nu), \varphi(\nu)\}$	$\vdash \varphi(\nu) \to \psi$	by (E∀)
$\{\forall\nu(\varphi(\nu) \to \psi), \exists\nu\varphi(\nu), \varphi(\nu)\}$	$\vdash \psi$	by (E→)
$\{\forall\nu(\varphi(\nu) \to \psi), \exists\nu\varphi(\nu)\}$	$\vdash \psi$	by (E∃)
$\{\forall\nu(\varphi(\nu) \to \psi)\}$	$\vdash \exists\nu\varphi(\nu) \to \psi$	by (I→)
	$\vdash \forall\nu(\varphi(\nu) \to \psi) \to (\exists\nu\varphi(\nu) \to \psi)$	by (I→).

The other axioms can be verified in a similar way. ⊣

Methods of Proof

The inference rules of the calculus of natural deduction are very useful because they resemble methods of proof which are commonly used in mathematics. For example, the elimination rule (E$\vee$) mimicks proofs by case distinction: Under the assumption that $\mathbf{\Phi} \vdash \varphi \vee \psi$, one can prove a formula χ by separately proving $\mathbf{\Phi} + \varphi \vdash \chi$ and $\mathbf{\Phi} + \psi \vdash \chi$.

In the following, we list several methods of proof such as proofs by contradiction, contraposition and case distinction.

Proposition 2.5 (Proof by Cases). *Let $\mathbf{\Phi}$ be a set of formulae and let φ, ψ, χ be some formulae. Then the following two statements hold:*

$$\mathbf{\Phi} \vdash \varphi \vee \psi \text{ and } \mathbf{\Phi} + \varphi \vdash \chi \text{ and } \mathbf{\Phi} + \psi \vdash \chi \implies \mathbf{\Phi} \vdash \chi \qquad (\vee 0)$$

$$\mathbf{\Phi} + \varphi \vdash \chi \text{ and } \mathbf{\Phi} + \neg\varphi \vdash \chi \implies \mathbf{\Phi} \vdash \chi \qquad (\vee 1)$$

Proof. Note that ($\vee$0) is exactly the statement of (E$\vee$) and ($\vee$1) is a special case of ($\vee$0), since $\mathbf{\Phi} \vdash \varphi \vee \neg\varphi$ by L_0. ⊣

Corollary 2.6 (Generalised Proof by Cases). *Let $\mathbf{\Phi}$ be a set of formulae and let $\psi_0, \ldots, \psi_n, \varphi$ be some formulae. Then we have:*

$$\mathbf{\Phi} \vdash \psi_0 \vee \cdots \vee \psi_n \text{ and } \mathbf{\Phi} + \psi_i \vdash \varphi \text{ for all } i \leq n \implies \mathbf{\Phi} \vdash \varphi$$

Since Corollary 2.6 is just a generalization of ($\vee$0), we will denote all instances of this form by ($\vee$0) as well.

Proof of Corollary 2.6. We proceed by induction on $n \geq 1$. For $n = 1$ the statement is exactly ($\vee$0). Now assume that $\mathbf{\Phi} \vdash \psi_0 \vee \ldots \vee \psi_n \vee \psi_{n+1}$ and $\mathbf{\Phi} + \psi_i \vdash \varphi$ for all $i \leq n+1$. Let $\mathbf{\Phi}' :\equiv \mathbf{\Phi} + \psi_0 \vee \ldots \vee \psi_n$ and observe that $\mathbf{\Phi}' \vdash \psi_0 \vee \ldots \vee \psi_n$ and $\mathbf{\Phi}' + \psi_i \vdash \varphi$, so by induction hypothesis $\mathbf{\Phi}' \vdash \varphi$. By the Deduction Theorem this implies $\mathbf{\Phi} \vdash \psi_0 \vee \ldots \vee \psi_n \to \varphi$. Moreover, by another application of (DT) we also have $\mathbf{\Phi} \vdash \psi_{n+1} \to \varphi$. Using L_8 and twice (DT), we obtain $\mathbf{\Phi} \vdash \psi_0 \vee \ldots \vee \psi_n \vee \psi_{n+1} \to \varphi$, hence (DT) yields the claim. ⊣

Proposition 2.7 (Ex Falso Quodlibet). *Let $\mathbf{\Phi}$ be a set of formulae and let φ an arbitrary formula. Then for every $\mathscr{L}$-formula ψ we have:*

$$\mathbf{\Phi} \vdash \varphi \wedge \neg\varphi \implies \mathbf{\Phi} \vdash \psi \qquad (🕮)$$

Proof. Let ψ be any formula and assume that $\mathbf{\Phi} \vdash \varphi \wedge \neg\varphi$ for some formula φ. By (E$\wedge$) we have $\mathbf{\Phi} \vdash \varphi$ and $\mathbf{\Phi} \vdash \neg\varphi$. Now the instance $\neg\varphi \to (\varphi \to \psi)$ of the logical axiom L_9 and two applications of Modus Ponens imply $\mathbf{\Phi} \vdash \psi$. ⊣

Notice that Proposition 2.7 implies that if we can derive a contradiction from $\mathbf{\Phi}$, we can derive *every* formula we like, even the impossible, denoted by the symbol

$$\text{📖}.$$

This is closely related to proofs by contradiction:

Corollary 2.8 (Proof by Contradiction). *Let $\mathbf{\Phi}$ be a set of formulae, and φ be an arbitrary formula. Then the following statements hold:*

$$\mathbf{\Phi} + \neg\varphi \vdash \text{📖} \quad \Longrightarrow \quad \mathbf{\Phi} \vdash \varphi,$$
$$\mathbf{\Phi} + \varphi \vdash \text{📖} \quad \Longrightarrow \quad \mathbf{\Phi} \vdash \neg\varphi$$

Proof. Note that the second statement is exactly the introduction rule ($\mathsf{I}\neg$). For the first statment, note that by ($\vee 1$) it is enough to check $\mathbf{\Phi} + \varphi \vdash \varphi$ and $\mathbf{\Phi} + \neg\varphi \vdash \varphi$. The first condition is clearly satisfied and the second one follows directly from ($\mathsf{I}\wedge$) and (📖). ⊣

Proposition 2.9 (Proof by Contraposition). *Let $\mathbf{\Phi}$ be a set of formulae and φ and ψ two arbitrary formulae. Then we have:*

$$\mathbf{\Phi} + \varphi \vdash \psi \quad \Longleftrightarrow \quad \mathbf{\Phi} + \neg\psi \vdash \neg\varphi \tag{CP}$$

Proof. Suppose first that $\mathbf{\Phi} + \varphi \vdash \psi$. Then by ($\mathsf{I}\wedge$), $\mathbf{\Phi} \cup \{\neg\psi, \varphi\} \vdash \psi \wedge \neg\psi$ and hence by ($\mathsf{I}\neg$) we obtain $\mathbf{\Phi} + \neg\psi \vdash \neg\varphi$. Conversely, assume that $\mathbf{\Phi} + \neg\psi \vdash \neg\varphi$. A similar argument as above yields $\mathbf{\Phi} + \varphi \vdash \neg\neg\psi$. An application of ($\mathsf{E}\neg$) completes the proof. ⊣

Note that Proposition 2.9 proves the logical equivalence

$$\varphi \to \psi \Leftrightarrow \neg\psi \to \neg\varphi.$$

Theorem 2.10 (Generalised Deduction Theorem). *If $\mathbf{\Phi}$ is an arbitrary set of formulae and $\mathbf{\Phi} \cup \{\psi_1, \ldots, \psi_n\} \vdash \varphi$, then $\mathbf{\Phi} \vdash (\psi_1 \wedge \cdots \wedge \psi_n) \to \varphi$; and vice versa:*

$$\mathbf{\Phi} \cup \{\psi_1, \ldots, \psi_n\} \vdash \varphi \quad \Longleftrightarrow \quad \mathbf{\Phi} \vdash (\psi_1 \wedge \cdots \wedge \psi_n) \to \varphi \tag{GDT}$$

Proof. This follows immediately from the Deduction Theorem and the DeMorgan's Laws (see Exercise 2.6). ⊣

The Normal Forms NNF & DNF

In many proofs it is convenient to convert a formula into an equivalent formula in some normal form. The simplest normal form is the following: A formula

is said to be in **Negation Normal Form**, denoted by NNF, if it does not contain the implication symbol $\rightarrow$ and if the negation symbol $\neg$ only occurs directly in front of atomic subformulae.

THEOREM 2.11 (NEGATION NORMAL FORM THEOREM). *Every formula is equivalent to some formula in* NNF.

Proof. We successively apply the following transformations to every non-atomic negated subformula ψ of φ, starting with the outermost negation symbols.

- Using TAUTOLOGY (K), we may replace subformulae of the form $\psi_1 \rightarrow \psi_2$ by $\neg\psi_1 \vee \psi_2$.
- If $\psi \equiv \neg\neg\psi'$ for some formula ψ', we replace ψ with ψ' using TAUTOLOGY (F).
- By the DEMORGAN'S LAWS (see EXERCISE 2.6), we replace subformulae of the form $\neg(\psi_1 \wedge \psi_2)$ and $\neg(\psi_1 \vee \psi_2)$, respectively, with $\neg\psi_1 \vee \neg\psi_2$ and $\neg\psi_1 \wedge \neg\psi_2$, respectively.
- If $\psi \equiv \neg\exists\nu\psi'$ then it follows from TAUTOLOGY (Q.0) that $\psi \Leftrightarrow \forall\nu\neg\psi'$, and hence we replace ψ with $\forall\nu\neg\psi'$. Similarly, using TAUTOLOGY (Q.1), we replace subformulae of the form $\neg\forall\nu\psi'$ with the equivalent formula $\exists\nu\neg\psi'$.

$\dashv$

A quantifier-free formula φ is said to be in **Disjunctive Normal Form**, denoted DNF, if it is a disjunction of conjunctions of atomic formulae or negated atomic formulae, i.e., it is of the form

$$(\varphi_{1,1} \wedge \ldots \wedge \varphi_{1,k_1}) \vee \cdots \vee (\varphi_{m,1} \wedge \cdots \wedge \varphi_{m,k_m})$$

for some quantifier-free formulae $\varphi_{i,j}$ which are either atomic or the negation of an atomic formula. In particular, each formula in DNF is also in NNF.

THEOREM 2.12 (DISJUNCTIVE NORMAL FORM THEOREM). *Every quantifier-free formula φ is equivalent to some formula in DNF.*

Proof. By the NEGATION NORMAL FORM THEOREM we may assume that φ is in NNF. Starting with the outermost conjunction symbol, we successively apply the distributive laws

$$\psi \wedge (\varphi_1 \vee \varphi_2) \Leftrightarrow (\psi \wedge \varphi_1) \vee (\psi \wedge \varphi_2) \quad \text{and}$$
$$(\varphi_1 \vee \varphi_2) \wedge \psi \Leftrightarrow (\varphi_1 \wedge \psi) \vee (\varphi_2 \wedge \psi)$$

until all conjunction symbols occur between atomic or negated atomic formulae. This process ends after FINITELY many steps, since there are only FINITELY many conjunction symbols. $\dashv$

A similar result holds for the so-called *Conjunctive Normal Form*, denoted CNF (see Exercise 2.8).

Substitution of Variables and the Prenex Normal Form

In Part II & III, we shall encode formulae by strings of certain symbols and by natural numbers, respectively. In order to do so, we have to make sure that the variables are among a well-defined set of symbols, namely among $v_0, v_1, \ldots$, where the index n of v_n is a natural number (i.e., a member of $\mathbb{N}$).

Theorem 2.13 (Variable Substitution Theorem). *For every sentence σ there is an equivalent sentence $\tilde{\sigma}$ which contains just variables among $v_0, v_1, \ldots$, where for any $m, n \in \mathbb{N}$ with $m < n$, if v_n appears in $\tilde{\sigma}$, then also v_m appears in $\tilde{\sigma}$.*

Proof. Let σ be an arbitrary sentence and let m be such that no variable v_k with $k \geq m$ appears in σ. Assume that $\exists\nu\varphi(\nu)$ is a sub-sentence of σ. Then $\exists\nu\varphi(\nu) \Leftrightarrow \exists v_k\varphi(\nu/v_k)$ for any $k \geq m$. To see this, first notice that since v_k does not appear in σ, the substitution $\varphi(\nu/v_k)$ is admissible. Furthermore, we have:

φ_0:	$\varphi(v_k) \to \exists\nu\varphi(\nu)$	instance of L_{11}
φ_1:	$\forall v_k\big(\varphi(v_k) \to \exists\nu\varphi(\nu)\big)$	from φ_0 by $(\forall)$
φ_2:	$\forall v_k\big(\varphi(v_k) \to \exists\nu\varphi(\nu)\big) \to \big(\exists v_k\varphi(v_k) \to \exists\nu\varphi(\nu)\big)$	instance of L_{13}
φ_3:	$\exists v_k\varphi(v_k) \to \exists\nu\varphi(\nu)$	from φ_2 and φ_1 by (MP)

Similarly, we obtain $\exists\nu\varphi(\nu) \to \exists v_k\varphi(v_k)$, which shows that $\exists\nu\varphi(\nu) \Leftrightarrow \exists v_k\varphi(v_k)$.

Assume now that $\forall\nu\varphi(\nu)$ is a sub-sentence of σ. Then $\forall\nu\varphi(\nu) \Leftrightarrow \forall v_k\varphi(\nu/v_k)$. Since the substitution $\varphi(\nu/v_k)$ is admissible, we have:

φ_0:	$\forall\nu\varphi(\nu) \to \varphi(v_k)$	instance of L_{10}
φ_1:	$\forall v_k\big(\forall\nu\varphi(\nu) \to \varphi(v_k)\big)$	from φ_0 by $(\forall)$
φ_2:	$\forall v_k\big(\forall\nu\varphi(\nu) \to \varphi(v_k)\big) \to \big(\forall\nu\varphi(\nu) \to \forall v_k\varphi(v_k)\big)$	instance of L_{12}
φ_3:	$\forall\nu\varphi(\nu) \to \forall v_k\varphi(v_k)$	from φ_2 and φ_1 by (MP)

Similarly we obtain $\forall v_k\varphi(v_k) \to \forall\nu\varphi(\nu)$, which shows that $\forall\nu\varphi(\nu) \Leftrightarrow \forall v_k\varphi(v_k)$.

Therefore, we can replace all the variables $\nu_0, \nu_1, \ldots$ appearing in σ step by step with variables $v_m, v_{m+1}, \ldots$ and obtain a sentence σ' which is equivalent

to σ. In a last step, we replace the variables $v_m, v_{m+1}, \ldots$ with $v_0, v_1, \ldots$ and finally obtain $\tilde{\sigma}$. ⊣

In Chapter 5, we will use the fact that every sentence can be transformed into a semantically equivalent sentence in the so-called *special Prenex Normal Form*:

A sentence σ is said to be in **Prenex Normal Form**, denoted by PNF, if it is of the form

$$\nexists_0 \nu_0 \ldots \nexists_n \nu_n \tilde{\sigma},$$

where the variables $\nu_0, \ldots, \nu_n$ are pairwise distinct, each $\nexists_m$ (for $0 \leq m \leq n$) stands for either $\exists$ or for $\forall$, and $\tilde{\sigma}$ is a quantifier-free formula. Furthermore, a sentence σ is in **special Prenex Normal form**, denoted by sPNF, if σ is in PNF and

$$\sigma \equiv \nexists_0 v_0 \nexists_1 v_1 \ldots \nexists_n v_n \tilde{\sigma}\,,$$

where each $\nexists_m$ (for $0 \leq m \leq n$) stands for either $\exists$ or $\forall$, $\tilde{\sigma}$ is quantifier-free, and in addition, each variable $v_0, \ldots, v_n$ appears free in $\tilde{\sigma}$.

Theorem 2.14 (Prenex Normal Form Theorem). *For every sentence σ there is an equivalent sentence $\tilde{\sigma}$ in* sPNF.

Sketch of the Proof. Using the Negation Normal Form Theorem we may suppose that φ is in NNF. Moreover, by Tautologies (O.1) and (O.2) (see in the proof of the Variable Substitution Theorem) we may additionally suppose that no variable is quantified more than once. The crucial part is now to show that for all formulae φ and ψ, where $\nu \notin \mathrm{free}(\psi)$, the following formulae are tautologies:

$$\exists\nu\varphi \circ \psi \;\Leftrightarrow\; \exists\nu\big(\varphi \circ \psi\big)\,,$$
$$\forall\nu\varphi \circ \psi \;\Leftrightarrow\; \forall\nu\big(\varphi \circ \psi\big)\,,$$

where $\circ$ stands for either $\vee$ or $\wedge$. We just prove that the formula

$$\exists\nu\varphi \vee \psi \;\rightarrow\; \exists\nu\big(\varphi \vee \psi\big)\,, \quad \text{where } \nu \notin \mathrm{free}(\psi),$$

is a tautology; all other cases are proved similarly.

By L_8, we obtain

$$\vdash \big(\exists\nu\varphi \rightarrow \exists\nu(\varphi \vee \psi)\big) \rightarrow \Big(\big(\psi \rightarrow \exists\nu(\varphi \vee \psi)\big) \rightarrow \big((\exists\nu\varphi \vee \psi) \rightarrow \exists\nu(\varphi \vee \psi)\big)\Big)\,.$$

Therefore, it is enough to show:

$$\vdash \exists \nu \varphi \to \exists \nu(\varphi \vee \psi) \quad \text{and} \quad \vdash \psi \to \exists \nu(\varphi \vee \psi)$$

We first prove $\vdash \exists \nu \varphi \to \exists \nu(\varphi \vee \psi)$:

φ_0:	$\varphi \to \varphi \vee \psi$	instance of L_6
φ_1:	$\varphi \vee \psi \to \exists \nu(\varphi \vee \psi)$	instance of L_{11}
φ_2:	$\varphi \to \exists \nu(\varphi \vee \psi)$	by Tautology (D.0)
φ_3:	$\forall \nu\big(\varphi \to \exists \nu(\varphi \vee \psi)\big)$	from φ_2 by $(\forall)$
φ_4:	$\forall \nu\big(\varphi \to \exists \nu(\varphi \vee \psi)\big) \to \big(\exists \nu \varphi \to \exists \nu(\varphi \vee \psi)\big)$	instance of L_{13}
φ_5:	$\exists \nu \varphi \to \exists \nu(\varphi \vee \psi)$	from φ_4 and φ_3 by (MP)

Now we show $\vdash \psi \to \exists \nu(\varphi \vee \psi)$:

φ_0:	$\psi \to \varphi \vee \psi$	instance of L_7
φ_1:	$\varphi \vee \psi \to \exists \nu(\varphi \vee \psi)$	instance of L_{11}
φ_2:	$\psi \to \exists \nu(\varphi \vee \psi)$	by Tautology (D.0)

After finitely many applications of the above tautologies, we obtain a sentence in PNF. Moreover, the Variable Substitution Theorem yields a formula in sPNF, i.e., a formula in which the quantified variables are $v_0, \ldots, v_n$ and in the quantifier part of the formula, the variables appear in the order $v_0, \ldots, v_n$.

⊣

Semi-formal Proofs

Previously, we have shown that formal proofs can be simplified by applying methods of proof such as case distinctions, proofs by contradiction or contraposition. However, in order to make proofs even more natural, it is useful to use natural language for describing a proof step as in an "informal" mathematical proof.

Example 2.15. We want to prove the tautology $\vdash \varphi \to \neg\neg\varphi$. Instead of writing out the whole formal proof, which is quite tedious, we can apply the methods of proof which we introduced above.

The first modification we make is to use (DT) to obtain the new goal

$$\{\varphi\} \vdash \neg\neg\varphi.$$

The easiest way to proceed is to make a proof by contradiction; hence it remains to show

$$\{\varphi, \neg\varphi\} \vdash \text{🕮}$$

which by $(\mathsf{I}\wedge)$ is again a consequence of the trivial goals

$$\{\varphi, \neg\varphi\} \vdash \varphi \quad \text{and} \quad \{\varphi, \neg\varphi\} \vdash \neg\varphi.$$

To sum up, this procedure can actually be transformed back into a formal proof, so it suffices as a proof of $\vdash \varphi \to \neg\neg\varphi$. Now this is still not completely satisfactory, since we would like to write the proof in natural language. A possible translation could thus be the following:

Semi-formal Proof. We want to prove that φ implies $\neg\neg\varphi$. We assume φ, and for a contradiction we assume also $\neg\varphi$. Then we have φ and $\neg\varphi$, which is a contradiction. Hence, we must have $\neg\neg\varphi$, i.e., $\varphi \to \neg\neg\varphi$. ⊣

We will now show in a systematic way how formal proofs can — in principle — be replaced by semi-formal proofs, which make use of a **controlled natural language**, i.e., a limited vocabulary consisting of natural language phrases such as "assume that" which are often used in mathematical proof texts. This language is controlled in the sense that its allowed vocabulary is only a subset of the entire English vocabulary and that every word and every phrase, respectively, has a unique precisely defined interpretation. However, for the sake of a nice proof style, we will not always stick to this limited vocabulary. Moreover, this section should be considered as a hint of how formal proofs can be formulated using a controlled natural language as well as a justification for working with natural language proofs rather than formal ones.

Every statement which we would like to prove formally is of the form $\mathbf{\Phi} \vdash \varphi$, where $\mathbf{\Phi}$ is a set of formulae and φ is a formula. Note that as in Example 2.15, in order to prove $\mathbf{\Phi} \vdash \varphi$ — which is actually a meta-proof — we perform operations both on the set of formulae $\mathbf{\Phi}$ and on the formula to be formally proved. We call a statement of the form $\mathbf{\Phi} \vdash \varphi$ a **goal**, the set $\mathbf{\Phi}$ is called **premises**, and the formula φ to be verified as **target**. Now instead of listing a formal proof, we can step by step reduce our current goal to a simpler one using the methods of proof from the previous section, until the target is tautological as in the case of Example 2.15.

In that sense, methods of proof are in that sense simply operations on the premises and the targets. For example, the proof by contraposition for example adds the negation of the target to the premises and replaces the original target by the negation of the premise from which it shall be derived:

If we want to show

$$\mathbf{\Phi} + \psi \vdash \varphi$$

we can prove $\mathbf{\Phi} + \neg\varphi \vdash \neg\psi$ instead A slightly different example is the proof of a conjunction $\mathbf{\Phi} \vdash \varphi \wedge \psi$, which is usually split into the two goals given by

$$\mathbf{\Phi} \vdash \varphi \quad \text{and} \quad \mathbf{\Phi} \vdash \psi.$$

Thus we have to revise our first attempt and interpret methods of proof as operations on FINITE lists of goals consisting of premises and targets.

We distinguish between two types of operations on goals: **Backward reasoning** means performing operations on targets, whereas **forward reasoning** denotes operations on the premises. We give some examples of both backward and forward reasoning and indicate how such proofs can be phrased in a semi-formal way.

Backward reasoning

- Targets are often of the universal conditional form $\forall\nu(\varphi(\nu) \to \psi(\nu))$. In particular, this pattern includes the purely universal formulae $\forall\nu\psi(\nu)$ by taking φ to be a tautology as well as simple conditionals of the form $\varphi \to \psi$. Now the usual procedure is to reduce $\mathbf{\Phi} \vdash \forall\nu(\varphi(\nu) \to \psi(\nu))$ to $\mathbf{\Phi} + \varphi(\nu) \vdash \psi(\nu)$ using $(\forall)$ and (DT). This can be rephrased as

  Assume $\varphi(\nu)$. *Then* ... *This shows* $\psi(\nu)$.

- As already mentioned above, if the target is a conjunction $\varphi \wedge \psi$, then one can show the conjuncts separately using $(\mathsf{I}\wedge)$. This step is usually executed without mentioning it explicitly.
- If the target is a negation $\neg\varphi$, one often uses a proof by contradiction or by contraposition: In the first case, we transform $\mathbf{\Phi} \vdash \neg\varphi$ to $\mathbf{\Phi} + \varphi \vdash \lightning$ and use the natural language notation

  Suppose for a contradiction that φ. *Then* ... *Contradiction.*

  In the latter case, we want to go from $\mathbf{\Phi} + \neg\psi \vdash \neg\varphi$ to $\mathbf{\Phi} + \varphi \vdash \psi$ or, in its positive version, from $\mathbf{\Phi} + \psi \vdash \neg\varphi$ to $\mathbf{\Phi} + \varphi \vdash \neg\psi$, respectively. In both cases, we can mark this with the keyword *contraposition*, e.g., as

  We proceed by contraposition ... *This shows* $\neg\varphi$.

Forwards reasoning

- By $(\mathsf{E}\wedge)$, conjunctive premises $\varphi \wedge \psi$ can be split into two premises φ, ψ; i.e., $\mathbf{\Phi} + \varphi \wedge \psi \vdash \chi$ can be reduced to $\mathbf{\Phi} \cup \{\varphi, \psi\} \vdash \chi$. This is usually performed automatically.
- Disjunctive premises are used for proofs by case distinction: If a goal of the form $\mathbf{\Phi} + \varphi \vee \psi \vdash \chi$ is given, we can reduce it to the new goals $\mathbf{\Phi} + \varphi \vdash \chi$ and $\mathbf{\Phi} + \psi \vdash \chi$. We can write this in a semi-formal way as

  Case 1: Assume φ ... *This proves* χ.
  Case 2: Assume ψ ... *This proves* χ.

- Intermediate proof steps: Often we first want to prove some intermediate statement which shall then be applied in order to resolve the target. Formally, this means that we want to show $\mathbf{\Phi} \vdash \varphi$ by first showing $\mathbf{\Phi} \vdash \psi$ and then adding ψ to the list of premises and checking $\mathbf{\Phi} + \psi \vdash \varphi$. Clearly, if we have $\mathbf{\Phi} \vdash \psi$ and $\mathbf{\Phi} + \psi \vdash \varphi$, using (DT) and (MP) we obtain that $\mathbf{\Phi} \vdash \varphi$. In a semi-formal proof, this can be described by

  We first show ψ ... This proves φ.

  Note that it is important to mark where the proof of the intermediate statement ψ ends, since from this point on, ψ can be used as a new premise.

Observe that in any case, once a goal $\mathbf{\Phi} \vdash \varphi$ is reduced to a tautology, it can be removed from the list of goals. This should be marked by a phrase like

This shows/proves φ.

so that it is clear that we move on to the next goal. The proof is complete as soon as no unresolved goals remain.

What is the use of such a formalised natural proof language? First of all, it increases readability. Secondly, by giving a precise formal definition to some of the common natural language phrases which appear in the proof texts, we show how — in principle — one could write formal proofs with a controlled natural language input. This input could then be parsed into a formal proof and subsequently be verified by a proof checking system.

We would like to emphasize that this section should only be considered as a motivation rather than a precise description of how formal proofs can be translated into semi-formal ones and vice versa. Nevertheless, it suffices to understand how this can theoretically be achieved. Therefore, in subsequent chapters, especially in Chapters 8 and 9, we will often present semi-formal proofs rather than formal ones.

Consistency & Compactness

We say that be a set of formulae $\mathbf{\Phi}$ is **consistent**, denoted $\mathrm{Con}(\mathbf{\Phi})$, if $\mathbf{\Phi} \nvdash \text{🕮}$, i.e., if there is *no* formula φ such that $\mathbf{\Phi} \vdash (\varphi \wedge \neg\varphi)$, otherwise $\mathbf{\Phi}$ is called **inconsistent**, denoted $\neg\,\mathrm{Con}(\mathbf{\Phi})$.

Fact 2.16. *Let $\mathbf{\Phi}$ be a set of formulae.*

(a) *If $\neg\,\mathrm{Con}(\mathbf{\Phi})$, then for all formulae ψ we have $\mathbf{\Phi} \vdash \psi$.*

(b) *If $\mathrm{Con}(\mathbf{\Phi})$ and $\mathbf{\Phi} \vdash \varphi$ for some formula φ, then $\mathbf{\Phi} \nvdash \neg\varphi$.*

(c) *If $\neg\,\mathrm{Con}(\mathbf{\Phi} + \varphi)$, for some formula φ, then $\mathbf{\Phi} \vdash \neg\varphi$.*

(d) *If $\mathbf{\Phi} \vdash \neg\varphi$, for some formula φ, then $\neg\,\mathrm{Con}(\mathbf{\Phi} + \varphi)$.*

Proof. Condition (a) is just PROPOSITION 2.7. For (b), notice that if $\mathbf{\Phi} \vdash \varphi$ and $\mathbf{\Phi} \vdash \neg\varphi$, then by $(\mathsf{I}\wedge)$ we get $\mathbf{\Phi} \vdash \square$ and thus also $\neg\,\mathrm{Con}(\mathbf{\Phi})$. Moreover, (c) coincides with the second statement of COROLLARY 2.8. Finally, for (d) note that if $\mathbf{\Phi} \vdash \neg\varphi$, then $\mathbf{\Phi} + \varphi \vdash \varphi \wedge \neg\varphi$ and hence $\mathbf{\Phi} + \varphi$ is inconsistent. ⊣

If we choose a set of formulae $\mathbf{\Phi}$ as the basis of a theory (e.g., a set of axioms), we have to make sure that $\mathbf{\Phi}$ is consistent. However, as we shall see later, in many cases this task is impossible.

We conclude this chapter with the COMPACTNESS THEOREM, which is a powerful tool in order to construct non-standard models of Peano Arithmetic or of Set Theory. On the one hand, it is just a consequence of the fact that formal proofs are FINITE sequences of formulae. On the other hand, the COMPACTNESS THEOREM is the main tool to prove that a given set of sentences is consistent with some given set of formulae $\mathbf{\Phi}$.

THEOREM 2.17 (COMPACTNESS THEOREM). *Let $\mathbf{\Phi}$ be an arbitrary set of formulae. Then $\mathbf{\Phi}$ is consistent if and only if every finite subset $\mathbf{\Phi}'$ of $\mathbf{\Phi}$ is consistent.*

Proof. Obviously, if $\mathbf{\Phi}$ is consistent, then every finite subset $\mathbf{\Phi}'$ of $\mathbf{\Phi}$ must be consistent. On the other hand, if $\mathbf{\Phi}$ is inconsistent, then there is a formula φ such that $\mathbf{\Phi} \vdash \varphi \wedge \neg\varphi$. In other words, there is a proof of $\varphi \wedge \neg\varphi$ from $\mathbf{\Phi}$. Now, since every proof is finite, there are only finitely many formulae of $\mathbf{\Phi}$ involved in this proof, and if $\mathbf{\Phi}'$ is this finite set of formulae, then $\mathbf{\Phi}' \vdash \varphi \wedge \neg\varphi$, which shows that $\mathbf{\Phi}'$, a finite subset of $\mathbf{\Phi}$, is inconsistent. ⊣

NOTES

Natural deduction in its modern form was developed by the German mathematician Gentzen in 1934 (see [11, 12]).

EXERCISES

2.0 (a) Show that quantifier-free formulae can be written with the only logical operator $\tilde{\wedge}$, where

$$\varphi \,\tilde{\wedge}\, \psi :\Longleftrightarrow \neg(\varphi \wedge \psi)\,.$$

(b) Show that quantifier-free formulae can be written with the only logical operator $\tilde{\vee}$, where

$$\varphi \,\tilde{\vee}\, \psi :\Longleftrightarrow \neg(\varphi \vee \psi)\,.$$

2.1 Show that logical equivalence $\Leftrightarrow$ defines an equivalence relation on the set of formulae.

2.2 Complete the proof of THEOREM 2.4.

2.3 Formalise the method of proof by counterexample and prove that it works.

2.4 Show that the TAUTOLOGIES (L.0) and (R) are provable from the logical axioms.

2.5 Prove the equivalences in the proof of THEOREM 1.7.

2.6 Let $\varphi_0, \dots, \varphi_n$ be formulae. Prove the DEMORGAN'S LAWS:

(a) $\neg(\varphi_0 \wedge \cdots \wedge \varphi_n) \Leftrightarrow (\neg\varphi_1 \vee \cdots \vee \neg\varphi_n)$

(b) $\neg(\varphi_0 \vee \cdots \vee \varphi_n) \Leftrightarrow (\neg\varphi_1 \wedge \cdots \wedge \neg\varphi_n)$

(c) $\varphi_0 \to \big(\varphi_1 \to (\cdots \to \varphi_n)\cdots\big) \Leftrightarrow (\neg\varphi_0 \vee \cdots \vee \neg\varphi_{n-1}) \vee \varphi_n$

2.7 Prove the following generalisation of L_{15} to an arbitrary formula φ:

$$\vdash (\tau_1 = \tau_1' \wedge \cdots \wedge \tau_n = \tau_n') \to \big(\varphi(\tau_1, \dots, \tau_n) \to \varphi(\tau_1', \dots, \tau_n')\big)\,,$$

where $\tau, \tau_1, \dots, \tau_n, \tau_1', \dots, \tau_n'$ are terms and φ is a formula with n free variables.

2.8 A quantifier-free formula φ is said to be in **Conjunctive Normal Form**, denoted CNF, if it is a conjunction of disjunctions of atomic formulae or negated atomic formulae, i.e., it is of the form

$$(\varphi_{1,1} \vee \cdots \vee \varphi_{1,k_1}) \wedge \cdots \wedge (\varphi_{m,1} \vee \cdots \vee \varphi_{m,k_m})$$

for some quantifier-free formulae $\varphi_{i,j}$ which are either atomic or the negation of an atomic formula.

Show that every quantifier-free formula φ is equivalent to some formula in CNF.

2.9 Let $\mathsf{L}_{9^3/_4}$ be the axiom schema $(\varphi \to \psi) \to \big((\varphi \to \neg\psi) \to \neg\varphi\big)$.

(a) Show that $\{\mathsf{L}_0, \mathsf{L}_1, \mathsf{L}_2, \mathsf{L}_8, \mathsf{L}_9\} \vdash \mathsf{L}_{9^3/_4}$, where $\vdash$ here means that we can only use Modus Ponens and the five logical axioms L_0, L_1, L_2, L_8, and L_9 — this applies analogously for the following exercises.

(b) Show that $\{\mathsf{L}_1, \mathsf{L}_2, \mathsf{L}_{9^3/_4}\} \vdash \varphi \to \neg\neg\varphi$.

(c) Show that $\{\mathsf{L}_1\text{–}\mathsf{L}_9\} \nvdash \mathsf{L}_{9^3/_4}$.

Hint: Define a mapping $|\cdot|$, which assigns to each formula φ a value $|\varphi| \in \{-1, 0, 1\}$ such that the following conditions are satisfied: $|\varphi \vee \psi| = \max\{|\varphi|, |\psi|\}$, $|\varphi \wedge \psi| = \min\{|\varphi|, |\psi|\}$, $|\neg\varphi| = -|\varphi|$, and the value of $|\varphi \to \psi|$ is given by the following table:

$\lvert\varphi\rvert$ \ $\lvert\psi\rvert$	-1	0	1	
-1	1	1	1	
0	1	1	1	$\lvert\varphi \to \psi\rvert$
1	-1	0	1	

Show that for every formula θ with $\{\mathsf{L}_1\text{–}\mathsf{L}_9\} \vdash \theta$ we have $|\theta| = 1$. On the other hand, for certain values of $|\varphi|$ and $|\psi|$ we have $|\mathsf{L}_{9^3/_4}| \neq 1$.

2.10 Let $\mathsf{L}_{9^1/_4}$ be the axiom schema $\neg\neg\varphi \to \varphi$.

(a) Show that $\{\mathsf{L}_0, \mathsf{L}_1, \mathsf{L}_2, \mathsf{L}_8, \mathsf{L}_9\} \vdash \mathsf{L}_{9^1/_4}$.

(b) Show that $\{\mathsf{L}_1, \mathsf{L}_2, \mathsf{L}_6, \mathsf{L}_7, \mathsf{L}_{9^3/_4}\} \vdash \neg\neg(\varphi \vee \neg\varphi)$.

(c) Show that $\{\mathsf{L}_1, \mathsf{L}_2, \mathsf{L}_6, \mathsf{L}_7, \mathsf{L}_{9^3/_4}, \mathsf{L}_{9^1/_4}\} \vdash \mathsf{L}_0$.

(d) Show that $\{\mathsf{L}_1\text{–}\mathsf{L}_9, \mathsf{L}_{9^3/_4}\} \nvdash \mathsf{L}_{9^1/_4}$ (compare with EXERCISE 2.9.(b)).

(e) Show that $\{\mathsf{L}_1\text{–}\mathsf{L}_9, \mathsf{L}_{9\,3/4}\} \nvdash \mathsf{L}_0$ (compare with Exercise 2.10.(c)).

Hint for parts (d) *&* (e): As in Exercise 2.9.(c), define a mapping $|\cdot|$, which assigns to each formula φ a value $|\varphi| \in \{-1, 0, 1\}$ such that the following conditions are satisfied: $|\varphi \vee \psi| = \max\{|\varphi|, |\psi|\}$, $|\varphi \wedge \psi| = \min\{|\varphi|, |\psi|\}$, and the values of $|\neg\varphi|$ and $|\varphi \to \psi|$ are given by the following tables:

$\lvert\varphi\rvert$	-1	0	1
$\lvert\neg\varphi\rvert$	1	-1	-1

$\lvert\varphi\rvert$ \ $\lvert\psi\rvert$	-1	0	1
-1	1	1	1
0	-1	1	1
1	-1	0	1

$|\varphi \to \psi|$

Show that for every formula θ with $\{\mathsf{L}_1\text{–}\mathsf{L}_9, \mathsf{L}_{9\,3/4}\} \vdash \theta$ we have $|\theta| = 1$. On the other hand, for certain values of $|\varphi|$ we have $|\mathsf{L}_0| \neq 1$ and $|\mathsf{L}_{9\,1/4}| \neq 1$, respectively.

2.11 Prove Glivenko's Theorem, which states that for every formula φ of Propositional Logic, i.e., for every formula φ which is quantifier-free, we have:

$$\{\mathsf{L}_0\text{–}\mathsf{L}_9\} \vdash \varphi \quad \textit{if and only if} \quad \{\mathsf{L}_1\text{–}\mathsf{L}_9, \mathsf{L}_{9\,3/4}\} \vdash \neg\neg\varphi$$

Chapter 3
Semantics: Making Sense of the Symbols

There are two different views on a given set of formulae $\mathbf{\Phi}$, namely the *syntactical* view and the *semantical* view.

From the syntactical point of view (presented in the previous chapters), we consider the set $\mathbf{\Phi}$ just as a set of well-formed formulae — regardless of their intended sense or meaning — from which we can prove some formulae. So, from a formal point of view there is no need to assign real objects (whatever this means) to our strings of symbols.

In contrast to this very formal syntactical view, there is also the semantical point of view according to which we consider the intended meaning of the formulae in $\mathbf{\Phi}$ and then seek for a *model* in which all formulae of $\mathbf{\Phi}$ become true. For this, we have to explain some basic notions of Model Theory like *structure* and *interpretation*, which we will do in a natural, informal language. In this language, we will use words like "or", "and", or phrases like "if. . .then". These words and phrases have the usual meaning. Furthermore, we assume that in our normal world, which we describe with our informal language, the basic rules of *common logic* apply. For example, a statement A is true or false, and if A is true, then the negation of A, denoted not-A, is false; and vice versa. Hence, the statement "A or not-A" is always true, which means that we tacitly assume the Law Of Excluded Middle, also known as Tertium Non Datur, which corresponds to the logical axiom L_0. Furthermore, we assume DeMorgan's Laws and apply Modus Ponens as an inference rule.

Structures & Interpretations

In order to define structures and interpretations, we have to assume some notions of Naive Set Theory like *subset*, *cartesian product*, or *relation*, which shall be properly defined in Part IV. On this occasion, we

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_4

also make use of the set theoretical symbol $\in$, which stands for the binary *membership relation*.

Let $\mathscr{L}$ be an arbitrary but fixed language. An **$\mathscr{L}$-structure $\mathbf{M}$** consists of a non-empty set A, called the **domain** of $\mathbf{M}$, together with a mapping which assigns to each constant symbol $c \in \mathscr{L}$ an element $c^{\mathbf{M}} \in A$, to each n-ary relation symbol $R \in \mathscr{L}$ a set of n-tuples $R^{\mathbf{M}}$ of elements of A, and to each n-ary function symbol $F \in \mathscr{L}$ a function $F^{\mathbf{M}}$ from n-tuples of A to A. In other words, the constant symbols denote elements of A, n-ary relation symbols denote subsets of A^n (i.e., subsets of the n-fold cartesian product of A), and n-ary functions symbols denote n-ary functions from A^n to A.

The interpretation of variables is given by a so-called assignment: An **assignment** in an $\mathscr{L}$-structure $\mathbf{M}$ is a mapping j which assigns to each variable an element of the domain A.

Finally, an **$\mathscr{L}$-interpretation $\mathbf{I}$** is a pair $(\mathbf{M}, j)$ consisting of an $\mathscr{L}$-structure $\mathbf{M}$ and an assignment j in $\mathbf{M}$. For a variable ν, an element $a \in A$, and an assignment j in $\mathbf{M}$, we define the assignment $j\frac{a}{\nu}$ by stipulating

$$j\tfrac{a}{\nu}(\nu') = \begin{cases} a & \text{if } \nu' \equiv \nu, \\ j(\nu') & \text{otherwise.} \end{cases}$$

Furthermore, for elements $a, a' \in A$ and variables ν, ν', we shall write $j\frac{a}{\nu}\frac{a'}{\nu'}$ instead of $\big(j\frac{a}{\nu}\big)\frac{a'}{\nu'}$.

For an interpretation $\mathbf{I} = (\mathbf{M}, j)$ and an element $a \in A$, we define:

$$\mathbf{I}\tfrac{a}{\nu} := (\mathbf{M}, j\tfrac{a}{\nu})$$

We associate with every interpretation $\mathbf{I} = (\mathbf{M}, j)$ and every $\mathscr{L}$-term τ an element $\mathbf{I}(\tau) \in A$ as follows:

- For a variable ν, let $\mathbf{I}(\nu) := j(\nu)$.
- For a constant symbol $c \in \mathscr{L}$, let $\mathbf{I}(c) := c^{\mathbf{M}}$.
- For an n-ary function symbol $F \in \mathscr{L}$ and terms $\tau_1, \ldots, \tau_n$, let

$$\mathbf{I}\big(F(\tau_1, \ldots, \tau_n)\big) := F^{\mathbf{M}}\big(\mathbf{I}(\tau_1), \ldots, \mathbf{I}(\tau_n)\big)\,.$$

Now, we are able to define precisely when a formula φ becomes *true* under an interpretation $\mathbf{I} = (\mathbf{M}, j)$; in which case we write $\mathbf{I} \vDash \varphi$ and say that φ is **true** in $\mathbf{I}$ (or that φ **holds** in $\mathbf{I}$). The definition is by induction on the complexity of the formula φ, where the truth value of expressions involving "NOT", "AND", "IF ... THEN", et cetera, is explained later. By the rules (F0)–(F4), φ must be of the form $\tau_1 = \tau_2$, $R(\tau_1, \ldots, \tau_n)$, $\neg\psi$, $\psi_1 \wedge \psi_2$, $\psi_1 \vee \psi_2$, $\psi_1 \to \psi_2$, $\exists\nu\psi$, or $\forall\nu\psi$:

$$\mathbf{I} \vDash \tau_1 = \tau_2 \quad :\Longleftrightarrow \quad \mathbf{I}(\tau_1) \text{ IS THE SAME OBJECT AS } \mathbf{I}(\tau_2)$$

$$\mathbf{I} \vDash R(\tau_1, \ldots, \tau_n) \quad :\Longleftrightarrow \quad \big\langle \mathbf{I}(\tau_1), \ldots, \mathbf{I}(\tau_n) \big\rangle \text{ BELONGS TO } R^{\mathbf{M}}$$

$$\begin{aligned}
\mathbf{I} \vDash \neg\psi \quad &:\Longleftrightarrow \quad \text{NOT } \mathbf{I} \vDash \psi \\
\mathbf{I} \vDash \psi_1 \wedge \psi_2 \quad &:\Longleftrightarrow \quad \mathbf{I} \vDash \psi_1 \text{ AND } \mathbf{I} \vDash \psi_2 \\
\mathbf{I} \vDash \psi_1 \vee \psi_2 \quad &:\Longleftrightarrow \quad \mathbf{I} \vDash \psi_1 \text{ OR } \mathbf{I} \vDash \psi_2 \\
\mathbf{I} \vDash \psi_1 \to \psi_2 \quad &:\Longleftrightarrow \quad \text{IF } \mathbf{I} \vDash \psi_1 \text{ THEN } \mathbf{I} \vDash \psi_2 \\
\mathbf{I} \vDash \exists\nu\psi \quad &:\Longleftrightarrow \quad \text{THERE EXISTS } a \text{ IN } A: \ \mathbf{I}^a_\nu \vDash \psi \\
\mathbf{I} \vDash \forall\nu\psi \quad &:\Longleftrightarrow \quad \text{FOR ALL } a \text{ IN } A: \ \mathbf{I}^a_\nu \vDash \psi
\end{aligned}$$

Notice that by the logical rules in our informal language, for *every* $\mathscr{L}$-formula φ we have either $\mathbf{I} \vDash \varphi$ or $\mathbf{I} \vDash \neg\varphi$. So, every $\mathscr{L}$-formula is either true or false in $\mathbf{I}$. On the syntactical level, however, we do not necessarily have $\mathbf{\Phi} \vdash \varphi$ or $\mathbf{\Phi} \vdash \neg\varphi$ for each set $\mathbf{\Phi}$ of $\mathscr{L}$-formulae and each $\mathscr{L}$-formula φ.

The following fact summarises a few immediate consequences of the above definitions:

FACT 3.0. (a) *If φ is a formula and $\nu \notin \mathrm{free}(\varphi)$, then:*

$$\mathbf{I}^a_\nu \vDash \varphi \quad \textit{if and only if} \quad \mathbf{I} \vDash \varphi$$

(b) *If $\varphi(\nu)$ is a formula and the substitution $\varphi(\nu/\tau)$ is admissible, then:*

$$\mathbf{I}^{\mathbf{I}(\tau)}_\nu \vDash \varphi(\nu) \quad \textit{if and only if} \quad \mathbf{I} \vDash \varphi(\tau)$$

Basic Notions of Model Theory

Let $\mathbf{\Phi}$ be an arbitrary set of $\mathscr{L}$-formulae. Then an $\mathscr{L}$-structure $\mathbf{M}$ is a **model of $\mathbf{\Phi}$** if for every assignment j and for each $\mathscr{L}$-formula $\varphi \in \mathbf{\Phi}$ we have $(\mathbf{M}, j) \vDash \varphi$, i.e., φ is true in the $\mathscr{L}$-interpretation $\mathbf{I} = (\mathbf{M}, j)$. Instead of saying "$\mathbf{M}$ is a model of $\mathbf{\Phi}$" we just write $\mathbf{M} \vDash \mathbf{\Phi}$. If φ fails in $\mathbf{M}$, then we write $\mathbf{M} \nvDash \varphi$. Notice that in the case when φ is a sentence, $\mathbf{M} \nvDash \varphi$ is equivalent to $\mathbf{M} \vDash \neg\varphi$, since for any $\mathscr{L}$-sentence φ we have *either* $\mathbf{M} \vDash \varphi$ *or* $\mathbf{M} \vDash \neg\varphi$.

EXAMPLE 3.1. Let $\mathscr{L} = \{\mathsf{c}, f\}$, where c is a constant symbol and f is a unitary function symbol. Furthermore, let $\mathbf{\Phi}$ consist of the following two $\mathscr{L}$-sentences

$$\underbrace{\forall x\big(x = \mathsf{c} \vee x = f(\mathsf{c})\big)}_{\varphi_1} \quad \text{and} \quad \underbrace{\exists x(x \neq \mathsf{c})}_{\varphi_2}.$$

We construct two models $\mathbf{M}_1$ and $\mathbf{M}_2$ with the same domain A, such that $\mathbf{M}_1 \vDash \mathbf{\Phi}$ and $\mathbf{M}_2 \nvDash \mathbf{\Phi}$: For this, let $A := \{0, 1\}$, and let

$$\mathsf{c}^{\mathbf{M}_1} := 0, \quad f^{\mathbf{M}_1}(0) := 1, \quad f^{\mathbf{M}_1}(1) := 0,$$

$$\mathsf{c}^{\mathbf{M}_2} := 0, \quad f^{\mathbf{M}_2}(0) := 0, \quad f^{\mathbf{M}_2}(1) := 1.$$

We leave it as an exercise to the reader to show that φ_2 holds in both models, whereas φ_1 holds just in the model $\mathbf{M}_1$. In fact, we have $\mathbf{M}_1 \vDash \varphi_1 \wedge \varphi_2$ and $\mathbf{M}_2 \vDash \neg\varphi_1 \wedge \varphi_2$.

As an immediate consequence of the definition of models, we get:

Fact 3.2. *If φ is an $\mathscr{L}$-formula, ν a variable, and $\mathbf{M}$ a model, then $\mathbf{M} \vDash \varphi$ if and only if $\mathbf{M} \vDash \forall\nu\varphi$.*

This leads to the following definition: Let $\langle \nu_1, \ldots, \nu_n \rangle$ be the sequence of variables which appear free in the $\mathscr{L}$-formula φ, where the variables appear in the sequence in the same order as they appear for the first time in φ if one reads φ from left to right. Then the **universal closure** of φ, denoted $\overline{\varphi}$, is defined by stipulating

$$\overline{\varphi} :\equiv \forall\nu_1 \cdots \forall\nu_n \, \varphi \,.$$

As a generalisation of Fact 3.2, we get:

Fact 3.3. *If φ is an $\mathscr{L}$-formula and $\mathbf{M}$ a model, then*

$$\mathbf{M} \vDash \varphi \quad \Longleftrightarrow \quad \mathbf{M} \vDash \overline{\varphi}$$

The following notation will be used later on to simplify the arguments when we investigate the truth-value of sentences in some model $\mathbf{M}$: Suppose that $\mathbf{M}$ is a model with domain A. Let $\varphi(\nu_1, \ldots, \nu_n)$ be an $\mathscr{L}$-formula whose free variables are $\nu_1, \ldots, \nu_n$ and let $a_1, \ldots, a_n \in A$. Then we write

$$\mathbf{M} \vDash \varphi(a_1, \ldots, a_n)$$

to denote that for every assignment j in $\mathbf{M}$ we have:

$$(\mathbf{M}, j\,\tfrac{a_1}{\nu_1} \cdots \tfrac{a_n}{\nu_n}) \vDash \varphi(\nu_1, \ldots, \nu_n)$$

Let us now have a closer look at models: For this, we fix a signature $\mathscr{L}$ (i.e., we fix a possibly empty set of constant symbols c, n-ary function symbols F, and n-ary relation symbols R). Two $\mathscr{L}$-structures $\mathbf{M}$ and $\mathbf{N}$ with domains A and B are **isomorphic**, denoted $\mathbf{M} \cong \mathbf{N}$, if there is a bijection $f : A \to B$ such that for all constant symbols $c \in \mathscr{L}$ we have

$$f\big(c^{\mathbf{M}}\big) = c^{\mathbf{N}} \,,$$

and for all natural numbers n, all n-ary function symbols $F \in \mathscr{L}$, all n-ary relation symbols $R \in \mathscr{L}$, and any $a_1, \ldots, a_n \in A$ we have:

$$f\big(F^{\mathbf{M}}(a_1, \ldots, a_n)\big) \;=\; F^{\mathbf{N}}\big(f(a_1), \ldots, f(a_n)\big)$$

$$\langle a_1, \ldots, a_n \rangle \in R^{\mathbf{M}} \;\Leftrightarrow\; \big\langle f(a_1), \ldots, f(a_n) \big\rangle \in R^{\mathbf{N}}$$

Since models are just $\mathscr{L}$-structures, we can extend the notion of an isomorphism to models and obtain the following:

FACT 3.4. *If* $\mathbf{M}$ *and* $\mathbf{N}$ *are isomorphic models of some given set of* $\mathscr{L}$*-formulae and* φ *is an* $\mathscr{L}$*-formula, then:*

$$\mathbf{M} \models \varphi \quad \Lleftarrow\!\Rrightarrow \quad \mathbf{N} \models \varphi$$

It may happen that although two $\mathscr{L}$-structures $\mathbf{M}$ and $\mathbf{N}$ are not isomorphic there is no $\mathscr{L}$-sentence that can distinguish between them. In this case, we say that $\mathbf{M}$ and $\mathbf{N}$ are elementarily equivalent. More formally, we say that $\mathbf{M}$ is **elementarily equivalent** to $\mathbf{N}$, denoted by $\mathbf{M} \equiv_e \mathbf{N}$, if each $\mathscr{L}$-sentence σ which is true in $\mathbf{M}$ is also true in $\mathbf{N}$. The following lemma shows that $\equiv_e$ is symmetric:

LEMMA 3.5. *If* $\mathbf{M}$ *and* $\mathbf{N}$ *are* $\mathscr{L}$*-structures and* $\mathbf{M} \equiv_e \mathbf{N}$*, then for each* $\mathscr{L}$*-sentence* σ *we have:*

$$\mathbf{M} \models \sigma \quad \Lleftarrow\!\Rrightarrow \quad \mathbf{N} \models \sigma$$

Proof. One direction follows immediately from the definition. For the other direction, assume that σ is not true in $\mathbf{M}$, i.e., $\mathbf{M} \not\models \sigma$. Then $\mathbf{M} \models \neg\sigma$, which implies $\mathbf{N} \models \neg\sigma$, and hence, σ is not true in $\mathbf{N}$. ⊣

As a consequence of FACT 3.3, we get:

FACT 3.6. *If* $\mathbf{M}$ *and* $\mathbf{N}$ *are elementarily equivalent models of some given set of* $\mathscr{L}$*-formulae and* φ *is an* $\mathscr{L}$*-formula, then we have:*

$$\mathbf{M} \models \varphi \quad \Lleftarrow\!\Rrightarrow \quad \mathbf{N} \models \varphi$$

In what follows, we investigate the relationship between syntax and semantic. In particular, we investigate the relationship between a formal proof of a formula φ from a set of formulae $\mathbf{\Phi}$ and the truth-value of φ in a model of $\mathbf{\Phi}$. In this context, two questions arise naturally:

- Is each formula φ which is provable from some set of formulae $\mathbf{\Phi}$ valid in every model $\mathbf{M}$ of $\mathbf{\Phi}$?
- Is every formula φ which is valid in each model $\mathbf{M}$ of $\mathbf{\Phi}$ provable from $\mathbf{\Phi}$?

Soundness Theorem

In this section, we give an answer to the former question in the previous paragraph; the answer to the latter question is postponed to Part II.

A logical calculus is called **sound** if all that we can prove is valid (i.e., true), which implies that we cannot derive a contradiction. The following theorem shows that First-Order Logic is sound.

THEOREM 3.7 (SOUNDNESS THEOREM). *Let* $\mathbf{\Phi}$ *be a set of* $\mathscr{L}$*-formulae and* $\mathbf{M}$ *a model of* $\mathbf{\Phi}$*. Then for every* $\mathscr{L}$*-formula* φ *we have:*

$$\mathbf{\Phi} \vdash \varphi \quad \Longrightarrow \quad \mathbf{M} \vDash \varphi$$

Somewhat shorter, we could say

$$\forall\, \varphi : \ \mathbf{\Phi} \vdash \varphi \quad \Longrightarrow \quad \forall\, \mathbf{M} \big(\mathbf{M} \vDash \mathbf{\Phi} \ \Longrightarrow \mathbf{M} \vDash \varphi \big) \,,$$

where the symbol $\forall$ *stands for* "FOR ALL".

Proof. First we show that all logical axioms are valid in $\mathbf{M}$. For this, we have to define truth-values of composite statements in the metalanguage. In the previous chapter, e.g., we defined:

$$\underbrace{\mathbf{M} \vDash \varphi \wedge \psi}_{\Theta} \quad \Longleftrightarrow \quad \underbrace{\mathbf{M} \vDash \varphi}_{\Phi} \ \text{AND} \ \underbrace{\mathbf{M} \vDash \psi}_{\Psi}$$

Thus, in the metalanguage the statement Θ is true if and only if the statement "Φ AND Ψ" is true. So, the truth-value of Θ depends on the truth-values of Φ and Ψ. In order to determine truth-values of composite statement like "Φ AND Ψ" or "IF Φ THEN Ψ", where the latter statement will get the same truth-value as "NOT Φ OR Ψ", we introduce so called *truth-tables*, in which **1** stands for **true** and **0** stands for **false**:

Φ	Ψ	NOT Φ	Φ AND Ψ	Φ OR Ψ	IF Φ THEN Ψ
0	0	1	0	0	1
0	1	1	0	1	1
1	0	0	0	1	0
1	1	0	1	1	1

With these truth-tables, one can show that all logical axioms are valid in $\mathbf{M}$. As an example, we show that every instance of L_1 is valid in $\mathbf{M}$. For this,

let φ_1 be an instance of L_1, i.e., $\varphi_1 \equiv \varphi \to (\psi \to \varphi)$ for some $\mathscr{L}$-formulae φ and ψ. Then $\mathbf{M} \vDash \varphi_1$ if and only if $\mathbf{M} \vDash \varphi \to (\psi \to \varphi)$:

$$\underbrace{\mathbf{M} \vDash \varphi \to (\psi \to \varphi)}_{\Theta} \quad \Longleftrightarrow \quad \text{IF } \underbrace{\mathbf{M} \vDash \varphi}_{\Phi} \text{ THEN } \underbrace{\mathbf{M} \vDash \psi \to \varphi}$$

$$\Theta \quad \Longleftrightarrow \quad \text{IF } \Phi \text{ THEN } (\text{IF } \underbrace{\mathbf{M} \vDash \psi}_{\Psi} \text{ THEN } \underbrace{\mathbf{M} \vDash \varphi}_{\Phi})$$

This shows that

$$\Theta \quad \Longleftrightarrow \quad \text{IF } \Phi \text{ THEN } (\text{IF } \Psi \text{ THEN } \Phi)\,.$$

Writing the truth-table of Θ, we see that the statement Θ is always true (i.e., φ_1 is valid in $\mathbf{M}$):

Φ	Ψ	IF Ψ THEN Φ	IF Φ THEN (IF Ψ THEN Φ)
0	0	1	1
0	1	0	1
1	0	1	1
1	1	1	1

Therefore, $\mathbf{M} \vDash \varphi_1$, and since φ_1 was an arbitrary instance of L_1, every instance of L_1 is valid in $\mathbf{M}$.

In order to show that the logical axioms L_{10}–L_{16} are also valid in $\mathbf{M}$, we need somewhat more than just truth-tables. For this purpose, let A be the domain of $\mathbf{M}$, let j be an arbitrary assignment, and let $\mathbf{I} = (\mathbf{M}, j)$ be the corresponding $\mathscr{L}$-interpretation. Now we show that every instance of L_{10} is valid in $\mathbf{M}$. For this, let φ_{10} be an instance of L_{10}, i.e., $\varphi_{10} \equiv \forall\nu\varphi(\nu) \to \varphi(\tau)$ for some $\mathscr{L}$-formula φ, where ν is a variable, τ an $\mathscr{L}$-term, and the substitution $\varphi(\nu/\tau)$ is admissible. We work with $\mathbf{I}$ and show that $\mathbf{I} \vDash \varphi_{10}$.

By definition, we have

$$\mathbf{I} \vDash \forall\nu\varphi(\nu) \to \varphi(\tau) \quad \Longleftrightarrow \quad \text{IF } \mathbf{I} \vDash \forall\nu\varphi(\nu) \text{ THEN } \mathbf{I} \vDash \varphi(\tau)\,,$$

and again by definition, we have

$$\mathbf{I} \vDash \forall\nu\varphi(\nu) \quad \Longleftrightarrow \quad \text{FOR ALL } a \text{ IN } A\colon \ \mathbf{I}^a_\nu \vDash \varphi(\nu)\,.$$

In particular, we obtain

$$\mathbf{I} \vDash \forall\nu\varphi(\nu) \quad \Longrightarrow \quad \mathbf{I}^{I(\tau)}_\nu \vDash \varphi(\nu)\,.$$

Furthermore, by FACT 3.0.(b) we get

$$\mathbf{I} \vDash \varphi(\tau) \iff \mathbf{I}^{\frac{I(\tau)}{\nu}} \vDash \varphi(\nu) .$$

Hence, we have

$$\text{IF } \mathbf{I} \vDash \forall\nu\varphi(\nu) \text{ THEN } \mathbf{I} \vDash \varphi(\tau)$$

which shows that

$$(\mathbf{M}, j) \vDash \forall\nu\varphi(\nu) \to \varphi(\tau) .$$

Since the assignment j was arbitrary, we finally have:

$$\mathbf{M} \vDash \forall\nu\varphi(\nu) \to \varphi(\tau)$$

Therefore, $\mathbf{M} \vDash \varphi_{10}$, and since φ_{10} was an arbitrary instance of L_{10}, every instance of L_{10} is valid in $\mathbf{M}$.

With similar arguments, one can show that every instance of L_{11}, L_{12}, or L_{13} is also valid in $\mathbf{M}$ (see EXERCISE 3.5.(a)). Furthermore, one can show that L_{14}, L_{15}, and L_{16} are also valid in $\mathbf{M}$ (see EXERCISE 3.5.(b)).

Let $\mathbf{\Phi}$ be a set of formulae, let $\mathbf{M}$ be a model of $\mathbf{\Phi}$, and assume that $\mathbf{\Phi} \vdash \varphi_0$ for some $\mathscr{L}$-formula φ_0. We shall show that $\mathbf{M} \vDash \varphi_0$. For this, we first notice the following facts:

- As we have seen above, each instance of a logical axiom is valid in $\mathbf{M}$.
- Since $\mathbf{M} \vDash \mathbf{\Phi}$, each formula of $\mathbf{\Phi}$ is valid in $\mathbf{M}$.
- By the truth-tables, we get

  $$\text{IF } (\mathbf{M} \vDash \varphi \to \psi \text{ AND } \mathbf{M} \vDash \varphi) \text{ THEN } \mathbf{M} \vDash \psi$$

  and therefore, every application of Modus Ponens in the proof of φ_0 from $\mathbf{\Phi}$ yields a valid formula (if the premises are valid).
- Since, by FACT 3.2,

  $$\mathbf{M} \vDash \varphi \iff \mathbf{M} \vDash \forall\nu\varphi(\nu)$$

  every application of the Generalisation in the proof of φ_0 from $\mathbf{\Phi}$ yields a valid formula.

From these facts, it follows immediately that *each* formula in the proof of φ_0 from $\mathbf{\Phi}$ is valid in $\mathbf{M}$. In particular, we get

$$\mathbf{M} \vDash \varphi_0$$

which completes the proof. ⊣

The following fact summarises a few consequences of the SOUNDNESS THEOREM 3.7.

Fact 3.8.

(a) *Every tautology is valid in each model:*

$$\forall\, \varphi : \ \vdash \varphi \quad \Longrightarrow \quad \forall\, \mathbf{M} : \ \mathbf{M} \vDash \varphi$$

(b) *If a set of formulae* $\mathbf{\Phi}$ *has a model, then* $\mathbf{\Phi}$ *is consistent:*

$$\exists!\, \mathbf{M} : \ \mathbf{M} \vDash \mathbf{\Phi} \quad \Longrightarrow \quad \mathrm{Con}(\mathbf{\Phi})$$

Here, the symbol $\exists!$ *stands for "*it exists*".*

(c) *The logical axioms are consistent:*

$$\mathrm{Con}(\mathsf{L}_0\text{-}\mathsf{L}_{16})$$

(d) *If a sentence* σ *is not valid in* $\mathbf{M}$*, where* $\mathbf{M}$ *is a model of* $\mathbf{\Phi}$*, then* σ *is not provable from* $\mathbf{\Phi}$*:*

$$\text{if} \quad (\, \mathbf{M} \nvDash \sigma \quad \text{and} \quad \mathbf{M} \vDash \mathbf{\Phi} \,) \quad \text{then} \quad \mathbf{\Phi} \nvdash \sigma$$

Completion of Theories

A set of $\mathscr{L}$-sentences is called an $\mathscr{L}$**-theory**, denoted by T. An $\mathscr{L}$-theory T is called **complete**, if for every $\mathscr{L}$-sentence σ we have *either* $\mathsf{T} \vdash \sigma$ *or* $\mathsf{T} \vdash \neg\sigma$. Notice that, by definition, an inconsistent theory is always **incomplete** (i.e., not complete). Furthermore, for an $\mathscr{L}$-theory T let $\mathbf{Th}(\mathsf{T})$ be the set of all $\mathscr{L}$-sentences σ, such that $\mathsf{T} \vdash \sigma$. By these definitions, we get that a consistent $\mathscr{L}$-theory T is complete if and only if for every $\mathscr{L}$-sentence σ we have *either* $\sigma \in \mathbf{Th}(\mathsf{T})$ *or* $\neg\sigma \in \mathbf{Th}(\mathsf{T})$.

Proposition 3.9. *If* T *is an* $\mathscr{L}$*-theory which has a model, then there exists a complete* $\mathscr{L}$*-theory* $\bar{\mathsf{T}}$ *which contains* T*. In particular, every* $\mathscr{L}$*-theory which has a model can be completed (i.e., can be extended to a complete theory).*

Proof. Let $\mathbf{M}$ be a model of some $\mathscr{L}$-theory T and let $\bar{\mathsf{T}}$ be the set of $\mathscr{L}$-sentences σ such that $\mathbf{M} \vDash \sigma$. Since for each $\mathscr{L}$-sentence σ_0 we have either $\mathbf{M} \vDash \sigma_0$ or $\mathbf{M} \vDash \neg\sigma_0$, we get either $\sigma_0 \in \bar{\mathsf{T}}$ or $\neg\sigma_0 \in \bar{\mathsf{T}}$, which shows that $\bar{\mathsf{T}}$ is complete, and since $\mathbf{M} \vDash \mathsf{T}$, we get that $\bar{\mathsf{T}}$ contains T. ⊣

Let $\mathbf{M}$ be a model of some $\mathscr{L}$-theory T and let $\bar{\mathsf{T}}$ be the set of $\mathscr{L}$-sentences σ such that $\mathbf{M} \vDash \sigma$. Then $\bar{\mathsf{T}}$ is called the **theory of** $\mathbf{M}$, denoted by $\mathbf{Th}(\mathbf{M})$. By definition, the theory $\mathbf{Th}(\mathbf{M})$ is always complete.

It is natural to ask whether the converse of Proposition 3.9 also holds, i.e., whether every $\mathscr{L}$-theory which can be completed has a model. Notice that if an $\mathscr{L}$-theory T can be completed, then T must be consistent. So, one

may ask whether every consistent theory has a model. An affirmative answer to this question together with FACT 3.8.(b) would imply that an $\mathscr{L}$-theory T is consistent if and only if T has a model — which is indeed the case, as we shall see below.

NOTES

The history of Model Theory can be traced back to the 19th century, when semantics began to play a role in Logic. However, one of the earliest results in modern Model Theory is Gödel's COMPLETENESS THEOREM (see Chapter 5). In the 1950's and 1960's, Model Theory was further developed, e.g., by Jerry Łoś (see Chapter 15) and Abraham Robinson (see Chapter 17).

EXERCISES

3.0 Show that the domain of a model is never empty.

Hint: Use EXAMPLE 1.0.

3.1 Let R be a binary relation symbol and let the three sentences $\varphi_1, \varphi_2, \varphi_3$ be defined as follows:

$$\varphi_1 :\equiv \forall x(xRx)\,, \quad \varphi_2 :\equiv \forall x\forall y(xRy \to yRx)\,, \quad \varphi_3 :\equiv \forall x\forall y\forall z\big((xRy \wedge yRz) \to xRz\big)$$

Find three models $\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3$ with domains as small as possible, such that

$$\mathbf{M}_1 \vDash \neg\varphi_1 \wedge \varphi_2 \wedge \varphi_3\,, \qquad \mathbf{M}_2 \vDash \varphi_1 \wedge \neg\varphi_2 \wedge \varphi_3\,, \qquad \mathbf{M}_3 \vDash \varphi_1 \wedge \varphi_2 \wedge \neg\varphi_3\,.$$

3.2 Let T be a set of $\mathscr{L}'$-sentences (for some signature $\mathscr{L}'$) and let $\mathbf{M}'$ be an $\mathscr{L}'$-structure such that $\mathbf{M}' \vDash \mathsf{T}$. Furthermore, let $\mathscr{L}$ be an extension of $\mathscr{L}'$ (i.e., $\mathscr{L}$ is a signature which contains $\mathscr{L}'$). Then there is an $\mathscr{L}$-structure $\mathbf{M}$ with the same domain as $\mathbf{M}'$, such that $\mathbf{M} \vDash \mathsf{T}$.

3.3 If two structures $\mathbf{M}$ and $\mathbf{N}$ are isomorphic, then they are elementarily equivalent.

3.4 Let DLO be the theory of dense linearly ordered sets without endpoints. More precisely, the signature $\mathscr{L}_{\mathsf{DLO}}$ contains just the binary relation symbol $<$, and the non-logical axioms of DLO are the following sentences:

DLO_0 $\forall x\neg(x < x)$
DLO_1 $\forall x\forall y\forall z\big((x < y \wedge y < z) \to x < z\big)$
DLO_2 $\forall x\forall y\big(x < y \vee x = y \vee y < x\big)$
DLO_3 $\forall x\forall y\exists z\big(x < y \to (x < z \wedge z < y)\big)$
DLO_4 $\forall x\exists y\exists z\big(y < x \wedge x < z\big)$

Show that every countable model of DLO is isomorphic to $(\mathbb{Q}, <)$.

Hint: Enumerate both $\mathbb{Q}$ and some model $\mathbf{M}$ of DLO, and construct an isomorphism by recursion in such a way that in the n-th step the n-th element of M is mapped to an element of $\mathbb{Q}$ with the order being preserved.

3.5 Let $\mathscr{L}$ be an arbitrary signature and let $\mathbf{M}$ be an arbitrary $\mathscr{L}$-structure.

(a) Show that L_{11}-L_{13} are valid in $\mathbf{M}$.

(b) Show that L_{14}-L_{16} are valid in $\mathbf{M}$.

3.6 We say that two $\mathscr{L}$-formulae φ and ψ are **semantically equivalent** if for all $\mathscr{L}$-structures $\mathbf{M}$ and every assignment j we have

$$(\mathbf{M}, j) \vDash \varphi \quad \Lsh\!\!\Rsh \quad (\mathbf{M}, j) \vDash \psi$$

(a) Show that for every sentence σ there is a semantically equivalent sentence $\tilde{\sigma}$ which contains just variables among $v_0, v_1, \ldots$, where for any $m, n \in \mathbb{N}$ with $m < n$, if v_n appears in $\tilde{\sigma}$, then also v_m appears in $\tilde{\sigma}$ (compare with THEOREM 2.13).

(b) Show that for every $\mathscr{L}$-sentence σ there is a semantically equivalent $\mathscr{L}$-sentence in sPNF (compare with THEOREM 2.14).

3.7 (a) Show that Group Theory GT is incomplete.

(b) Let $\psi \equiv \forall x \forall y\, (x \circ y = y \circ x)$. Show that $\mathsf{GT} + \psi$ is incomplete.

Part II
Gödel's Completeness Theorem

In this part of the book, we shall prove Gödel's Completeness Theorem and show several consequences of it. Roughly speaking, Gödel's Completeness Theorem states that every consistent $\mathscr{L}$-theory T has a model $\mathbf{M} \models \mathsf{T}$. With respect to the model $\mathbf{M}$, every $\mathscr{L}$-sentence σ is either true or false (i.e., either σ or $\neg\sigma$ is true in $\mathbf{M}$). Hence, the set of $\mathscr{L}$-sentences which are true in $\mathbf{M}$ is with this respect *complete*. Therefore, as a consequence of Gödel's Completeness Theorem we obtain that every consistent theory is contained in a complete theory. However, this result should not be confused with Gödel's First Incompleteness Theorem (presented in Part III), which states that the theory of Peano Arithmetic PA is incomplete.

Gödel proved his famous theorem in his doctoral dissertation *Über die Vollständigkeit des Logikkalküls* [14] which he completed in 1929. In 1930, he published the same material as in the doctoral dissertation in a rewritten and shortened form in [15]. However, instead of presenting Gödel's original proof we decided to follow Henkin's construction, which can be found in [23] (see also [25]), since it fits better in the logical framework as developed in Part I. Even though Henkin's construction also works for uncountable signatures, we shall prove in Chapter 15 the general Completeness Theorem with an ultraproduct construction, using Łoś's Theorem.

We would like to mention that in our proof of Gödel's Completeness Theorem — in contrast to Henkin's proof — we only have to assume the existence of *potentially infinite* sets, but no instance of an *actually infinite* set is required (see also Chapter 0).

Chapter 4
Maximally Consistent Extensions

Throughout this chapter, we require that all formulae are written in Polish notation and that the variables are among $v_0, v_1, v_2, \ldots$ Notice that the former requirement is just another notation which does not involve brackets, and that by the VARIABLE SUBSTITUTION THEOREM 2.13, the latter requirement gives us semantically equivalent formulae.

Maximally Consistent Theories

Let $\mathscr{L}$ be an arbitrary signature and let T be an $\mathscr{L}$-theory (i.e., a set of $\mathscr{L}$-sentences). We say that T is **maximally consistent** if T is consistent and for every $\mathscr{L}$-sentence σ we have *either* $\sigma \in \mathsf{T}$ *or* $\neg\,\mathrm{Con}(\mathsf{T} + \sigma)$. In other words, a consistent theory T is maximally consistent if no proper extension of T is consistent. The following fact is just a reformulation of this definition.

FACT 4.0. *Let $\mathscr{L}$ be a signature and let T be a consistent $\mathscr{L}$-theory. Then T is maximally consistent if and only if for every $\mathscr{L}$-sentence σ, either $\sigma \in \mathsf{T}$ or $\mathsf{T} \vdash \neg\sigma$.*

Proof. By FACT 2.16(c) & (d) we have:

$$\neg\,\mathrm{Con}(\mathsf{T} + \sigma) \quad \Longleftrightarrow \quad \mathsf{T} \vdash \neg\sigma$$

Hence, an $\mathscr{L}$-theory is maximally consistent if and only if for every $\mathscr{L}$-sentence σ, either $\sigma \in \mathsf{T}$ or $\mathsf{T} \vdash \neg\sigma$. ⊣

As a consequence of FACT 4.0, we get

LEMMA 4.1. *Let $\mathscr{L}$ be a signature and let T be a consistent $\mathscr{L}$-theory. Then T is maximally consistent if and only if for every $\mathscr{L}$-sentence σ, either $\sigma \in \mathsf{T}$ or $\neg\sigma \in \mathsf{T}$.*

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_5

Proof. We have to show that the following equivalence holds:

$$\forall\sigma\big(\sigma \in \mathsf{T} \;\; or \;\; \mathsf{T} \vdash \neg\sigma\big) \quad \Longleftrightarrow \quad \forall\sigma\big(\sigma \in \mathsf{T} \;\; or \;\; \neg\sigma \in \mathsf{T}\big)$$

($\Rightarrow$) Assume that for every $\mathscr{L}$-sentence σ we have $\sigma \in \mathsf{T}$ or $\mathsf{T} \vdash \neg\sigma$. If $\sigma \in \mathsf{T}$, then the implication obviously holds. If $\sigma \notin \mathsf{T}$, then $\mathsf{T} \vdash \neg\sigma$, and since T is consistent, this implies $\mathsf{T} \nvdash \sigma$. Now, by TAUTOLOGY (F), this implies $\mathsf{T} \nvdash \neg\neg\sigma$ and by our assumption we finally get $\neg\sigma \in \mathsf{T}$.

($\Leftarrow$) Assume that for every $\mathscr{L}$-sentence σ we have $\sigma \in \mathsf{T}$ or $\neg\sigma \in \mathsf{T}$. If $\sigma \in \mathsf{T}$, then the implication obviously holds. Now, if $\sigma \notin \mathsf{T}$, then by our assumption we have $\neg\sigma \in \mathsf{T}$, which obviously implies $\mathsf{T} \vdash \neg\sigma$. ⊣

Maximally consistent theories have similar features as complete theories: Recall that an $\mathscr{L}$-theory T is complete if for every $\mathscr{L}$-sentence σ we have *either* $\mathsf{T} \vdash \sigma$ *or* $\mathsf{T} \vdash \neg\sigma$.

As an immediate consequence of the definitions, we get

FACT 4.2. *Let $\mathscr{L}$ be a signature, let T be a consistent $\mathscr{L}$-theory, and let* $\mathbf{Th}(\mathsf{T})$ *be the set of all $\mathscr{L}$-sentences which are provable from T.*

(a) *If T is complete, then* $\mathbf{Th}(\mathsf{T})$ *is maximally consistent.*

(b) *If T is maximally consistent, then* $\mathbf{Th}(\mathsf{T})$ *is the same as T.*

The next result gives a condition under which a theory can be extended to a maximally consistent theory. In fact, it is just a reformulation of PROPOSITION 3.9.

FACT 4.3. *If an $\mathscr{L}$-theory T has a model, then T has a maximally consistent extension.*

Proof. Let $\mathbf{M}$ be a model of the $\mathscr{L}$-theory T and let $\mathbf{Th}(\mathbf{M})$ be the set of $\mathscr{L}$-sentences σ such that $\mathbf{M} \vDash \sigma$. Then $\mathbf{Th}(\mathbf{M})$ is obviously a maximally consistent theory which contains T. ⊣

Later we shall see that every consistent theory has a model. For this, we first show how a consistent theory can be extended to a maximally consistent theory.

Universal List of Sentences

Let $\mathscr{L}$ be an arbitrary but fixed countable signature, where by "countable" we mean that the symbols in $\mathscr{L}$ can be listed in a FINITE or POTENTIALLY INFINITE list $L_{\mathscr{L}}$.

First, we encode the symbols of $\mathscr{L}$ corresponding to the order in which they appear in the list $L_{\mathscr{L}}$: The first symbol is encoded with "2", the second with "22", the third with "222", and so on. For every symbol $\zeta \in L_{\mathscr{L}}$, let $\#\zeta$ denote the code of ζ. Therefore, the code of a symbol of $\mathscr{L}$ is just a sequence of 2's.

Furthermore, we encode the logical symbols as follows:

Symbol ζ	Code $\#\zeta$
$=$	11
$\neg$	1111
$\wedge$	111111
$\vee$	11111111
$\rightarrow$	1111111111
$\exists$	111111111111
$\forall$	11111111111111
v_0	1
v_1	111
$\vdots$	$\vdots$
v_n	$\underbrace{1111 \ \ldots \ 11111}_{(2n+1)\text{ 1's}}$

In the next step, we encode strings of symbols: Let $\bar{\zeta} \equiv \zeta_0\zeta_1\zeta_2\ldots\zeta_n$ be a finite string of symbols, then

$$\#\bar{\zeta} := \#\zeta_0 0 \#\zeta_1 0 \#\zeta_2 \ldots 0 \#\zeta_n$$

For a string $\#\zeta$ (i.e., a string of 0's, 1's, and 2's), let $|\#\zeta|$ be the length of $\#\zeta$ (i.e., the number of 0's, 1's, and 2's which appear in $\#\zeta$).

Now, we order the codes of strings of symbols by their length and strings of the same length lexicographically, where $0 < 1 < 2$. If, with respect to this ordering, $\#\zeta$ is less than $\#\zeta'$, then we write $\#\zeta \prec \#\zeta'$.

Finally, let

$$\Lambda_{\mathscr{L}} := [\sigma_1, \sigma_2, \ldots]$$

be the potentially infinite list of all $\mathscr{L}$-sentences written in Polish notation (notice that we did not encode brackets), where we require

$$\#\sigma_i \prec \#\sigma_j \iff i < j\,.$$

We call $\Lambda_{\mathscr{L}}$ the **universal list of $\mathscr{L}$-sentences**.

Lindenbaum's Lemma

In this section, we show that every consistent set of $\mathscr{L}$-sentences T can be extended to a maximally consistent set of $\mathscr{L}$-sentences $\overline{\mathsf{T}}$. Since the universal list of $\mathscr{L}$-sentences contains all possible $\mathscr{L}$-sentences, every set T of $\mathscr{L}$-sentences can be listed in a finite or potentially infinite list.

THEOREM 4.4 (LINDENBAUM'S LEMMA). *Let $\mathscr{L}$ be a countable signature and let T be a consistent set of $\mathscr{L}$-sentences. Furthermore, let σ_0 be an $\mathscr{L}$-sentence which cannot be proved from T, i.e., $\mathsf{T} \nvdash \sigma_0$. Then there exists a maximally consistent set $\overline{\mathsf{T}}$ of $\mathscr{L}$-sentences which contains $\neg\sigma_0$ as well as all the sentences of T.*

Proof. Let $\Lambda_{\mathscr{L}} = [\sigma_1, \sigma_2, \ldots]$ be the universal list of all $\mathscr{L}$-sentences. First we extend $\Lambda_{\mathscr{L}}$ with the $\mathscr{L}$-sentence $\neg\sigma_0$; let $\Lambda^0_{\mathscr{L}} = [\neg\sigma_0, \sigma_1, \sigma_2, \ldots]$.

Now, we go through the list $\Lambda^0_{\mathscr{L}}$ and define step by step a list $\overline{\mathsf{T}}$ of $\mathscr{L}$-sentences. For this, we define T_0 as the list which contains just $\neg\sigma_0$, i.e., $\mathsf{T}_0 := [\neg\sigma_0]$. If T_n is already defined, then

$$\mathsf{T}_{n+1} := \begin{cases} \mathsf{T}_n + [\sigma_{n+1}] & \text{if } \operatorname{Con}(\mathsf{T} + \mathsf{T}_n + \sigma_{n+1}), \\ \mathsf{T}_n & \text{otherwise.} \end{cases}$$

Let $\overline{\mathsf{T}} = [\neg\sigma_0, \sigma_{i_1}, \ldots]$ be the resulting list, i.e., $\overline{\mathsf{T}}$ is the potentially infinite list which contains each finite list T_n as an initial segment. Notice that this construction only works if we assume the metamathematical LAW OF EXCLUDED MIDDLE or a similar principle like the WEAK KÖNIG'S LEMMA (see EXERCISE 4.1): Even in the case when we cannot decide whether $\mathsf{T} + T_n + \sigma_n$ is consistent or not, we assume, from a metamathematical point of view, that *either* $\mathsf{T} + T_n + \sigma_n$ is consistent *or* $\mathsf{T} + T_n + \sigma_n$ is inconsistent (and *neither* both, *nor* none).

The following claim states that we cannot derive a contradiction from finitely many $\mathscr{L}$-sentences in $\overline{\mathsf{T}}$ and that we cannot add any new $\mathscr{L}$-sentence to $\overline{\mathsf{T}}$ without destroying this property. However, in order to simplify our terminology, we shall consider the potentially infinite list $\overline{\mathsf{T}}$ as an actual infinite set — notice that this is just a "façon-de-parler" since we do not have to assume the existence of actual infinite sets.

CLAIM. *$\overline{\mathsf{T}}$ is a maximally consistent set of $\mathscr{L}$-sentences which contains $\neg\sigma_0$ as well as all the sentences of T.*

Proof of Claim. First we show that $\overline{\mathsf{T}}$ contains $\mathsf{T} + \neg\sigma_0$, then we show that $\overline{\mathsf{T}}$ is consistent, and finally we show that for every $\mathscr{L}$-sentence σ we have either $\sigma \in \overline{\mathsf{T}}$ or $\neg \operatorname{Con}(\overline{\mathsf{T}} + \sigma)$.

$\overline{\mathsf{T}}$ contains all sentences of $\mathsf{T} + \neg\sigma_0$: By definition, $\mathsf{T}_0 = [\neg\sigma_0]$, and since T_0 is an initial segment of the list $\overline{\mathsf{T}}$, $\neg\sigma_0$ belongs to $\overline{\mathsf{T}}$. For every $\sigma \in \mathsf{T}$,

there is $n \in \mathbb{N}$ with $n \geq 1$ such that $\sigma \equiv \sigma_n$. By induction, we show that if $\sigma_n \in \mathsf{T}$ then $\sigma_n \in \overline{\mathsf{T}}$. For this, suppose that the claim holds for all $m \leq n$ and that $\sigma_{n+1} \in \mathsf{T}$. Let m be the largest number $m \leq n$ such that $\sigma_m \in \overline{\mathsf{T}}$. Then we have $\mathsf{T}_n = \mathsf{T}_m$. If $\neg\,\mathrm{Con}(\mathsf{T} + \mathsf{T}_n + \sigma_{n+1})$, then since $\sigma_{n+1} \in \mathsf{T}$, we have $\neg\,\mathrm{Con}(\mathsf{T} + \mathsf{T}_m)$, contradicting $\sigma_m \in \overline{\mathsf{T}}$. Hence we have $\mathrm{Con}(\mathsf{T} + \mathsf{T}_n + \sigma_{n+1})$ and therefore $\sigma_{n+1} \in \mathsf{T}_{n+1}$.

$\overline{\mathsf{T}}$ *is consistent*: By the Compactness Theorem 2.17 it is enough to show that every finite subset of $\overline{\mathsf{T}}$ is consistent. Since every finite subset of $\overline{\mathsf{T}}$ is contained in T_n for some n, it suffices to prove by induction that T_n is consistent for every $n \in \mathbb{N}$. Since $\mathsf{T} \nvdash \sigma_0$, $\mathsf{T}_0 = [\neg\sigma_0]$ is consistent. Now suppose $\mathrm{Con}(\mathsf{T}_m)$ for all $m \leq n$. If $\neg\,\mathrm{Con}(\mathsf{T} + \mathsf{T}_n + \sigma_{n+1})$, then $\mathsf{T}_{n+1} = \mathsf{T}_n$ is consistent by our induction hypothesis. Otherwise, $\mathrm{Con}(\mathsf{T} + \mathsf{T}_n + \sigma_{n+1})$ and therefore T_{n+1} is consistent, too.

For every σ, *either* $\sigma \in \overline{\mathsf{T}}$ *or* $\neg\,\mathrm{Con}(\overline{\mathsf{T}} + \sigma)$: For every $\mathscr{L}$-sentence σ, there is a $n \in \mathbb{N}$ with $n \geq 1$ such that $\sigma \equiv \sigma_n$. By the Law of Excluded Middle, we have *either* $\mathrm{Con}(\mathsf{T} + \mathsf{T}_{n-1} + \sigma_n)$ *or* $\neg\,\mathrm{Con}(\mathsf{T} + \mathsf{T}_{n-1} + \sigma_n)$. In the former case we obtain $\sigma_n \in \mathsf{T}_n$, which implies $\sigma \in \overline{\mathsf{T}}$. In the latter case we obtain $\neg\,\mathrm{Con}(\overline{\mathsf{T}} + \sigma_n)$, which is the same as $\neg\,\mathrm{Con}(\overline{\mathsf{T}} + \sigma)$. $\dashv_{\text{Claim}}$

Thus, the list $\overline{\mathsf{T}}$ has all the required properties, which completes the proof. $\dashv$

The following fact summarises the main properties of $\overline{\mathsf{T}}$.

Fact 4.5. *Let* T, $\overline{\mathsf{T}}$, *and* σ_0 *be as above, and let* σ *and* σ' *be any* $\mathscr{L}$*-sentences.*

(a) $\neg\sigma_0 \in \overline{\mathsf{T}}$.

(b) *Either* $\sigma \in \overline{\mathsf{T}}$ *or* $\neg\sigma \in \overline{\mathsf{T}}$.

(c) *If* $\mathsf{T} \vdash \sigma$, *then* $\sigma \in \overline{\mathsf{T}}$.

(d) $\overline{\mathsf{T}} \vdash \sigma$ *if and only if* $\sigma \in \overline{\mathsf{T}}$.

(e) *If* $\sigma \Leftrightarrow \sigma'$, *then* $\sigma \in \overline{\mathsf{T}}$ *if and only if* $\sigma' \in \overline{\mathsf{T}}$.

Proof. (a) follows by construction of $\overline{\mathsf{T}}$.

Since $\overline{\mathsf{T}}$ is maximally consistent, (b) follows by Lemma 4.1.

For (c), notice that $\mathsf{T} \vdash \sigma$ implies $\neg\,\mathrm{Con}(\mathsf{T} + \neg\sigma)$, hence $\neg\sigma \notin \overline{\mathsf{T}}$ and by (b) we get $\sigma \in \overline{\mathsf{T}}$.

For (d), let us first assume $\overline{\mathsf{T}} \vdash \sigma$, where $\sigma \equiv \sigma_n$. This implies $\mathrm{Con}(\overline{\mathsf{T}} + \sigma)$, hence $\mathrm{Con}(\mathsf{T} + T_n + \sigma_n)$, and by construction of $\overline{\mathsf{T}}$ we get $\sigma_n \in \overline{\mathsf{T}}$. On the other hand, if $\sigma \in \overline{\mathsf{T}}$, then we obviously have $\overline{\mathsf{T}} \vdash \sigma$.

For (e), recall that $\sigma \Leftrightarrow \sigma'$ is just an abbreviation for $\vdash \sigma \leftrightarrow \sigma'$. Thus, (e) follows immediately from (d). $\dashv$

Fact 4.5 shows that the $\mathscr{L}$-sentences in $\overline{\mathsf{T}}$ "behave" like valid sentences in a model, which is indeed the case—as the following proposition shows.

PROPOSITION 4.6. *Let* $\overline{\mathsf{T}}$ *be as above, and let* $\sigma, \sigma_1, \sigma_2$ *be any* $\mathscr{L}$*-sentences in Polish notation.*

(a) $\neg\sigma \in \overline{\mathsf{T}} \iff \text{NOT } \sigma \in \overline{\mathsf{T}}$

(b) $\wedge\sigma_1\sigma_2 \in \overline{\mathsf{T}} \iff \sigma_1 \in \overline{\mathsf{T}} \text{ AND } \sigma_2 \in \overline{\mathsf{T}}$

(c) $\vee\sigma_1\sigma_2 \in \overline{\mathsf{T}} \iff \sigma_1 \in \overline{\mathsf{T}} \text{ OR } \sigma_2 \in \overline{\mathsf{T}}$

(d) $\rightarrow\sigma_1\sigma_2 \in \overline{\mathsf{T}} \iff \text{IF } \sigma_1 \in \overline{\mathsf{T}} \text{ THEN } \sigma_2 \in \overline{\mathsf{T}}$

Proof. (a) Follows immediately from FACT 4.5.(b).

(b) First notice that by FACT 4.5.(d), $\wedge\sigma_1\sigma_2 \in \overline{\mathsf{T}}$ if and only if $\overline{\mathsf{T}} \vdash \wedge\sigma_1\sigma_2$. Thus, by L_3 and L_4 and (MP) we get $\overline{\mathsf{T}} \vdash \sigma_1$ and $\overline{\mathsf{T}} \vdash \sigma_2$. Therefore, by FACT 4.5.(d), we get $\sigma_1 \in \overline{\mathsf{T}}$ AND $\sigma_2 \in \overline{\mathsf{T}}$. On the other hand, if $\sigma_1 \in \overline{\mathsf{T}}$ AND $\sigma_2 \in \overline{\mathsf{T}}$, then, by FACT 4.5.(d), we get $\overline{\mathsf{T}} \vdash \sigma_1$ and $\overline{\mathsf{T}} \vdash \sigma_2$. Now, by TAUTOLOGY (B), this implies $\overline{\mathsf{T}} \vdash \wedge\sigma_1\sigma_2$, and by FACT 4.5.(d) we finally get $\wedge\sigma_1\sigma_2 \in \overline{\mathsf{T}}$.

(c) and (d) follow from FACT 4.5.(e) and from the 3-SYMBOLS THEOREM 1.7 which states that for each formula σ there is an equivalent formula σ' which contains neither $\vee$ nor $\rightarrow$. ⊣

EXERCISES

4.0 Show that all the logical axioms of propositional logic (i.e., L_0–L_9) were used in the proofs of FACT 4.0, LEMMA 4.1, FACT 4.5, and PROPOSITION 4.6. Notice that in the proof of FACT 4.0, we used FACT 2.16.(c) & (d).

4.1 The WEAK KÖNIG'S LEMMA is a very weak choice principle: A tree T is a 0-1-*tree* if it is a sub-tree of a binary tree in which the two successors of a node are always labelled with 0 and 1, respectively. Now, the WEAK KÖNIG'S LEMMA states that

every infinite 0-1*-tree contains an infinite branch.*

Show that in the proof of LINDENBAUM'S LEMMA 4.4, the LAW OF EXCLUDED MIDDLE can be replaced be the metamathematical WEAK KÖNIG'S LEMMA.

Hint: First, let Λ be the set of all finite lists $\lambda = [\neg\sigma_0, \varrho_1, \ldots, \varrho_n]$ of $\mathscr{L}$-sentences, where for each $1 \leq i \leq n$, either $\varrho_i \equiv \sigma_i$ or $\varrho_i \equiv \neg\sigma_i$. Now, encode formal proofs, which are finite sequences of $\mathscr{L}$-formulae, with natural numbers. Finally, construct the tree T consisting of all lists $\lambda \in \Lambda$, such that there is no formal proof of an inconsistency from $\mathsf{T} + \lambda$ with a code-number less than the length of λ. Then T corresponds to an infinite 0-1-tree with the property that each infinite branch through T corresponds to a maximally consistent set of $\mathscr{L}$-sentences.

4.2 Show that in the case when the theory $\mathsf{T} + \neg\sigma_0$ already has a model $\mathbf{M}$, then we can just set $\overline{\mathsf{T}} = \mathbf{Th}(\mathbf{M})$. Therefore, we do not need the LAW OF EXCLUDED MIDDLE in this case.

Chapter 5
The Completeness Theorem

As in the previous chapter, we require that all formulae are written in Polish notation and that the variables are among $v_0, v_1, v_2, \ldots$ Let $\mathscr{L}$ be a countable signature, where by "countable" we mean that the symbols in $\mathscr{L}$ can be listed in a FINITE or POTENTIALLY INFINITE list. Furthermore, let T be a consistent $\mathscr{L}$-theory and let σ_0 be an $\mathscr{L}$-sentence which is not provable from T. Finally, let $\overline{\mathsf{T}}$ be the maximally consistent extension of $\mathsf{T} + \neg\sigma_0$ obtained with LINDENBAUM'S LEMMA 4.4.

We shall construct a model of $\overline{\mathsf{T}}$ as follows: In a first step, we extend the signature $\mathscr{L}$ to a signature $\mathscr{L}_c$ by adding countably many new constant symbols, so-called *special constants*. In a second the step, we extend the $\mathscr{L}$-theory $\overline{\mathsf{T}}$ to an $\mathscr{L}_c$-theory $\overline{\mathsf{T}}_c$ by adding so-called *witnesses* to existential sentences in $\overline{\mathsf{T}}$. In particular, for each sentence $\exists\nu\sigma(\nu) \in \overline{\mathsf{T}}$ we add an $\mathscr{L}_c$-sentence $\sigma(c)$, where c is some special constant. In a third step, we extend the $\mathscr{L}_c$-theory $\overline{\mathsf{T}}_c$ to a maximally consistent $\mathscr{L}_c$-theory $\tilde{\mathsf{T}}$, and in a last step, we build the domain of the model of $\tilde{\mathsf{T}}$ as a list of lists of closed $\mathscr{L}_c$-terms.

Extending the Language

A string of symbols is called a **term-constant**, if it results from applying FINITELY many times the following rules:

(C0) Each closed (i.e., variable-free) $\mathscr{L}$-term is a term-constant.
(C1) If $\tau_0, \ldots, \tau_{n-1}$ are any term-constants which we have already built and F is an n-ary function symbol, then $F\tau_0 \cdots \tau_{n-1}$ is a term-constant.
(C2) For any natural numbers i, n, if $\tau_0, \ldots, \tau_{n-1}$ are any term-constants which we have already built, then $(i, \tau_0, \ldots, \tau_{n-1}, n)$ is a term-constant.

The strings $(i, \tau_0, \ldots, \tau_{n-1}, n)$ which are built with rule (C2) are called **special constants**. Notice that for $n = 0$, $(i, \tau_0, \ldots, \tau_{n-1}, n)$ becomes $(i, 0)$.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_6

Let $\mathscr{L}_c$ be the signature $\mathscr{L}$ extended by the countably many special constants. In order to write the special constants in a list, we first encode them and then define an ordering on the set of codes.

First we encode closed $\mathscr{L}$-terms with strings of 0's, 1's, and 2's as in Chapter 4. Now, let $c \equiv (i, \tau_0, \ldots, \tau_{n-1}, n)$ be a special constant, where the codes $\#\tau_0, \ldots, \#\tau_{n-1}$ of $\tau_0, \ldots, \tau_{n-1}$ are already defined. Then we encode c as follows:

$$\begin{array}{rcl} c & \equiv & (\quad \boldsymbol{i} \quad , \ \boldsymbol{\tau_0} \ , \quad \ldots \quad , \ \boldsymbol{\tau_{n-1}} \ , \quad \boldsymbol{n} \quad) \\ & & \downarrow \quad \downarrow \quad \downarrow \ \downarrow \ \downarrow \qquad\quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \\ \#c & \equiv & 6 \ \underbrace{1 \ldots 1}_{i\text{-times } 1} \ 8 \ \#\tau_0 \ 8 \quad \ldots \quad 8 \ \#\tau_{n-1} \ 8 \ \underbrace{1 \ldots 1}_{n\text{-times } 1} \ 9 \end{array}$$

The codes of special constants are ordered by their length and lexicographically, where $0 < 1 < 2 < 6 < 8 < 9$.

Finally, let $\Lambda_\tau = [\tau_0, \tau_1, \ldots]$ be the potentially infinite list of all term-constants, and let $\Lambda_c = [c_0, c_1, \ldots]$ be the potentially infinite list of all special constants, both ordered with respect to the ordering of their codes.

Extending the Theory

In this section, we shall add witnesses for certain existential $\mathscr{L}_c$-sentences σ_i in the list $\overline{\mathsf{T}} = [\neg\sigma_0, \sigma_1, \ldots, \sigma_i, \ldots]$, where an $\mathscr{L}_c$-sentence is existential if it is of the form $\exists\nu\varphi$. We choose the witnesses from the list Λ_c of special constants. In order to make sure that we have a witness for each existential $\mathscr{L}_c$-sentence (and not just for $\mathscr{L}$-sentences), and also to make sure that the choice of witnesses does not lead to a contradiction, we have to choose the witnesses carefully.

Let $\sigma_i \in \overline{\mathsf{T}}$ and let $c_j \equiv (i, t_0, \ldots, t_{n-1}, n)$ be a special constant. Then we say that c_j witnesses σ_i, or that c_j is a **witness** for σ_i, if

- $i \geq 1$ and σ_i is in special Prenex Normal Form sPNF (see Chapter 2),
- the string of symbols $\exists v_n$ appears in σ_i,
- for all $m < n$: if $\exists v_m$ appears in σ_i, then $t_m \equiv (i, t_0, \ldots, t_{m-1}, m)$.

On the one hand, we have only witnesses c_j for $\mathscr{L}$-sentences σ_i with $i \geq 1$. On the other hand, notice that since $\neg\sigma_0$ is not necessarily in sPNF, by construction of $\overline{\mathsf{T}}$ there exists an $i \geq 1$ such that σ_i and $\neg\sigma_0$ are semantically equivalent, which will be sufficient for our purposes.

If an $\mathscr{L}$-sentence $\sigma_i \in \overline{\mathsf{T}}$ is in sPNF and either $\exists v_n$ or $\forall v_n$ appears in σ_i, then

$$\sigma_i \equiv \forall_0 v_0 \forall_1 v_1 \cdots \forall_n v_n \sigma_{i,n}(v_0, \ldots, v_n)\,,$$

where $\sigma_{i,n}(v_0,\dots,v_n)$ is an $\mathscr{L}$-formula in which each variable among $v_0,\dots,v_n$ appears free. In particular, if $c_j \equiv (i,t_0,\dots,t_{n-1},n)$ witnesses σ_i, then

$$\sigma_i \equiv \forall v_0 \forall v_1 \cdots \forall v_{n-1} \exists v_n \sigma_{i,n}(v_0,\dots,v_n)\,,$$

i.e., $\exists v_n$ appears in σ_i. Furthermore, if $\sigma_i \in \overline{\mathsf{T}}$ is in sPNF, $c_j \equiv (i,t_0,\dots,t_{n-1},n)$ is a special constant, and c_j witnesses σ_i, then let

$$\sigma_{i,n}[c_j] :\equiv \sigma_{i,n}(v_0/t_0,\dots,v_{n-1}/t_{n-1},v_n/c_j)\,.$$

Now, we go through the list $\Lambda_c = [c_0,c_1,\dots]$ of special constants and extend step by step the list $\overline{\mathsf{T}} = [\sigma_0,\sigma_1,\dots]$. For this, we first stipulate $\mathsf{T}_0 := \overline{\mathsf{T}}$. Now assume that T_j is already defined and that $c_j \equiv (i,t_0,\dots,t_{n-1},n)$ for some natural numbers i,n and term-constants $t_0,\dots,t_{n-1}$. Then we have the following two cases:

Case 1. The special constant c_j does not witness the $\mathscr{L}$-sentence $\sigma_i \in \overline{\mathsf{T}}$. In this case, we set $\mathsf{T}_{j+1} := \mathsf{T}_j$.

Case 2. The special constant c_j witnesses $\sigma_i \in \overline{\mathsf{T}}$. In this case, we insert the $\mathscr{L}_c$-sentence $\sigma_{i,n}[c_j]$ into the list T_j at the place which corresponds to the code $\#\sigma_{i,n}[c_j]$. The extended list is then T_{j+1}.

Finally, let $\overline{\mathsf{T}}_c$ be the resulting list, i.e., $\overline{\mathsf{T}}_c$ is the union of all the T_j's.

LEMMA 5.0. *$\overline{\mathsf{T}}_c$ is consistent.*

Proof. By construction of $\overline{\mathsf{T}}$ we have $\mathrm{Con}(\overline{\mathsf{T}})$ with respect to the signature $\mathscr{L}$. We first show that $\overline{\mathsf{T}}$ is also consistent with respect to the signature $\mathscr{L}_c$: Assume towards a contradiction that $\overline{\mathsf{T}} \vdash \square$ with respect to the signature $\mathscr{L}_c$. In that proof, we replace each special constant c by a variable ν_c which does not occur in any of the finitely many formulae of the proof, such that ν_c and $\nu_{c'}$ are distinct variables whenever c and c' are distinct special constants. Notice that every logical axiom becomes a logical axiom of the same type. Moreover, notice that all $\mathscr{L}$-sentences of $\overline{\mathsf{T}}$ remain unchanged since they do not contain special constants. Furthermore, each application of Modus Ponens or Generalisation becomes a new application of the same inference rule. To see this, notice that we do not apply Generalisation to any of the ν_c's, since otherwise, we would have applied Generalisation to a special constant c, but c is a term-constant and not a variable. Since the obtained proof does not contain any special constants, we get $\overline{\mathsf{T}} \vdash \square$ (with respect to $\mathscr{L}$), which contradicts the fact that $\overline{\mathsf{T}}$ is consistent (with respect to $\mathscr{L}$). Therefore we have $\mathrm{Con}(\overline{\mathsf{T}})$ with respect to $\mathscr{L}_c$.

Now, assume towards a contradiction that $\overline{\mathsf{T}}_c$ is inconsistent, i.e., $\neg\,\mathrm{Con}(\overline{\mathsf{T}}_c)$. Then, by the COMPACTNESS THEOREM 2.17, we find finitely many pairwise distinct $\mathscr{L}_c$-sentences $\sigma_{i,n}[c_j]$ in $\overline{\mathsf{T}}_c$ such that

$$\neg\,\mathrm{Con}\left(\overline{\mathsf{T}} + \left\{\sigma_{i_1,n_1}[c_{j_1}],\dots,\sigma_{i_k,n_k}[c_{j_k}]\right\}\right).$$

Notice that since the $\mathscr{L}_c$-sentences $\sigma_{i_1,n_1}[c_{j_1}], \ldots, \sigma_{i_k,n_k}[c_{j_k}]$ are pairwise distinct, so are the special constants $c_{j_1}, \ldots, c_{j_k}$. Without loss of generality we may assume that $\sigma_{i_1,n_1}[c_{j_1}], \ldots, \sigma_{i_k,n_k}[c_{j_k}]$ are such that the sum $n_1 + \ldots + n_k + k$ is minimal.

For term-constants τ we define the **height** $h(\tau)$ as follows: If τ is a closed $\mathscr{L}$-term, then $h(\tau) := 0$; if $\tau_0, \ldots, \tau_{n-1}$ are term-constants and $F \in \mathscr{L}$ is an n-ary function symbol, then

$$h(F\tau_0 \cdots \tau_{n-1}) := \max\big\{h(\tau_0), \ldots, h(\tau_{n-1})\big\};$$

and finally, if $\tau \equiv (i, \tau_0, \ldots, \tau_{n-1}, n)$ is a special constant, then

$$h(\tau) := 1 + \max\big\{h(\tau_0), \ldots, h(\tau_{n-1})\big\},$$

where $\max \emptyset := 0$. Without loss of generality we may assume that

$$h(c_{j_k}) = \max\big\{h(c_{j_1}), \ldots, h(c_{j_k})\big\},$$

i.e., for each special constant c_j occurring in $c_{j_1}, \ldots, c_{j_k}$ we have $h(c_j) < h(c_{j_k})$. In particular, it follows that c_{j_k} does not occur in any such special constant c_j.

Let us now consider the formula $\sigma_{i_k,n_k}[c_{j_k}]$. In order to simplify the notation, we write i, n, j instead of i_k, n_k, j_k respectively; in particular, $\sigma_{i_k,n_k}[c_{j_k}]$ becomes $\sigma_{i,n}[c_j]$. Furthermore, let

$$\Sigma := \big\{\sigma_{i_1,n_1}[c_{j_1}], \ldots, \sigma_{i_{k-1},n_{k-1}}[c_{j_{k-1}}]\big\},$$

and let $c_j \equiv (i, t_0, \ldots, t_{n-1}, n)$, i.e.,

$$\sigma_{i,n}[c_j] \equiv \sigma_{i,n}(v_0/t_0, \ldots, v_{n-1}/t_{n-1}, v_n/c_j).$$

Since c_j witnesses σ_i, $\exists v_n$ appears in σ_i, i.e.,

$$\sigma_{i,n-1}(v_0, \ldots, v_{n-1}) \equiv \exists v_n \sigma_{i,n}(v_0, \ldots, v_{n-1}, v_n).$$

In order to simplify the notation again, we set

$$\tilde{\sigma}(v_n) :\equiv \sigma_{i,n}(v_0/t_0, \ldots, v_{n-1}/t_{n-1}, v_n).$$

Notice that v_n is the only variable which appears free in $\tilde{\sigma}$.

Claim. $\neg\,\mathrm{Con}\big(\overline{\mathsf{T}} + \Sigma + \sigma_{i,n}[c_j]\big) \quad \Longrightarrow \quad \neg\,\mathrm{Con}\big(\overline{\mathsf{T}} + \Sigma + \exists v_n \tilde{\sigma}(v_n)\big)$

Proof of Claim. If $\overline{\mathsf{T}} + \Sigma + \sigma_{i,n}[c_j]$ is inconsistent, then

$$\overline{\mathsf{T}} + \Sigma + \sigma_{i,n}[c_j] \vdash \not\Box, \tag{$\vdash_1$}$$

and with the Deduction Theorem we get

$$\overline{\mathsf{T}} + \Sigma \vdash \sigma_{i,n}[c_j] \to \text{⌀}. \tag{$\vdash_2$}$$

In the latter proof ($\vdash_2$) we replace the special constant c_j throughout the proof by a variable ν which does not occur in $\sigma_{i,n}$ and which does not occur in any of the finitely many formulae of the former proof ($\vdash_1$). Notice that every logical axiom becomes a logical axiom of the same type. Moreover, notice that $\mathscr{L}$-sentences of $\overline{\mathsf{T}}$ are not affected (since they do not contain special constants). Furthermore, $\mathscr{L}_c$-sentences of Σ are not affected either, since they do not contain the special constant c_j (recall that the special constants $c_{j_1}, \ldots, c_{j_k}$ are pairwise distinct). Finally, each application of Modus Ponens or Generalisation becomes a new application of the same inference rule (notice that we do not apply Generalisation to ν, since otherwise we would have applied Generalisation to c_j, but c_j is a term-constant). Now, we construct a proof of $\exists v_n \tilde{\sigma}(v_n) \to \text{⌀}$ from $\overline{\mathsf{T}} + \Sigma$ as follows:

$\overline{\mathsf{T}} + \Sigma \vdash \tilde{\sigma}(\nu) \to \text{⌀}$	by assumption
$\overline{\mathsf{T}} + \Sigma \vdash \forall \nu \big(\tilde{\sigma}(\nu) \to \text{⌀}\big)$	by Generalisation
$\overline{\mathsf{T}} + \Sigma \vdash \forall \nu \big(\tilde{\sigma}(\nu) \to \text{⌀}\big) \to \big(\exists \nu \tilde{\sigma}(\nu) \to \text{⌀}\big)$	L_{13}
$\overline{\mathsf{T}} + \Sigma \vdash \exists \nu \tilde{\sigma}(\nu) \to \text{⌀}$	by Modus Ponens
$\overline{\mathsf{T}} + \Sigma \vdash \exists v_n \tilde{\sigma}(v_n) \to \text{⌀}$	Tautology (Q.1)

Therefore, we finally have $\neg \operatorname{Con}\big(\overline{\mathsf{T}} + \Sigma + \exists v_n \tilde{\sigma}(v_n)\big)$. $\dashv_{\text{Claim}}$

Let us now consider σ_i. Let $m \leq n$ be the largest natural number such that for each l with $1 \leq l \leq m$ we have that $\forall v_{n-l}$ appears in σ_i. For example, if $m = 0$, then $\nexists_{n-1}$ is the quantifier $\exists$, and if $m = n$, then for no $n' < n$, $\nexists_{n'}$ is the quantifier $\exists$. In general, σ_i is of the form

$$\sigma_i \equiv \underbrace{\nexists_0 v_0 \quad \cdots}_{\exists \text{ or } \forall} \exists v_{n-m-1} \underbrace{\forall v_{n-m} \quad \cdots \quad \forall v_{n-1}}_{\text{only } \forall} \exists v_n \, \sigma_{i,n}(v_0, \ldots, v_n) \,.$$

Consider now the formula

$$\tilde{\sigma}_m :\equiv \sigma_{i,n-m-1}(v_0/t_0, \ldots, v_{n-m-1}/t_{n-m-1}) \,.$$

Then either $\tilde{\sigma}_m \in \overline{\mathsf{T}}$ (in case $m = n$), or $\exists v_{n-m-1}$ appears in σ_i (in case $m < n$), and therefore, we are in one of following two cases:

Case 1. $\tilde{\sigma}_m \in \overline{\mathsf{T}}$: First notice that in this case,

$$\tilde{\sigma}_m \equiv \forall v_0 \cdots \forall v_{n-1} \exists v_n \sigma_{i,n}(v_0, \ldots, v_n) \,.$$

Since $\tilde{\sigma}_m \in \overline{\mathsf{T}}$ and $t_0, \ldots, t_{n-1}$ are closed terms, by L_{10} we get $\overline{\mathsf{T}} \vdash \exists v_n \tilde{\sigma}(v_n)$. Hence, by the Claim, $\neg \operatorname{Con}(\overline{\mathsf{T}} + \Sigma)$. This shows that we do not need $\sigma_{i_k,n_k}[c_{j_k}]$ to derive a contradiction from

$$\overline{\mathsf{T}} + \left\{\sigma_{i_1,n_1}[c_{j_1}], \ldots, \sigma_{i_k,n_k}[c_{j_k}]\right\},$$

which is a contradiction to the minimality of the sum $n_1 + \ldots n_k + k$.

Case 2. $\exists v_{n-m-1}$ appears in σ_i: Note that since $c_j \equiv (i, t_0, \ldots, t_{n-1}, n)$ witnesses σ_i,

$$t_{n-m-1} \equiv (i, t_0, \ldots, t_{n-m-2}, n-m-1)$$

witnesses σ_i, too. Similar as above, by L_{10} we get

$$\overline{\mathsf{T}} + \sigma_{i,n-m-1}[t_{n-m-1}] \vdash \exists v_n \tilde{\sigma}(v_n),$$

and with the Deduction Theorem we obtain

$$\overline{\mathsf{T}} \vdash \sigma_{i,n-m-1}[t_{n-m-1}] \to \exists v_n \tilde{\sigma}(v_n).$$

This shows that if we derive a contradiction from

$$\overline{\mathsf{T}} + \Sigma + \exists v_n \tilde{\sigma}(v_n),$$

then we also derive a contradiction from

$$\overline{\mathsf{T}} + \Sigma + \sigma_{i,n-m-1}[t_{n-m-1}],$$

which is again a contradiction to the minimality of the sum $n_1 + \ldots n_k + k$.

Therefore, $\overline{\mathsf{T}} + \left\{\sigma_{i_1,n_1}[c_{j_1}], \ldots, \sigma_{i_k,n_k}[c_{j_k}]\right\}$ is consistent, and since the finitely many $\mathscr{L}_c$-sentences $\sigma_{i_1,n_1}[c_{j_1}], \ldots, \sigma_{i_k,n_k}[c_{j_k}]$ were arbitrary, we obtain that $\overline{\mathsf{T}}_c$ is consistent, which completes the proof. ⊣

The Completeness Theorem for Countable Signatures

In this section, we shall construct a model of the $\mathscr{L}_c$-theory $\overline{\mathsf{T}}_c$, which is of course also a model of the $\mathscr{L}$-theory $\mathsf{T} + \neg\sigma_0$. However, since we extended the signature $\mathscr{L}$, we first have to extend the binary relation $=$, as well as relation symbols in $\mathscr{L}$, to the new closed $\mathscr{L}_c$-terms.

Lemma 5.1. *The list $\overline{\mathsf{T}}_c$ can be extended to a consistent list $\tilde{\mathsf{T}}$ of $\mathscr{L}_c$-sentences, such that the additional $\mathscr{L}_c$-sentences are variable-free (i.e., they contain neither quantifiers nor free variables) and for each variable-free $\mathscr{L}_c$-sentence σ we have*

$$\textit{either} \quad \sigma \in \tilde{\mathsf{T}} \quad \textit{or} \quad \neg\sigma \in \tilde{\mathsf{T}}.$$

Proof. Like in the proof of Lindenbaum's Lemma 4.4, we go through the list of all variable-free $\mathscr{L}_c$-sentences and successively extend the list $\overline{\mathsf{T}}_c$ to a consistent list $\tilde{\mathsf{T}}$. ⊣

Notice that by the construction of $\overline{\mathsf{T}}_c$, if the $\mathscr{L}_c$-sentence

$$\underbrace{\forall v_k \cdots \forall v_{n-1}}_{\text{only } \forall\text{-quantifiers}} \exists v_n \, \sigma_{i,n}(v_0/t_0, \ldots, v_{k-1}/t_{k-1}, v_k, \ldots, v_n)$$

belongs to $\overline{\mathsf{T}}_c$, where $n-1 \geq k$, then for all $t_k, \ldots, t_{n-1} \in \Lambda_\tau$ and

$$c_j \equiv (i, t_0, \ldots, t_{k-1}, t_k, \ldots, t_{n-1}, n)$$

the sentence

$$\sigma_{i,n}(v_0/t_0, \ldots, v_{k-1}/t_{k-1}, v_k/t_k, \ldots, v_{n-1}/t_{n-1}, v_n/c_j)$$

belongs to $\overline{\mathsf{T}}_c$, and hence, this sentence belongs to $\tilde{\mathsf{T}}$. The reason is because in the witness c_j, $t_k, \ldots, t_{n-1}$ can be any term-constants in Λ_τ.

Furthermore, as a consequence of the construction of $\tilde{\mathsf{T}}$ we obtain the following result.

FACT 5.2. *If the $\mathscr{L}_c$-sentence*

$$\sigma \equiv \underbrace{\forall v_k \cdots \forall v_n}_{\textit{only } \forall\textit{-quantifiers}} \sigma_{i,n}(v_0/t_0, \ldots, v_{k-1}/t_{k-1}, v_k, \ldots, v_n)$$

belongs to $\overline{\mathsf{T}}_c$, where $\sigma_{i,n}$ does not contain quantifiers, then for all $t_k, \ldots, t_n \in \Lambda_\tau$ the sentence

$$\sigma' \equiv \sigma_{i,n}(v_0/t_0, \ldots, v_{k-1}/t_{k-1}, v_k/t_k, \ldots, v_n/t_n)$$

belongs to $\tilde{\mathsf{T}}$.

Proof. If $\sigma \in \overline{\mathsf{T}}_c$, then by L_{10} we have $\overline{\mathsf{T}}_c \vdash \sigma'$. Thus, by construction of $\tilde{\mathsf{T}}$ we get $\sigma' \in \tilde{\mathsf{T}}$. ⊣

Now we are ready to construct the domain of a model of $\tilde{\mathsf{T}}$, which shall be a list of lists. For this, let

$$\Lambda_\tau = [t_0, t_1, \ldots, t_n, \ldots]$$

be the list of all term-constants (ordered with respect to their codes). We go through the list Λ_τ and construct step by step a list of lists: First, we set $A_0 := [\,]$. Now, assume that A_n is already defined and consider the $\mathscr{L}_c$-sentences

$$t_n = t_0, \quad t_n = t_1, \quad \ldots \quad t_n = t_{n-1}\,.$$

If, for some m with $0 \leq m < n$, the sentence $t_n = t_m$ belongs to $\tilde{\mathsf{T}}$, then we append t_n to that list in A_n which contains t_m; the resulting list is A_{n+1}. If none of the sentences $t_n = t_m$ (for $0 \leq m < n$) belongs to $\tilde{\mathsf{T}}$ (e.g., if

$n = 0$), then $A_{n+1} := A_n + \big[[t_n]\big]$. Finally, let $A = \big[[t_{n_0}, \ldots], [t_{n_1}, \ldots] \ldots\big]$ be the resulting list. Then, A is a finite or potentially infinite list of potentially infinite lists.

The lists in the list A are the objects of the domain of our model $\mathbf{M} \models \tilde{\mathsf{T}}$. In order to simplify the notation, for term-constants τ let $\tilde{\tau}$ be the unique list of A which contains τ.

In order to get an $\mathscr{L}_c$-structure $\mathbf{M}$ with domain A, we have to define a mapping which assigns to each constant symbol $c \in \mathscr{L}_c$ an element $c^{\mathbf{M}} \in A$, to each n-ary function symbol $F \in \mathscr{L}$ a function $F^{\mathbf{M}} : A^n \to A$, and to each n-ary relation symbol $R \in \mathscr{L}$ a set $R^{\mathbf{M}} \subseteq A^n$:

- If $c \in \mathscr{L}_c$ is a constant symbol of $\mathscr{L}$ or a special constant, then let
$$c^{\mathbf{M}} := \tilde{c}\,.$$
- If $F \in \mathscr{L}$ is an n-ary function symbol and $\tilde{t}_1, \ldots, \tilde{t}_n$ are elements of A, then let
$$F^{\mathbf{M}}\tilde{t}_1 \cdots \tilde{t}_n := \widetilde{Ft_1 \cdots t_n}\,.$$
- If $R \in \mathscr{L}$ is an n-ary relation symbol and $\tilde{t}_1, \ldots, \tilde{t}_n$ are elements of A, then we define
$$\langle \tilde{t}_1, \ldots, \tilde{t}_n \rangle \in R^{\mathbf{M}} \quad :\Longleftrightarrow \quad Rt_1 \cdots t_n \in \tilde{\mathsf{T}}\,.$$

Fact 5.3. *The definitions above, which rely on representatives of the lists in A, are well-defined.*

Proof. This follows easily by L_{14}, L_{15}, and L_{16}, and the construction of $\tilde{\mathsf{T}}$; the details are left as an exercise to the reader. ⊣

Theorem 5.4. *The $\mathscr{L}_c$-structure $\mathbf{M}$ is a model of $\tilde{\mathsf{T}}$, and therefore also of the $\mathscr{L}$-theory $\mathsf{T} + \neg\sigma_0$.*

Proof. We have to show that for each $\mathscr{L}_c$-sentence σ we have

$$\sigma \in \tilde{\mathsf{T}} \implies \mathbf{M} \models \sigma\,, \quad \text{or equivalently} \quad \mathbf{M} \not\models \sigma \implies \sigma \notin \tilde{\mathsf{T}}\,.$$

First, we consider the case when σ is variable-free: The proof is by induction on the number of logical operators. By Lemma 5.1 we know that for each variable-free $\mathscr{L}_c$-sentence σ we have either $\sigma \in \tilde{\mathsf{T}}$ or $\neg\sigma \in \tilde{\mathsf{T}}$. Hence, we must show that for each variable-free $\mathscr{L}_c$-sentences σ we have $\sigma \in \tilde{\mathsf{T}}$ if and only if $\mathbf{M} \models \sigma$.

If σ is variable-free and does not contain logical operators, then σ is atomic. In this case, we have either $\sigma \equiv t_1 = t_2$ (for some term-constants t_1 and t_1) or $\sigma \equiv Rt_1 \cdots t_n$ (for an n-ary relation symbol $R \in \mathscr{L}$ and term-constants $t_1, \ldots, t_n$), and by construction of $\mathbf{M}$, we get $\sigma \in \tilde{\mathsf{T}}$ if and only if $\mathbf{M} \models \sigma$ in both cases.

Before we consider the case when σ is variable-free and contains logical operators, recall that for any $\mathscr{L}_c$-sentence $\tilde{\sigma}$ with $\sigma \Leftrightarrow \tilde{\sigma}$, by the SOUNDNESS THEOREM 3.7 we get $\mathbf{M} \vDash \sigma$ if and only if $\mathbf{M} \vDash \tilde{\sigma}$. Therefore, by the THREE-SYMBOLS THEOREM 1.7, we may assume that σ is either of the form $\neg\sigma'$ or of the form $\wedge\sigma_1\sigma_2$. Now, let σ be a non-atomic, variable-free $\mathscr{L}_c$-sentence, and assume that for each variable-free $\mathscr{L}_c$-sentence σ' which contains fewer logical operators than σ, we have $\sigma' \in \tilde{\mathsf{T}}$ if and only if $\mathbf{M} \vDash \sigma'$. By our former assumption, we just have to consider the following two cases:

Case 1. $\sigma \equiv \neg\sigma'$: Since σ' has fewer logical operators than σ, we have $\sigma' \in \tilde{\mathsf{T}}$ if and only if $\mathbf{M} \vDash \sigma'$. This shows that

$$\neg\sigma' \notin \tilde{\mathsf{T}} \iff \mathbf{M} \nvDash \neg\sigma', \quad \text{or equivalently} \quad \sigma \in \tilde{\mathsf{T}} \iff \mathbf{M} \vDash \sigma.$$

Case 2. $\sigma \equiv \wedge\sigma_1\sigma_2$: Since each if σ_1 and σ_2 has fewer logical operators than $\tilde{\sigma}$, we have $\sigma_1 \in \tilde{\mathsf{T}}$ if and only if $\mathbf{M} \vDash \sigma_1$, and $\sigma_2 \in \tilde{\mathsf{T}}$ if and only if $\mathbf{M} \vDash \sigma_2$. Hence, we obtain

$$\wedge\sigma_1\sigma_2 \in \tilde{\mathsf{T}} \iff \sigma_1 \in \tilde{\mathsf{T}} \text{ AND } \sigma_2 \in \tilde{\mathsf{T}} \iff$$
$$\mathbf{M} \vDash \sigma_1 \text{ AND } \mathbf{M} \vDash \sigma_2 \iff \mathbf{M} \vDash \wedge\sigma_1\sigma_2$$

which shows that $\sigma \in \tilde{\mathsf{T}} \iff \mathbf{M} \vDash \sigma$.

We now consider the case when the $\mathscr{L}_c$-sentence $\sigma \in \tilde{\mathsf{T}}$ contains variables: The proof is by induction on the number of different variables which appear in σ. If $\sigma \in \tilde{\mathsf{T}}$ is an $\mathscr{L}_c$-sentence which contains variables, then, by construction of $\overline{\mathsf{T}}_c$, there is a $\tilde{\sigma} \in \overline{\mathsf{T}}_c$ in sPNF, say

$$\tilde{\sigma} \equiv \forall_0 v_0 \cdots \forall_n v_n \sigma_{i,n}(v_0, \ldots, v_n), \quad \text{where } \sigma_{i,n} \text{ is quantifier free,}$$

such that for some natural numbers i, k, n with $k \leq n$ and some term-constants $t_0, \ldots, t_{k-1}$ we have

$$\sigma \equiv \forall_k v_k \cdots \forall_n v_n \sigma_{i,n}(v_0/t_0, \ldots, v_{k-1}/t_{k-1}, v_k, \ldots, v_n).$$

Now, let σ be an $\mathscr{L}_c$-sentence of the above form and assume that for each $\mathscr{L}_c$-sentence σ' which contains fewer variables than σ we have

$$\sigma' \in \tilde{\mathsf{T}} \implies \mathbf{M} \vDash \sigma', \quad \text{or equivalently} \quad \mathbf{M} \nvDash \sigma' \implies \sigma' \notin \tilde{\mathsf{T}}.$$

We are in exactly one of the following two cases:

Case 1. $\forall_k$ is the quantifier $\exists$: Suppose that $\sigma \in \tilde{\mathsf{T}}$. Then we have $\sigma \in \overline{\mathsf{T}}_c$ and for the special constant

$$t_k \equiv (i, t_0, \ldots, t_{k-1}, k)$$

and the $\mathscr{L}_c$-sentence

$$\sigma' \equiv \forall_{k+1} v_{k+1} \cdots \forall_n v_n \sigma_{i,n}(v_0/t_0, \ldots, v_k/t_k, v_{k+1}, \ldots),$$

we have $\sigma' \in \overline{\mathsf{T}}_c$, and consequently $\sigma' \in \tilde{\mathsf{T}}$. Now, since σ' has fewer variables than σ, by our assumption we conclude that $\mathbf{M} \vDash \sigma'$, and therefore, by L_{11} and the SOUNDNESS THEOREM 3.7, we obtain $\mathbf{M} \vDash \sigma$. Hence,

$$\sigma \in \tilde{\mathsf{T}} \implies \mathbf{M} \vDash \sigma\,.$$

Case 2. $\exists_k$ is the quantifier $\forall$: Suppose that $\sigma \in \tilde{\mathsf{T}}$, where

$$\sigma \equiv \underbrace{\forall v_k \cdots \forall v_{k+s}}_{\text{only } \forall\text{-quantifiers}} \exists v_{k+s+1}\, \sigma_{i,k+s+1}(v_0/t_0, \ldots, v_{k-1}/t_{k-1}, v_k, \ldots, v_{k+s+1})$$

or

$$\sigma \equiv \underbrace{\forall v_k \cdots \forall v_{k+s}}_{\text{only } \forall\text{-quantifiers}} \sigma_{i,k+s}(v_0/t_0, \ldots, v_{k-1}/t_{k-1}, v_k, \ldots, v_{k+s})$$

where $\sigma_{i,k+s}$ does not contain quantifiers. Then, by FACT 5.2 and the construction of $\tilde{\mathsf{T}}$, for all $t_k, \ldots, t_{k+s} \in \Lambda_\tau$ and

$$c_j \equiv (i, t_0, \ldots, t_{k-1}, t_k, \ldots, t_{k+s}, k+s+1)\,,$$

in the former case we have

$$\underbrace{\sigma_{i,k+s+1}(v_0/t_0, \ldots, v_{k+s}/t_{k+s}, v_{k+s+1}/c_j)}_{\sigma'} \in \tilde{\mathsf{T}}\,,$$

and in the latter case we have

$$\underbrace{\sigma_{i,k+s}(v_0/t_0, \ldots, v_{k+s}/t_{k+s})}_{\sigma'} \in \tilde{\mathsf{T}}\,.$$

In both cases, σ' has fewer variables than σ and by our assumption we have $\mathbf{M} \vDash \sigma'$. Thus, by FACT 3.0.(b), for all $\tilde{t}_k, \ldots, \tilde{t}_{k+s} \in A$ we have

$$\left(\mathbf{M}, j\,\tfrac{\tilde{t}_k}{v_k} \cdots \tfrac{\tilde{t}_{k+s}}{v_{k+s}}\right) \vDash \exists v_{k+s+1} \sigma_{i,k+s+1}(v_0/t_0, \ldots, v_{k+s}/t_{k+s}, v_{k+s+1})$$

or

$$\left(\mathbf{M}, j\,\tfrac{\tilde{t}_k}{v_k} \cdots \tfrac{\tilde{t}_{k+s}}{v_{k+s}}\right) \vDash \sigma_{i,k+s}(v_0/t_0, \ldots, v_{k+s}/t_{k+s})\,,$$

and in both cases we have $\mathbf{M} \vDash \sigma$.

So, for each $\mathscr{L}_c$-sentence σ we have that $\sigma \in \tilde{\mathsf{T}}$ implies $\mathbf{M} \vDash \sigma$. This shows that $\mathbf{M} \vDash \tilde{\mathsf{T}}$, in particular, $\mathbf{M} \vDash \mathsf{T} + \neg\sigma_0$. ⊣

The following theorem just summarises what we have achieved so far:

THEOREM 5.5 (GÖDEL'S COMPLETENESS THEOREM). *If $\mathscr{L}$ is a countable signature and T is a consistent set of $\mathscr{L}$-sentences, then T has a model. Moreover, if $\mathsf{T} \nvdash \sigma_0$ (for some $\mathscr{L}$-sentence σ_0), then $\mathsf{T} + \neg\sigma_0$ has a model.*

In our construction, it was essential that the signature $\mathscr{L}$ was countable, so that the symbols in $\mathscr{L}$ could be encoded by finite strings. However, in the more formal setting of axiomatic Set Theory, we can also prove the Completeness Theorem for arbitrarily large signatures (see Chapter 15).

Some Consequences and Equivalents

We conclude this chapter by discussing some consequences and equivalent formulations of Gödel's Completeness Theorem 5.5 which follow directly or in combination with the Compactness Theorem 2.17.

Let $\mathscr{L}$ be a countable signature, T a set of $\mathscr{L}$-sentences, and σ_0 an $\mathscr{L}$-sentence.

- If $\mathsf{T} \nvdash \sigma_0$, then there is an $\mathscr{L}$-structure $\mathbf{M}$ such that $\mathbf{M} \vDash \mathsf{T} + \neg\sigma_0$:

$$\mathsf{T} \nvdash \sigma_0 \quad \Longrightarrow \quad \exists\, \mathbf{M}\, (\mathbf{M} \vDash \mathsf{T} + \neg\sigma_0)$$

  This is just a reformulation of Gödel's Completeness Theorem 5.5.

- If T is consistent, then T has a model:

$$\mathrm{Con}(\mathsf{T}) \quad \Longrightarrow \quad \exists\, \mathbf{M}\, (\mathbf{M} \vDash \mathsf{T})$$

  This follows from the fact that $\mathrm{Con}(\mathsf{T})$ is equivalent to the existence of an $\mathscr{L}$-sentence σ_0 such that $\mathsf{T} \nvdash \sigma_0$.

- If each model of T is also a model of σ_0, then $\mathsf{T} \vdash \sigma_0$:

$$\forall\, \mathbf{M}\, (\mathbf{M} \vDash \mathsf{T} \ \Longrightarrow\ \mathbf{M} \vDash \sigma_0) \quad \Longrightarrow \quad \mathsf{T} \vdash \sigma_0$$

  This follows by contraposition: If $\mathsf{T} \nvdash \sigma_0$, then, by Gödel's Completeness Theorem 5.5, there is a model $\mathbf{M} \vDash \mathsf{T} + \neg\sigma_0$.

- In combination with the Compactness Theorem 2.17, we obtain the following implication:

  If every finite subset T' of T has a model, then T has a model.

  If every finite subset T' of T has a model, then every finite subset T' of T is consistent, and therefore, by the Compactness Theorem 2.17, T is consistent. Thus, T has a model.

The most important consequence of Gödel's Completeness Theorem 5.5 and the Soundness Theorem 3.7 is the following equivalence:

$$\underbrace{\forall\, \mathbf{M}\, (\mathbf{M} \vDash \mathsf{T} \ \Longrightarrow\ \mathbf{M} \vDash \sigma_0)}_{\text{denoted by } \mathsf{T} \vDash \sigma_0} \quad \Longleftrightarrow \quad \mathsf{T} \vdash \sigma_0$$

This equivalence allows us to replace *formal proofs* by *mathematical proofs*. For example, instead of proving formally the uniqueness of the neutral element in groups from the axioms of Group Theory GT, we just show that in every model of GT (i.e., in every group), the neutral element is unique. So, instead of $\mathsf{GT} \vdash \sigma_0$, we just show $\mathsf{GT} \models \sigma_0$.

As a last consequence, we would like to mention the so-called *Skolem's Paradox*, which is in fact just the countable version of the Downward Löwenheim–Skolem Theorem 15.8.

Theorem 5.6 (Skolem's Paradox). *If $\mathscr{L}$ is a countable signature and T is a consistent set of $\mathscr{L}$-sentences, then T has a countable model.*

Proof. In the previous chapter, when we have constructed the universal list of $\mathscr{L}$-sentences, we began with a countable signature $\mathscr{L}$ and a consistent set of $\mathscr{L}$-sentences T, and at the end, we obtained a model of T whose domain was a finite or potentially infinite list of lists. So, the model of T which we constructed is countable. ⊣

What is paradoxical about this statement? For example, the signature of the axioms of Zermelo-Fraenkel Set Theory ZFC only consists of the membership relation $\in$. Hence, if ZFC is consistent, it has a countable model. However, it is easy to prove from the axiom system ZFC that there exist uncountable sets. Nevertheless, this is no contradiction, since countability in the formal theory and on the metalevel simply do not coincide.

Notes

The Completeness Theorem for countable signatures was first proved by Gödel [14, 15]. Later, a modified proof was given by Henkin [23] (see also [25]). The proof given here is essentially Henkin's proof, but in contrast to Henkin's proof, our construction does not rely on the assumption that an *actually infinite* set exists (in Halbeisen [21] a framework for metamathematics is provided which is sufficient to build models of first-order theories).

Exercises

5.0 Let $\mathscr{L}$ be a countable signature and let T be a consistent $\mathscr{L}$-theory. Show that if T has up to isomorphisms a unique model, then T is complete.

5.1 Let $\mathscr{L}$ be a countable signature, let Σ be the collection of all $\mathscr{L}$-structures $\mathbf{M}$, and for each $\mathscr{L}$-sentence φ, let X_φ be the collection of all $\mathbf{M} \in \Sigma$ such that $\mathbf{M} \models \varphi$.

(a) Show that the set $\{X_\varphi : \varphi \text{ is an } \mathscr{L}\text{-sentence}\}$ is a basis for a topology on Σ.

(b) Show that X_φ is closed for each $\mathscr{L}$-sentence φ.

(c) Show that the topological space Σ is compact, i.e., show that each open covering of Σ contains a finite sub-covering.

5.2 Let DLO be the — assumingly consistent — theory of dense linearly ordered sets without endpoints (see Exercise 3.4).

(a) Show that the theory DLO is complete, i.e., for all $\mathscr{L}_{\mathsf{DLO}}$-sentences σ we have *either* $\mathsf{DLO} \vdash \sigma$ *or* $\mathsf{DLO} \vdash \neg\sigma$.

(b) Show that the converse of Exercise 3.3 does not hold.

Hint: Consider $(\mathbb{Q}, <)$ and $(\mathbb{R}, <)$ where $\mathbb{Q}$ is the set of rational numbers, $\mathbb{R}$ the set of real numbers and $<$ is the natural ordering on $\mathbb{Q}$ and $\mathbb{R}$, respectively.

5.3 Let $\mathscr{L}$ be a countable signature and let T be a consistent set of $\mathscr{L}$-sentences such that T has arbitrarily large finite models. Show that T has an infinite model.

Hint: Use the Compactness Theorem 2.17.

5.4 Let $\mathscr{L}$ be the language that consists of two constant symbols $\mathtt{0}$ and $\mathtt{1}$, two binary function symbols $+$ and $\cdot$, and a binary relation symbol $<$. Let $\mathbf{M} = (\mathbb{Q}, 0, 1, +, \cdot, <)$ be the usual $\mathscr{L}$-interpretation and let $\mathsf{T} = \mathbf{Th}(\mathbf{M})$.

Prove that there is a model of T which contains an infinitesimal number, i.e., a number $x > \mathtt{0}$ such that $n \cdot x < \mathtt{1}$ for all natural numbers n.

Hint: Note that n is not included in the language $\mathscr{L}$. Hence, you need to find a way to formalise n within the language.

Chapter 6
Language Extensions by Definitions

Sometimes it is convenient to extend a given signature $\mathscr{L}$ by adding new non-logical symbols which have to be properly defined within the language $\mathscr{L}$ or with respect to a given $\mathscr{L}$-theory T. Let the extended signature be $\mathscr{L}^*$ and let the corresponding extended $\mathscr{L}^*$-theory be T^*. Since T is an $\mathscr{L}$-theory, we can only prove $\mathscr{L}$-sentences from T but no $\mathscr{L}^*$-sentences which contain symbols from $\mathscr{L}^* \setminus \mathscr{L}$. However, this does not imply that we can prove substantially more from T^* than from T: It might be that for each $\mathscr{L}^*$-sentence σ^* which is provable from T^* there is an $\mathscr{L}$-sentence $\tilde{\sigma}$ such that $\mathsf{T}^* \vdash \sigma^* \leftrightarrow \tilde{\sigma}$ and $\mathsf{T} \vdash \tilde{\sigma}$ which is indeed the case as we shall see below.

Defining new Relation Symbols

Let us first consider an example from Peano Arithmetic: Extend the signature $\mathscr{L}_{\mathsf{PA}}$ of Peano Arithmetic by adding the binary relation symbol $<$ and denote the extended signature by $\mathscr{L}^*_{\mathsf{PA}} := \mathscr{L}_{\mathsf{PA}} \cup \{<\}$. In order to define the binary relation $<$, we give an $\mathscr{L}_{\mathsf{PA}}$-formula $\psi_<$ with two free variables (e.g., x and y) and say that the relation $x < y$ holds if and only if $\psi_<(x, y)$ holds. In our case, $\psi_<(x, y) \equiv \exists z(x + \mathtt{s}z = y)$. Therefore, we would define the symbol $<$ by stipulating:

$$x < y :\Longleftrightarrow \exists z(x + \mathtt{s}z = y)$$

The problem is now to find for each $\mathscr{L}^*_{\mathsf{PA}}$-sentence σ^* an $\mathscr{L}_{\mathsf{PA}}$-sentence $\tilde{\sigma}$ and an extension PA^* of PA, such that $\mathsf{PA}^* \vdash \sigma^* \leftrightarrow \tilde{\sigma}$ and $\mathsf{PA} \vdash \tilde{\sigma}$ whenever $\mathsf{PA}^* \vdash \sigma^*$.

The following result provides an algorithm which transforms sentences σ^* in the extended language into equivalent sentences $\tilde{\sigma}$ in the original language. In order to prove that the algorithm works, we will make use of Gödel's Completeness Theorem 5.5.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_7

THEOREM 6.0. *Let $\mathscr{L}$ be a countable signature, let R be an n-ary relation symbol which does not belong to $\mathscr{L}$, and let $\mathscr{L}^* := \mathscr{L} \cup \{R\}$. Furthermore, let $\psi_R(v_1, \ldots, v_n)$ be an $\mathscr{L}$-formula with $\mathrm{free}(\psi_R) = \{v_1, \ldots, v_n\}$ and let*

$$\vartheta_R \equiv \forall v_1 \cdots \forall v_n \big(R v_1 \cdots v_n \leftrightarrow \psi_R(v_1, \ldots, v_n) \big) .$$

Finally, let T be a consistent $\mathscr{L}$-theory and let $\mathsf{T}^* := \mathsf{T} + \vartheta_R$. Then there exists an effective algorithm which transforms each $\mathscr{L}^*$-formula φ^* into an $\mathscr{L}$-formula $\tilde{\varphi}$ such that:

(a) *If R does not appear in φ^*, then $\tilde{\varphi} \equiv \varphi^*$.*

(b) $\widetilde{\neg\varphi} \equiv \neg\tilde{\varphi}$ *(for $\varphi^* \equiv \neg\varphi$)*

(c) $\widetilde{\varphi_1 \wedge \varphi_2} \equiv \tilde{\varphi}_1 \wedge \tilde{\varphi}_2$ *(for $\varphi^* \equiv \varphi_1 \wedge \varphi_2$)*

(d) $\widetilde{\exists \nu \varphi} \equiv \exists \nu \tilde{\varphi}$ *(for $\varphi^* \equiv \exists \nu \varphi$)*

(e) $\mathsf{T}^* \vdash \varphi^* \leftrightarrow \tilde{\varphi}$

(f) *If $\mathsf{T}^* \vdash \varphi^*$, then $\mathsf{T} \vdash \tilde{\varphi}$.*

Proof. Let φ^* be an arbitrary $\mathscr{L}^*$-formula. In φ^*, we replace each occurrence of $R(v_1/\tau_1, \ldots, v_n/\tau_n)$ (where $\tau_1, \ldots, \tau_n$ are $\mathscr{L}$-terms) by a particular $\mathscr{L}^*$-formula $\psi'_R(v_1/\tau_1, \ldots, v_n/\tau_n)$ such that

$$\psi'_R(v_1, \ldots, v_n) \Leftrightarrow_{\mathsf{T}} \psi_R(v_1, \ldots, v_n)$$

and none of the bound variables in ψ'_R is among $v_1, \ldots, v_n$ or appears in one of the $\mathscr{L}$-terms $\tau_1, \ldots, \tau_n$. In fact, in order to obtain ψ'_R we just have to rename the bound variables in ψ_R using the VARIABLE SUBSTITUTION THEOREM. For the resulting $\mathscr{L}$-formula $\tilde{\varphi}$, (a)–(d) are obviously satisfied.

We prove (e) and (f) on the semantic level: For this, we first show how we can extend a model $\mathbf{M} \models \mathsf{T}$ to a model $\mathbf{M}^* \models \mathsf{T}^*$. Let $\mathbf{M}$ be an $\mathscr{L}$-structure with domain A such that for each assignment j we have $(\mathbf{M}, j) \models \mathsf{T}$ (i.e., $\mathbf{M} \models \mathsf{T}$). We extend $\mathbf{M}$ to an $\mathscr{L}^*$-structure $\mathbf{M}^*$ with the same domain A by stipulating $\mathbf{M}^*|_{\mathscr{L}} := \mathbf{M}$, and for any $a_1, \ldots, a_n \in A$:

$$R^{\mathbf{M}^*}(a_1, \ldots, a_n) \;:\Longleftrightarrow\; \big(\mathbf{M}, j\tfrac{a_1}{v_1} \cdots \tfrac{a_n}{v_n}\big) \models \psi_R(v_1, \ldots, v_n) .$$

Then $\mathbf{M}^*$ is an $\mathscr{L}^*$-structure and for every assignment j we have

$$(\mathbf{M}^*, j) \models \mathsf{T} \quad \text{and} \quad (\mathbf{M}^*, j) \models \vartheta_R .$$

Therefore, we obtain

$$\mathbf{M}^* \models \mathsf{T}^* .$$

In order to prove (e), by GÖDEL'S COMPLETENESS THEOREM 5.5 it is enough to show that $\varphi^* \leftrightarrow \tilde{\varphi}$ holds in every model $\mathbf{M}^*$ of T^*. So, let $\mathbf{M}^*$ be

an arbitrary model of T^*. In particular, $\mathbf{M}^* \vDash \vartheta_R$. If φ^* does not contain R, then we are done. Otherwise, if φ^* is atomic, then $\varphi^* \equiv Rt_1 \cdots t_n$ for some $\mathscr{L}$-terms $t_1, \ldots, t_n$. Since $\mathbf{M}^* \vDash \vartheta_R$, we get:

$$\mathbf{M}^* \vDash Rt_1 \cdots t_n \leftrightarrow \psi'_R(t_1, \ldots, t_n)$$

This shows $\mathbf{M}^* \vDash \varphi^* \leftrightarrow \tilde{\varphi}$ for atomic formulae, and by (b)–(d) we get the result for arbitrary formulae.

For (f), we first extend an arbitrary model $\mathbf{M} \vDash \mathsf{T}$ to a model $\mathbf{M}^* \vDash \mathsf{T}^*$. By (e), for each $\mathscr{L}^*$-formula φ^* we have:

$$\mathbf{M}^* \vDash \varphi^* \quad \Lsh\!\!\Rsh \quad \mathbf{M}^* \vDash \tilde{\varphi}$$

Now, if $\mathsf{T}^* \vdash \varphi^*$, then $\mathbf{M}^* \vDash \varphi^*$, which implies that $\mathbf{M}^* \vDash \tilde{\varphi}$. Since $\tilde{\varphi}$ is an $\mathscr{L}$-formula, we get $\mathbf{M} \vDash \tilde{\varphi}$, and since the model $\mathbf{M}$ of T was arbitrary, by Gödel's Completeness Theorem 5.5 we get $\mathsf{T} \vdash \tilde{\varphi}$. ⊣

Defining new Function Symbols

If we define new functions, the situation is slightly more subtle. However, there is also an algorithm which transforms sentences σ^* in the extended language into equivalent sentences $\tilde{\sigma}$ in the original language:

Theorem 6.1. *Let $\mathscr{L}$ be a countable signature, let f be an n-ary function symbol which does not belong to $\mathscr{L}$, let $\mathscr{L}^* := \mathscr{L} \cup \{f\}$ and let T be a consistent $\mathscr{L}$-theory. Furthermore, let $\psi_f(v_1, \ldots, v_n, y)$ be an $\mathscr{L}$-formula with $\mathrm{free}(\psi_f) = \{v_1, \ldots, v_n, y\}$ such that*

$$\mathsf{T} \vdash \forall v_1 \cdots \forall v_n \exists! y \, \psi_f(v_1, \ldots, v_n, y)\,.$$

Finally, let

$$\vartheta_f \equiv \forall v_1 \cdots \forall v_n \forall y \big(f v_1 \cdots v_n = y \leftrightarrow \psi_f(v_1, \ldots, v_n, y) \big)$$

and let $\mathsf{T}^* := \mathsf{T} + \vartheta_f$. Then there exists an effective algorithm which transforms each $\mathscr{L}^*$-formula φ^* into an $\mathscr{L}$-formula $\tilde{\varphi}$ such that:

(a) *If f does not appear in φ^*, then $\tilde{\varphi} \equiv \varphi^*$.*

(b) $\widetilde{\neg\varphi} \equiv \neg\tilde{\varphi}$ *(for $\varphi^* \equiv \neg\varphi$)*

(c) $\widetilde{\varphi_1 \wedge \varphi_2} \equiv \tilde{\varphi}_1 \wedge \tilde{\varphi}_2$ *(for $\varphi^* \equiv \varphi_1 \wedge \varphi_2$)*

(d) $\widetilde{\exists \nu \varphi} \equiv \exists \nu \tilde{\varphi}$ *(for $\varphi^* \equiv \exists \nu \varphi$)*

(e) $\mathsf{T}^* \vdash \varphi^* \leftrightarrow \tilde{\varphi}$

(f) *If $\mathsf{T}^* \vdash \varphi^*$, then $\mathsf{T} \vdash \tilde{\varphi}$.*

Proof. By an *elementary f-term* we mean an $\mathscr{L}^*$-term of the form $ft_1 \cdots t_n$, where $t_1, \ldots, t_n$ are $\mathscr{L}^*$-terms which do not contain the symbol f. We first prove the theorem for atomic $\mathscr{L}^*$-formulae φ^* (i.e., for formulae which are free of quantifiers and logical operators). Let $\varphi^*(f|w)$ be the result of replacing the leftmost occurence of an elementary f-term in φ^* by a new symbol w, which stands for a new variable. Then, the formula

$$\exists w\big(\psi_f(t_1, \ldots, t_n, w) \wedge \varphi^*(f|w)\big)$$

is called the *f-transform of φ^**. If φ^* does not contain f, then let φ^* be its own f-transform. Before we procceed, let us prove the following

Claim. $\mathsf{T}^* \vdash \exists w\big(\psi_f(t_1, \ldots, t_n, w) \wedge \varphi^*(f|w)\big) \leftrightarrow \varphi^*$

Proof of Claim. Let $\mathbf{M}^*$ be a model of T^* with domain A, let j be an arbitrary assignment which assigns an element of A to w, and let $\mathbf{M}^*_j := (\mathbf{M}^*, j)$ be the corresponding $\mathscr{L}^*$-interpretation. Assume that

$$\mathbf{M}^*_j \vDash \exists w\big(\psi_f(t_1, \ldots, t_n, w) \wedge \varphi^*(f|w)\big)\,.$$

Then, since $\mathsf{T}^* \vdash \forall v_1 \cdots \forall v_n \exists! y\, \psi_f(v_1, \ldots, v_n, y)$, there exists a unique $b \in A$ such that

$$\mathbf{M}^*_{j\,\frac{b}{w}} \vDash \psi_f(t_1, \ldots, t_n, w) \wedge \varphi^*(f|w)\,,$$

which is the same as saying that

$$\mathbf{M}^*_j \vDash \psi_f(t_1, \ldots, t_n, b) \wedge \varphi^*(f|b)\,.$$

Now, since $\mathbf{M}^*_j \vDash \vartheta_f$, b is the same object as $f^{\mathbf{M}^*_j} t_1^{\mathbf{M}^*_j} \cdots t_n^{\mathbf{M}^*_j}$. This implies

$$\mathbf{M}^*_j \vDash ft_1 \cdots t_n = b\,,$$

which shows that

$$\mathbf{M}^*_j \vDash \varphi^*\,.$$

For the reverse implication, assume that $\mathbf{M}^*_j \vDash \varphi^*$ and let b be the same object as $f^{\mathbf{M}^*_j} t_1^{\mathbf{M}^*_j} \cdots t_n^{\mathbf{M}^*_j}$. Then $\mathbf{M}^*_j \vDash \varphi^*(f|b)$ and, since $\mathbf{M}^*_j \vDash \vartheta_f$,

$$\mathbf{M}^*_j \vDash \psi_f(t_1, \ldots, t_n, w) \leftrightarrow ft_1 \cdots t_n = w\,.$$

In particular, we get

$$\mathbf{M}^*_{j\,\frac{b}{w}} \vDash \psi_f(t_1, \ldots, t_n, b) \leftrightarrow ft_1 \cdots t_n = b\,,$$

and because $f^{\mathbf{M}^*_j} t_1^{\mathbf{M}^*_j} \cdots t_n^{\mathbf{M}^*_j}$ is the same object as b, we get $\mathbf{M}^*_j \vDash \psi_f(t_1, \ldots, t_n, b)$. Since we already know $\mathbf{M}^*_j \vDash \varphi^*(f|b)$, we have:

$$\mathbf{M}_j^* \vDash \psi_f(t_1, \ldots, t_n, b) \wedge \varphi^*(f|b)$$

So, there exists a b in A such that

$$\mathbf{M}_{j\,\frac{b}{w}}^* \vDash \psi_f(t_1, \ldots, t_n, w) \wedge \varphi^*(f|w)\,,$$

which is the same as saying that

$$\mathbf{M}_j^* \vDash \exists w\big(\psi_f(t_1, \ldots, t_n, w) \wedge \varphi^*(f|w)\big)\,.$$

Since the model $\mathbf{M}^*$ of T^* was arbitrary, by Gödel's Completeness Theorem 5.5 we get $\mathsf{T}^* \vdash \exists w\big(\psi_f(t_1, \ldots, t_n, w) \wedge \varphi^*(f|w)\big) \leftrightarrow \varphi^*$. $\dashv_{\text{Claim}}$

Since the f-transform $\exists w\big(\psi_f(t_1, \ldots, t_n, w) \wedge \varphi^*(f|w)\big)$ of φ^* contains one f less than φ^*, if we take successive f-transforms (always introducing new variables), we obtain eventually an atomic $\mathscr{L}$-formula $\tilde{\varphi}$ (i.e., a formula which does not contain f) such that $\mathsf{T}^* \vdash \varphi^* \leftrightarrow \tilde{\varphi}$. We call $\tilde{\varphi}$ the *f-less transform of φ^**.

In order to get f-less transforms of non-atomic $\mathscr{L}^*$-formulae φ^*, we just extend the definition by letting $\widetilde{\neg\varphi}$ be $\neg\tilde{\varphi}$, $\widetilde{\wedge\varphi_1\varphi_2}$ be $\wedge\tilde{\varphi}_1\tilde{\varphi}_2$, and $\widetilde{\exists\nu\varphi}$ be $\exists\nu\tilde{\varphi}$; properties (a)–(e) are then obvious.

It remains to prove property (f). For this, let $\mathbf{M}$ be an abitrary model of T with domain A. Then, since $\mathsf{T} \vdash \forall v_1 \cdots \forall v_n \exists! y\, \psi_f(v_1, \ldots, v_n, y)$, for all $a_1, \ldots, a_n$ in A there exists a unique b in A such that

$$\mathbf{M} \vDash \psi_f(a_1, \ldots, a_n, b),$$

and we define the n-ary function f^* on A by stipulating:

$$f^*(a_1, \ldots, a_n) := b$$

With this definition, we can extend the $\mathscr{L}$-structure $\mathbf{M}$ to an $\mathscr{L}^*$-structure $\mathbf{M}^*$, where we still have $\mathbf{M}^* \vDash \mathsf{T}$. With the definition of f^*, we additionally get $\mathbf{M}^* \vDash \vartheta_f$, which implies $\mathbf{M}^* \vDash \mathsf{T}^*$. If we have $\mathsf{T}^* \vdash \varphi^*$ for some $\mathscr{L}^*$-formula φ^*, then there exists an $\mathscr{L}$-formula $\tilde{\varphi}$ such that $\mathsf{T}^* \vdash \varphi^* \leftrightarrow \tilde{\varphi}$, i.e., $\mathsf{T}^* \vdash \tilde{\varphi}$. Since $\mathsf{T}^* \vdash \tilde{\varphi}$ implies $\mathbf{M}^* \vDash \tilde{\varphi}$, and because $\tilde{\varphi}$ is an $\mathscr{L}$-formula, we have $\mathbf{M} \vDash \tilde{\varphi}$. Now, since the model $\mathbf{M}$ of T was arbitrary, by Gödel's Completeness Theorem 5.5 we get $\mathsf{T} \vdash \tilde{\varphi}$. $\dashv$

Defining new Constant Symbols

Constant symbols can be handled like 0-ary function symbols:

Fact 6.2. *Let $\mathscr{L}$ be a countable signature, let c be a constant symbol which does not belong to $\mathscr{L}$, let $\mathscr{L}^* := \mathscr{L} \cup \{c\}$ and let T be a consistent $\mathscr{L}$-theory.*

Furthermore, let $\psi_c(y)$ *be an* $\mathscr{L}$*-formula with* $\mathrm{free}(\psi_c) = \{y\}$ *such that* $\mathsf{T} \vdash \exists! y\, \psi_c(y)$*. Finally, let*

$$\vartheta_c \equiv \forall y\big(c = y \leftrightarrow \psi_c(y)\big)$$

and let $\mathsf{T}^* := \mathsf{T} + \vartheta_c$*. Then there exists an effective algorithm which transforms each* $\mathscr{L}^*$*-formula* φ^* *into an* $\mathscr{L}$*-formula* $\tilde{\varphi}$ *such that:*

(a) *If* f *does not appear in* φ^**, then* $\tilde{\varphi} \equiv \varphi^*$*.*

(b) $\widetilde{\neg\varphi} \equiv \neg\tilde{\varphi}$ *(for* $\varphi^* \equiv \neg\varphi$*)*

(c) $\widetilde{\varphi_1 \wedge \varphi_2} \equiv \tilde{\varphi}_1 \wedge \tilde{\varphi}_2$ *(for* $\varphi^* \equiv \varphi_1 \wedge \varphi_2$*)*

(d) $\widetilde{\exists \nu \varphi} \equiv \exists \nu \tilde{\varphi}$ *(for* $\varphi^* \equiv \exists \nu \varphi$*)*

(e) $\mathsf{T}^* \vdash \varphi^* \leftrightarrow \tilde{\varphi}$

(f) *If* $\mathsf{T}^* \vdash \varphi^*$*, then* $\mathsf{T} \vdash \tilde{\varphi}$*.*

Proof. The algorithm is constructed in exactly the same way as in the proof of Theorem 6.1. ⊣

Notes

In the proof of Theorem 6.1, we essentially followed the proof of Proposition 2.28 of Mendelson [35].

Exercises

6.0 (a) Write the axioms of Group Theory in the language $\mathscr{L}_{\mathsf{GT}^*} = \{\circ\}$, where $\circ$ is a binary function symbol.

(b) Extend the language $\mathscr{L}_{\mathsf{GT}^*}$ with the constant symbol e for the neutral element.

(c) Extend the language $\mathscr{L}_{\mathsf{GT}^*} \cup \{\mathsf{e}\}$ with the unary function symbol $\mathrm{inv}(\,\cdot\,)$, where $\mathrm{inv}(x)$ is the inverse of x with respect to "$\circ$".

6.1 Show that in a signature $\mathscr{L}$, constant symbols and function symbols are dispensable (i.e., we only need relation symbols as non-logical symbols).

Hint: Notice that n-ary function symbols can be replaced by $(n+1)$-ary relation symbols, and that constant symbols can be replaced by unary relation symbols.

6.2 (a) Write the axioms of Group Theory in the language $\mathscr{L} = \{R\}$, where R is a ternary relation symbol.

(b) Extend the language $\mathscr{L}$ with the unary relation symbol R_{e} for the neutral element.

(c) Extend the language $\mathscr{L} \cup \{R_{\mathsf{e}}\}$ with the binary relation symbol R_{inv}, where $R_{\mathrm{inv}}(x, y)$ holds if and only if y is the inverse of x.

Hint: Use Exercises 6.1 and 6.0.

6.3 Let $\mathscr{L} = \{+, 0\}$ and $\mathscr{L}^* = \mathscr{L} \cup \{<\}$. Let $\mathsf{T}^* = \mathbf{Th}(\mathbb{Z}, +, 0, <)$. Show that there is no $\mathscr{L}$-formula $\psi_<$ such that and $\mathsf{T}^* \vdash \forall x \forall y (x < y \leftrightarrow \psi_<(x, y))$.

Part III
Gödel's Incompleteness Theorems

On the syntactical level, an $\mathscr{L}$-theory T is complete if for every $\mathscr{L}$-sentence σ, either $\mathsf{T} \vdash \sigma$ or $\mathsf{T} \vdash \neg\sigma$. On the semantical level, a consistent $\mathscr{L}$-theory is T complete if any two models of T are elementary equivalent.
In this part of the book we shall first provide a few models of Peano Arithmetic PA, where we assume that PA is consistent. Then, we shall prove Gödel's First Incompleteness Theorem, which states that Peano Arithmetic PA is not complete, i.e., there is a $\mathscr{L}_{\mathsf{PA}}$-sentence σ, such that neither $\mathsf{PA} \vdash \sigma$ nor $\mathsf{PA} \vdash \neg\sigma$. In a second step we shall prove Gödel's Second Incompleteness Theorem, In a second step we shall prove Gödel's Second Incompleteness Theorem, which states that no theory which is at least as strong as Peano Arithmetic can prove its own consistency.

Chapter 7
Countable Models of Peano Arithmetic

By Gödel's Completeness Theorem 5.5 we know that every consistent theory T has a model. Later we will see that if T has an infinite model, then it also has arbitrarily large models.

In this chapter, we provide different countable models of PA and investigate their structure. First, we construct the so-called *standard model*, and then we extend this model to countable *non-standard models*, where by "countable" we mean that the elements of the domain of the model can be listed in a potentially infinite list.

The Standard Model

For the sake of completeness, let us first recall the language and the seven axioms of Peano Arithmetic PA. The language of PA is $\mathscr{L}_{\mathsf{PA}} = \{\mathsf{0}, \mathsf{s}, +, \cdot\}$, where $\mathsf{0}$ is a constant symbol, s is a unary function symbol, and $+$ and $\cdot$ are binary function symbols.

PA_0: $\neg\exists x(\mathsf{s}x = \mathsf{0})$

PA_1: $\forall x\forall y(\mathsf{s}x = \mathsf{s}y \to x = y)$

PA_2: $\forall x(x + \mathsf{0} = x)$

PA_3: $\forall x\forall y(x + \mathsf{s}y = \mathsf{s}(x + y))$

PA_4: $\forall x(x \cdot \mathsf{0} = \mathsf{0})$

PA_5: $\forall x\forall y(x \cdot \mathsf{s}y = (x \cdot y) + x)$

If φ is any $\mathscr{L}_{\mathsf{PA}}$-formula with $x \in \mathrm{free}(\varphi)$, then:

PA_6: $\big(\varphi(\mathsf{0}) \wedge \forall x(\varphi(x) \to \varphi(\mathsf{s}(x)))\big) \to \forall x\varphi(x)$

The domain $\mathbb{N}$ of our standard model consists of the elements in the list of natural numbers as introduced in Chapter 0. So, each natural number in the

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_8

set $\mathbb{N}$ is either $\mathbf{0}$ or of the form $\mathbf{s} \cdots \mathbf{s0}$ for some FINITE string $\mathbf{s} \cdots \mathbf{s}$. Notice the difference between $\mathtt{s}$ (which is an unary function symbol) and $\mathbf{s}$ (which is a symbol which we use to build the elements of the set $\mathbb{N}$, i.e., the objects in the domain of our standard model of Peano Arithmetic). In order to write this more formally, we extend the signature $\mathscr{L}_{\mathsf{PA}}$ by the unary relation symbol $\mathtt{N}$ and add the following statement as a kind of meta-axiom to PA:

$$\Phi \equiv \forall x \left(\left\{ \mathtt{N}(0), \forall \mathtt{z} \big(\mathtt{N}(\mathtt{z}) \to \mathtt{N}(\mathtt{sz}) \big) \right\} \vdash \mathtt{N}(\mathtt{x}) \right)$$

Notice that this statement is *not* a statement in first-order logic since it involves the symbol $\vdash$, which implicitly incorporates the metamathematical notion of FINITENESS. However, the statement Φ makes sure that every model of $\mathsf{PA} + \Phi$ is isomorphic to the standard model.

Now, we are going to define the standard model of PA with domain $\mathbb{N}$. For this, we first have to define first an $\mathscr{L}_{\mathsf{PA}}$-structure $\mathbb{N}$. If σ and τ are both (possibly empty) finite strings of the form $\mathbf{s} \cdots \mathbf{s}$, then we can interpret the non-logical symbols in $\mathscr{L}_{\mathsf{PA}}$ as follows:

$$0^{\mathbb{N}} := \mathbf{0}$$

$$\begin{aligned} \mathtt{s}^{\mathbb{N}} : \ & \mathbb{N} \to \mathbb{N} \\ & \sigma\mathbf{0} \mapsto \mathbf{s}\sigma\mathbf{0} \end{aligned}$$

$$\begin{aligned} +^{\mathbb{N}} : \quad & \mathbb{N} \times \mathbb{N} \ \to \ \mathbb{N} \\ & \langle \sigma\mathbf{0}, \tau\mathbf{0} \rangle \mapsto \sigma\tau\mathbf{0} \end{aligned}$$

$$\begin{aligned} \cdot^{\mathbb{N}} : \quad & \mathbb{N} \times \mathbb{N} \ \to \mathbb{N} \\ & \langle \sigma\mathbf{0}, \tau\mathbf{0} \rangle \mapsto \sigma\sigma \ \cdots \ \sigma\mathbf{0} \\ & \qquad\qquad\quad \uparrow\uparrow \ \cdots \ \uparrow \\ & \qquad\qquad \underbrace{\mathbf{s}\,\mathbf{s} \ \cdots \ \mathbf{s}}_{\tau} \end{aligned}$$

Note that if either σ or τ is the empty string, then $\sigma\mathbf{0} \cdot^{\mathbb{N}} \tau\mathbf{0}$ is $\mathbf{0}$. The main feature of the $\mathscr{L}_{\mathsf{PA}}$-structure $\mathbb{N}$ is that every element of $\mathbb{N}$ corresponds to a certain $\mathscr{L}_{\mathsf{PA}}$-term. In order to prove this, we introduce the following notion: To each FINITE string $\sigma \equiv \mathbf{s} \cdots \mathbf{s}$ we assign a FINITE string $\underline{\sigma} \equiv \mathtt{s} \cdots \mathtt{s}$ such that $\underline{\sigma}$ is obtained from σ by replacing each occurrence of $\mathbf{s}$ by $\mathtt{s}$. As a consequence of this definition, we get the following

FACT 7.0. *For all* FINITE *strings σ and τ of the form $\mathbf{s} \cdots \mathbf{s}$, we have:*

(a) *If σ is not the empty string, then $\mathsf{PA} \vdash \underline{\sigma}0 \neq 0$.*

(b) $\mathsf{PA} \vdash \underline{\sigma}0 = \underline{\tau}0 \iff \sigma\mathbf{0} \equiv \tau\mathbf{0}$.

Proof. (a) follows from PA_0, and (b) follows from PA_1 and L_{14}. ⊣

LEMMA 7.1. *Every element of* $\mathbb{N}$ *corresponds to a unique* FINITE *application of the function* $\mathtt{s}$ *to* 0, *or in other words, every element of* $\mathbb{N}$ *is equal to a unique* FINITE *application of the function* $\mathtt{s}^{\mathbb{N}}$ *to* $0^{\mathbb{N}}$. *More formally, for every element* $\sigma\mathbf{0}$ *of* $\mathbb{N}$ *there is a unique* $\mathscr{L}_{\mathsf{PA}}$*-term* $\underline{\sigma}0$ *such that*

$$(\underline{\sigma}0)^{\mathbb{N}} \quad \text{IS THE SAME OBJECT AS} \quad \sigma\mathbf{0}\,,$$

or equivalently,

$$(\underline{\sigma}0)^{\mathbb{N}} \equiv \sigma\mathbf{0}\,.$$

Proof. By definition of $\mathtt{s}^{\mathbb{N}}$, for every FINITE string $\tau \equiv \mathbf{s}\cdots\mathbf{s}$ we get that $\mathtt{s}^{\mathbb{N}}(\tau\mathbf{0})$ is the same element of $\mathbb{N}$ as $\mathbf{s}\tau\mathbf{0}$, and after applying this fact FINITELY many times we get:

$$\overbrace{\underbrace{\mathtt{s}^{\mathbb{N}}\ \mathtt{s}^{\mathbb{N}}\ \cdots\ \mathtt{s}^{\mathbb{N}}\ 0^{\mathbb{N}}}_{\sigma\mathbf{0}}}^{(\underline{\sigma}0)^{\mathbb{N}}}$$

The uniqueness of $\underline{\sigma}0$ follows from FACT 7.0. ⊣

Now, we are ready to prove that the $\mathscr{L}_{\mathsf{PA}}$-structure $\mathbb{N}$, which is called the **standard model** of Peano Arithmetic, is indeed a model of PA.

THEOREM 7.2. $\mathbb{N} \vDash \mathsf{PA}$.

Proof. By definition of $\mathtt{s}^{\mathbb{N}}$ we get $\mathbb{N} \vDash \mathsf{PA}_0$ and by FACT 7.0 we also have $\mathbb{N} \vDash \mathsf{PA}_1$. Further, by definition of $+^{\mathbb{N}}$ and $\cdot^{\mathbb{N}}$ we get $\mathbb{N} \vDash \mathsf{PA}_2$ and $\mathbb{N} \vDash \mathsf{PA}_4$ respectively. For PA_3, let σ and τ be (possibly empty) finite strings of the form $\mathbf{s}\cdots\mathbf{s}$. Then

$$\sigma\mathbf{0} +^{\mathbb{N}} \mathtt{s}^{\mathbb{N}}\tau\mathbf{0} \equiv \sigma\mathbf{s}\tau\mathbf{0} \equiv \mathbf{s}\sigma\tau\mathbf{0} \equiv \mathtt{s}^{\mathbb{N}}(\sigma\mathbf{0} +^{\mathbb{N}} \tau\mathbf{0}).$$

Similarly, we can show $\mathbb{N} \vDash \mathsf{PA}_5$ (see EXERCISE 7.0). In order to show that $\mathbb{N} \vDash \mathsf{PA}_6$, let $\varphi(x)$ be an $\mathscr{L}_{\mathsf{PA}}$-formula and let us assume that

$$\mathbb{N} \vDash \varphi(0) \wedge \forall x\big(\varphi(x) \to \varphi(\mathtt{s}x)\big)\,. \qquad (*)$$

We have to show that $\mathbb{N} \vDash \forall x\varphi(x)$. By definition of models we get that $\varphi(\mathbf{0})$ holds in $\mathbb{N}$ and for all $n \in \mathbb{N}$: If $\varphi(n)$ holds in $\mathbb{N}$, then also $\varphi(\mathtt{s}^{\mathbb{N}}n)$ holds in $\mathbb{N}$. Let $\sigma\mathbf{0}$ be an arbitrary element of $\mathbb{N}$. Since σ is a FINITE string,

by $(*)$, the logical axiom L_{10}, and by applying FINITELY many times Modus Ponens, we get $\mathbb{N} \vDash \varphi(\sigma\mathbf{0})$. Hence, since $\sigma\mathbf{0}$ was arbitrary, $\varphi(n)$ holds in $\mathbb{N}$ for every string $n \in \mathbb{N}$, and therefore, $\mathbb{N} \vDash \forall x \varphi(x)$. ⊣

As a matter of fact, we would like to mention that from a metamathematical point of view, every model of PA must contain an isomorphic copy of the standard model $\mathbb{N}$. Therefore, it would also make sense to call $\mathbb{N}$ the **minimal model** of Peano Arithmetic.

One might be tempted to think that $\mathbb{N}$ is essentially the only model of PA, but this is not the case, as we shall now see.

Countable Non-Standard Models

The previous section shows that every natural number in the standard model $\mathbb{N}$ corresponds to a unique $\mathscr{L}_{\mathsf{PA}}$-term; more precisely, every element $\sigma\mathbf{0}$ of $\mathbb{N}$ is the same object as the term $\underline{\sigma}0$. In order to simplify notations, we will from now on use variables such as $n, m, \ldots$ to denote elements of $\mathbb{N}$ and $\underline{n}, \underline{m}, \ldots$ their counterpart in the formal language $\mathscr{L}_{\mathsf{PA}}$, i.e., if n stands for $\sigma\mathbf{0}$, then $\underline{n}$ denotes $\underline{\sigma}0$.

Since every model $\mathbf{M}$ of PA contains $\underline{n}^{\mathbf{M}}$, the **standard natural numbers**, for every $n \in \mathbb{N}$, it is clear that $\mathbf{M}$ contains a copy of the standard model. However, $\mathbf{M}$ can also have **non-standard natural numbers**, i.e., elements which are not interpretations of terms of the form $\underline{n}$. In the following, we present the simplest way to construct such non-standard models.

Let $\mathscr{L}_{\mathsf{PA}^+}$ be the language $\mathscr{L}_{\mathsf{PA}}$ augmented by an additional constant symbol c, which is different from $\mathsf{0}$. Note that by setting

$$x < y :\Longleftrightarrow \exists r(x + \mathsf{s}r = y)$$

one can introduce an ordering in PA, which in the standard model corresponds to the usual ordering of natural numbers (for further details see Chapters 8 and 9). Let PA^+ be the theory whose axioms are PA_0–PA_6 together with the axioms

$$\begin{aligned} &\mathsf{c} > \mathsf{0} \\ &\mathsf{c} > \mathsf{s0} \\ &\mathsf{c} > \mathsf{ss0} \\ &\mathsf{c} > \mathsf{sss0} \\ &\quad\vdots \end{aligned}$$

Hence, PA^+ is $\mathsf{PA} \cup \{\mathsf{c} > \underline{n} : n \in \mathbb{N}\}$.

LEMMA 7.3. $\mathrm{Con}(\mathsf{PA}^+)$, *i.e., the theory* PA^+ *is consistent.*

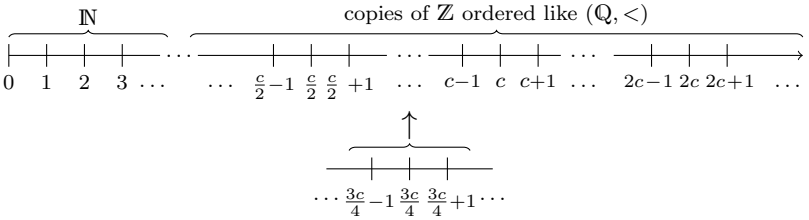
Proof. By the COMPACTNESS THEOREM it suffices to prove that every FINITE subset of PA^+ is consistent. Let T be a FINITE subset of PA^+. Now let $n \in \mathbb{N}$ be maximal such that the formula $\mathsf{c} > \underline{n}$ belongs to T. Notice that such n exists, since T is FINITE. Then we can define a model $\mathbf{M}$ of T with domain $\mathbb{N}$ by interpreting the constant and function symbols by $\mathsf{0}^{\mathbf{M}} \equiv \mathbf{0}$, $\mathsf{s}^{\mathbf{M}} \equiv \mathsf{s}$, $+^{\mathbf{M}} \equiv +^{\mathbb{N}}$, $\cdot^{\mathbf{M}} \equiv \cdot^{\mathbb{N}}$ and $\mathsf{c}^{\mathbf{M}} \equiv \mathsf{s}n$. Since $\mathbb{N} \models \mathsf{PA}$, we get that $\mathbf{M} \models \mathsf{PA}$ and by construction $\mathbf{M} \models \mathsf{c} > \underline{m}$ for every $m \leq n$, and hence $\mathbf{M} \models \mathsf{T}$. ⊣

Now, since $\mathscr{L}_{\mathsf{PA}^+}$ is a countable signature, by THEOREM 5.6 it follows that PA^+ has a countable model $\mathbf{M}$ which is also a **non-standard model** of PA, i.e., a model which is not isomorphic to the standard model $\mathbb{N}$. What does the order structure of $\mathbf{M}$ look like?

Note that c has a successor $\mathsf{s}c = c + \mathsf{1}$, and $c + \mathsf{1}$ in turn has a successor, and so on. Furthermore, since

$$\mathsf{PA} \vdash \forall x \big(x = \mathsf{0} \vee \exists y (x = \mathsf{s}y)\big)$$

(see LEMMA 8.3), c also has a predecessor, i.e., there exists $c - \mathsf{1}$ with the property that $\mathsf{s}(c - \mathsf{1}) = c$, and the same argument yields that c in fact has infinitely many predecessors, which are all non-standard. Hence, the order structure of c and its predecessors and successors corresponds to $(\mathbb{Z}, <)$, so there are infinitely many such $\mathbb{Z}$-chains. Moreover, each multiple of c yields a further copy of a $\mathbb{Z}$-chain. Now, one can easily prove in PA that every number is either even or odd (see EXERCISE 8.1), and hence there is d such that $\mathsf{2}d = c$ or $\mathsf{2}d = c + \mathsf{1}$. We denote d by $\frac{c}{2}$. This shows that between the copy of the standard model and the $\mathbb{Z}$-chain given by c, there is a further $\mathbb{Z}$-chain given by $\frac{c}{2}$ and its predecessors and successors. In fact, the $\mathbb{Z}$-chains are ordered like $(\mathbb{Q}, <)$ (see EXERCISE 7.6).



Note that the proof of LEMMA 7.3 implies that there are non-standard models of PA which are elementarily equivalent to $\mathbb{N}$. To see this, let $\mathbf{Th}(\mathbb{N})$ denote the theory of all $\mathscr{L}_{\mathsf{PA}}$-sentences which are true in $\mathbb{N}$. Then one could simply replace PA by $\mathbf{Th}(\mathbb{N})$ in LEMMA 7.3 and thus obtain a model of $\mathbf{Th}(\mathbb{N})$ augmented by all formulae of the form $\mathsf{c} > \underline{n}$ for $n \in \mathbb{N}$. By construction, this model is elementarily equivalent to $\mathbb{N}$. For a more general result see EXERCISE 7.3.

Notes

An early attempt at formalising arithmetic was given by Grassmann [20] in 1861, who defined addition and multiplication and proved elementary results such as the associative and commutative laws using induction. Dedekind [6] also identified induction as a key principle in 1888, as well as the first two axioms of Peano Arithmetic; however, he introduced them as a definition rather than as axioms. Peano [42] presented his five axioms in 1889, where he only introduced zero and the successor function axiomatically, and the induction axiom is given in second-order logic in the following form: Every set of natural numbers which contains 0 and is closed under the successor function is the set of all natural numbers. The version of Peano's Axioms formalised in first-order logic — where the induction axiom is replaced by an axiom schema, and the axioms defining addition and multipliation are included — goes back to the advent of first-order logic in the 1920's. The first explicit construction of a non-standard model of arithmetic was given by Skolem in [52]. For further reading on non-standard models consult [29].

Exercises

7.0 Verify that $\mathbb{N} \vDash \mathsf{PA}_5$.

7.1 Prove that PA_0 and PA_1 are independent of the other axioms of PA.

7.2 Show that there is a non-standard model $\mathbf{M}$ of PA with domain M, such that there is a non-zero $a \in M$ divisible by every standard prime number.

Hint: Notice that divisibility can be formalised in PA by $x \mid y :\Longleftrightarrow \exists r(rx = y)$.

7.3 Show that there are uncountably many countable models of PA which are all elementarily equivalent (with respect to the signature $\mathscr{L}_{\mathsf{PA}}$) and pairwise non-isomorphic.

Hint: Let $\mathbb{P} \subseteq \mathbb{N}$ be the set of prime numbers and let c be a constant symbol which is different from $\mathsf{0}$. For any distinct prime numbers p and q, let $\varphi_{p,q}$ be the formula

$$p \mid \mathsf{c} \wedge q \nmid \mathsf{c}\,.$$

For every subset $S \subseteq \mathbb{P}$, let Φ_S be the collection of all formulae $\varphi_{p,q}$ such that $p \in S$ and $q \notin S$. For each set $S \subseteq \mathbb{P}$ construct a model for $\mathsf{PA} + \Phi_S$ and use that, by Cantor's Theorem 13.8, the power-set of $\mathbb{P}$ is uncountable.

7.4 Prove the following so-called *Overspill Principle*: If $\mathbf{M}$ is a non-standard model of PA with domain M and φ is a formula with one free variable then

$$\mathbf{M} \vDash \varphi(\underline{n}) \quad \text{for all } n \in \mathbb{N}$$

implies that there is a non-standard element $a \in M$ such that

$$\mathbf{M} \vDash \forall x(x < a \to \varphi(x)).$$

7.5 Show that it is not possible to introduce a relation standard(x) by a language extension of $\mathscr{L}_{\mathsf{PA}}$ such that for every model $\mathbf{M}$ of PA with domain M and for every $a \in M$ we have $\mathbf{M} \vDash \text{standard}(a)$ if and only if $a = \underline{n}^{\mathbf{M}}$ for some $n \in \mathbb{N}$.

7.6 Let $\mathbf{M}$ be a countable non-standard model of PA with domain M. For every non-standard element c of M, let

$$\mathbb{Z}_c :\equiv \big\{d \in M : \text{there exists } n \in \mathbb{N} \text{ such that } d + n = c \text{ or } c + n = d\big\},$$

and let $\mathbb{Z}_c \prec \mathbb{Z}_d$ if and only if $c + n < d$ for all $n \in \mathbb{N}$, where $<$ is the linear ordering in the model $\mathbf{M}$.

Show that the set $\{\mathbb{Z}_c : c \in M \text{ is non-standard}\}$ together with the binary relation $\prec$ is a dense linearly ordered set and use Exercise 3.4 to conclude that the order structure of $\mathbf{M}$ corresponds to the disjoint union of $\mathbb{N}$ and $\mathbb{Q} \times \mathbb{Z}$.

Chapter 8
Arithmetic in Peano Arithmetic

In this chapter, we take a closer look at Peano Arithmetic (PA) which we have defined in Chapter 1. In particular, we prove within PA some basic arithmetical results, starting with the commutativity and associativity of addition and multiplication, culminating in some results about coprimality. This paves the way for the coding of finite sequences of numbers, which will be covered in the next chapter. Furthermore, we introduce some alternative formulations of the Induction Schema PA_6.

Addition & Multiplication

In this section, we verify the basic computation rules of PA involving addition and multiplication. Since the complete proofs are very long and tedious, we will only show the commutativity of $+$ in an elaborate way. Subsequently, we will use semi-formal proofs as described in Chapter 1 which include enough details to allow the reader to reconstruct a corresponding formal proof.

Lemma 8.0. $\mathsf{PA} \vdash \forall x \forall y (x + y = y + x)$

Proof. We proceed by induction on x. Thus, we have to show

(a) $\mathsf{PA} \vdash \forall y (\mathsf{0} + y = y + \mathsf{0})$, and

(b) $\mathsf{PA} \vdash \forall y (x + y = y + x) \to \forall y (\mathsf{s}x + y = y + \mathsf{s}x)$.

For (a), we first prove

$$\mathsf{PA} \vdash \forall y (\mathsf{0} + y = y)$$

by induction on y. The base case $\mathsf{0} + \mathsf{0} = \mathsf{0}$ is clearly an instance of PA_2, and for the induction step, we assume $\mathsf{0} + y = y$ for some y. Then $\mathsf{0} + \mathsf{s}y = \mathsf{s}(\mathsf{0} + y)$ by PA_3 and $\mathsf{s}(\mathsf{0} + y) = \mathsf{s}y$ by assumption. In order to keep the notation short,

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_9

we just write $0+\mathsf{s}y = \mathsf{s}(0+y) = \mathsf{s}y$ instead of $0+\mathsf{s}y = \mathsf{s}(0+y) \wedge \mathsf{s}(0+y) = \mathsf{s}y$. So, by PA_6 we obtain $\forall y(0+y=y)$, and since by PA_2 we have $\forall y(y+0=y)$, by symmetry and transitivity of $=$ we have $\forall y(0+y=y+0)$.

As a prerequisite for (b) we need

$$\mathsf{PA} \vdash \forall y\big(\mathsf{s}x + y = \mathsf{s}(x+y)\big)$$

which is again verified by induction on y: If $y=0$, note that by PA_2 we have $\mathsf{s}x+0 = \mathsf{s}x = \mathsf{s}(x+0)$. For the induction step, assume $\mathsf{s}x+y = \mathsf{s}(x+y)$. Then, by PA_3, we have $\mathsf{s}x+\mathsf{s}y = \mathsf{s}(\mathsf{s}x+y) = \mathsf{s}(\mathsf{s}(x+y)) = \mathsf{s}(x+\mathsf{s}y)$.

Now, we are ready to prove (b): Assume that $x+y=y+x$ for some x and for all y. Then $\mathsf{s}x+y = \mathsf{s}(x+y) = \mathsf{s}(y+x) = y+\mathsf{s}x$ by our computation above and PA_3, which, by PA_6, shows (b). ⊣

In a similar manner, we can derive other basic calculation rules whose proofs are left as an exercise for the reader.

Lemma 8.1.

(a) $\mathsf{PA} \vdash \forall x\forall y\forall z\big((x+y)+z = x+(y+z)\big)$

(b) $\mathsf{PA} \vdash \forall x\forall y\forall z\big((x\cdot y)\cdot z = x\cdot(y\cdot z)\big)$

(c) $\mathsf{PA} \vdash \forall x\forall y(x\cdot y = y\cdot x)$

(d) $\mathsf{PA} \vdash \forall x\forall y\forall z\big(x\cdot(y+z) = (x\cdot y)+(x\cdot z)\big)$

From now on, we will make use of these rules without explicitly mentioning them anymore. The next lemma shows injectivity of left — and by commutativity also right — addition.

Lemma 8.2. $\mathsf{PA} \vdash \forall x\forall y\forall z(x+y=x+z \to y=z)$

Proof. The proof is by induction on x. The base case follows from the proof of Lemma 8.0. For the induction step, assume

$$\forall y\forall z(x+y=x+z \to y=z)$$

and let $\mathsf{s}x+y = \mathsf{s}x+z$. Then $\mathsf{s}(x+y) = \mathsf{s}x+y = \mathsf{s}x+z = \mathsf{s}(x+z)$, where the first and the third equality again follow from Lemma 8.0 and PA_3. Then by PA_2 we obtain $x+y=x+z$ and in particular $y=z$. ⊣

The next result is crucial, because — as we will see in Chapter 10 — it is the only application of PA_6 which is indispensable for the proof of the First Incompleteness Theorem 10.5.

Lemma 8.3. $\mathsf{PA} \vdash \forall x\big(x=0 \vee \exists y(x=\mathsf{s}y)\big)$

Proof. We proceed by induction on x. The base case is trivial and the induction step follows from the fact that x witnesses $\exists y(\mathsf{s}x=\mathsf{s}y)$. ⊣

From now on, we will use the convention that $\cdot$ binds stronger than $+$ and omit the multiplication sign, e.g., the term $xy + z$ stands for $(x \cdot y) + z$. Furthermore, by associativity of $+$ and $\cdot$ we may omit parentheses whenever we have pure products of pure sums of terms.

In order to keep the notation short, for $\mathscr{L}_{\mathsf{PA}}$-formulae φ we define

$$\forall x \neq \mathsf{0}\,\big(\varphi(x)\big) :\Longleftrightarrow \forall x\big(x \neq \mathsf{0} \to \varphi(x)\big).$$

The next result shows a property of multiplication which is similar to the one given in Lemma 8.2 for addition.

Lemma 8.4.

(a) $\mathsf{PA} \vdash \forall x \forall y\big(xy = \mathsf{0} \leftrightarrow (x = \mathsf{0} \vee y = \mathsf{0})\big)$

(b) $\mathsf{PA} \vdash \forall x \neq \mathsf{0}\, \forall y \forall z(xy = xz \to y = z)$

Proof. For (a), let $xy = \mathsf{0}$ and suppose towards a contradiction that $x, y \neq \mathsf{0}$. Then by Lemma 8.3 there are x', y' such that $x = \mathsf{s}x'$ and $y = \mathsf{s}y'$. By PA_5 and PA_3, we obtain

$$\mathsf{0} = xy = \mathsf{s}x' \cdot \mathsf{s}y' = \mathsf{s}x' \cdot y' + \mathsf{s}x' = \mathsf{s}(\mathsf{s}x' \cdot y' + x')\,,$$

which contradicts PA_0.

For (b), suppose that $x \neq \mathsf{0}$. We proceed by induction on y. If $y = \mathsf{0}$, then $xy = \mathsf{0}$. So, $xy = xz$ implies $xz = \mathsf{0}$ and by (a) we obtain $z = \mathsf{0}$ and consequently $y = z$. Now assume that

$$\forall z(xy = xz \to y = z).$$

Let z be arbitrary such that $x \cdot \mathsf{s}y = xz$. By (a), we can rule out the possibility that $z = \mathsf{0}$. Hence, by Lemma 8.3, there is a z' such that $z = \mathsf{s}z'$. Therefore, by PA_5,

$$xy + x = x \cdot \mathsf{s}y = xz = x \cdot \mathsf{s}z' = xz' + x\,.$$

Using Lemmata 8.0 and 8.2 we obtain that $xy = xz'$ and thus the induction hypothesis implies $y = z'$. Therefore, we finally get $\mathsf{s}y = \mathsf{s}z' = z$ as desired. ⊣

The Natural Ordering on Natural Numbers

In Chapter 6, we have seen how to extend languages by incorporating new symbols for relations, functions or constants. In this sense, we can now introduce the binary relations $\leq$ and $<$ in PA by stipulating

$$x \leq y :\Longleftrightarrow \exists r(x + r = y)\,,$$

$$x < y :\Longleftrightarrow \exists r(x + \mathsf{s}r = y)\,.$$

An alternative definition of $x < y$ is given by

$$x < y :\Longleftrightarrow \exists r \neq 0\,(x + r = y)\,.$$

Furthermore, we define

$$x \geq y :\Longleftrightarrow y \leq x$$
$$x > y :\Longleftrightarrow y < x.$$

Now, we define **bounded quantification** by stipulating

$$\exists x \lhd y\, \varphi(x) :\Longleftrightarrow \exists x\big(x \lhd y \wedge \varphi(x)\big),$$

$$\forall x \lhd y\, \varphi(x) :\Longleftrightarrow \forall x\big(x \lhd y \to \varphi(x)\big),$$

where $\lhd$ stands either for $<$ or for $\leq$. The next result shows some properties of $<$ and $\leq$.

Lemma 8.5.

(a) $\mathsf{PA} \vdash \forall x \forall y (x < \mathtt{s}y \leftrightarrow x \leq y)$

(b) $\mathsf{PA} \vdash \forall x \forall y (x < y \leftrightarrow \mathtt{s}x \leq y)$

Proof. We only consider (a) and leave (b) as an excercise. Fix x and y. Firstly, assume that $x < \mathtt{s}y$ and take $r \neq 0$ such that $x + r = \mathtt{s}y$. By Lemma 8.3 we find an r' such that $r = \mathtt{s}r'$. Then $\mathtt{s}(x + r') = x + \mathtt{s}r' = x + r = \mathtt{s}y$ by PA_3, and by PA_2 we obtain $x + r' = y$, which shows that $x \leq y$.

Conversely, let $x \leq y$ and take r such that $x + r = y$. Then $x + \mathtt{s}r = \mathtt{s}(x + r) = \mathtt{s}y$, which shows that $x < \mathtt{s}y$. ⊣

The next result implies that $\leq$ defines a total ordering on the natural numbers.

Lemma 8.6.

(a) $\mathsf{PA} \vdash \forall x (x \leq x)$

(b) $\mathsf{PA} \vdash \forall x \forall y (x \leq y \wedge y \leq x \to x = y)$

(c) $\mathsf{PA} \vdash \forall x \forall y \forall z (x \leq y \wedge y \leq z \to x \leq z)$

(d) $\mathsf{PA} \vdash \forall x \forall y (x < y \vee x = y \vee x > y)$

Proof. Condition (a) is a trivial consequence of PA_2.

For (b), assume that $x \leq y$ and $y \leq x$. Then there are r, s such that $x + r = y$ and $y + s = x$. We obtain that

$$y + (s + r) = (y + s) + r = x + r = y = y + 0.$$

By Lemma 8.2, this implies $s + r = \mathsf{0}$ and hence, by PA_0, $s = \mathsf{0} = r$, which shows that $x = y$.

For (c), let $x \leq y$ and $y \leq z$ and take witnesses r, s satisfying $x + r = y$ and $y + s = z$, respectively. Then $x + (r + s) = (x + r) + s = y + s = z$ and thus $x \leq z$.

We show (d) by induction on x. If $x = \mathsf{0}$, we can make a case distinction according to Lemma 8.3: If $y = \mathsf{0}$ then $x = y$ and otherwise $x < y$. For the induction step, fix y and assume that $x < y \vee x = y \vee x > y$. Now, we make a case distinction, where in the case of $x < y$, Lemma 8.5 implies that $\mathsf{s}x \leq y$ and thus either $\mathsf{s}x < y$ or $\mathsf{s}x = y$. Secondly, if $x = y$ then

$$\mathsf{s}x = \mathsf{s}y = \mathsf{s}(y + \mathsf{0}) = y + \mathsf{s0},$$

which shows that $\mathsf{s}x > y$. The case of $x > y$ is similar. ⊣

Finally, one can show that addition and multiplication with non-zero numbers preserve the natural odering (the proof is left as an excercise to the reader):

Lemma 8.7.

(a) $\mathsf{PA} \vdash \forall x \forall y \forall z \big(x \leq y \leftrightarrow (x + z \leq y + z)\big)$

(b) $\mathsf{PA} \vdash \forall x \forall y \forall z \neq \mathsf{0} \big(x \leq y \leftrightarrow (x \cdot z \leq y \cdot z)\big)$

Subtraction & Divisibility

With the help of the ordering that we have introduced in the previous section, we are ready to define a version of subtraction which rounds up to $\mathsf{0}$ in order to preserve non-negativity. For this, we first show the following

Lemma 8.8. $\mathsf{PA} \vdash \forall x \forall y \big(x \leq y \to \exists! r (x + r = y)\big)$

Proof. Assume that $x \leq y$. The existence of r follows directly from the definition of $\leq$, and the uniqueness of r is a consequence of Lemma 8.2. ⊣

Therefore, we can define within PA the binary function $-$, called **bounded subtraction**, by stipulating

$$x - y = z :\Longleftrightarrow (y \leq x \wedge y + z = x) \vee (x < y \wedge z = \mathsf{0})\,.$$

Observe that $\mathsf{PA} \vdash \forall x \forall y \leq x ((x - y) + y = x)$, from which we can easily derive computation rules for bounded subtraction such as

$$\mathsf{PA} \vdash \forall x \forall y \forall z \big(x(y - z) = xy - xz\big), \text{ or}$$

$$\mathsf{PA} \vdash \forall x \forall y \forall z \big(x \leq y \to (x - z \leq y - z)\big).$$

Let us now turn to divisibility, which can easily be formalised by stipulating

$$x \mid y :\Longleftrightarrow \exists r(rx = y).$$

If the binary divisibility relation $\mid$ holds for the ordered pair (x, y), then we say that x *divides* y. Without much effort, one can verify that the divisibility relation is reflexive, antisymmetric, and transitive. For this reason, we will omit the proof of the next result.

Lemma 8.9.

(a) $\mathsf{PA} \vdash \forall x(x \mid x)$
(b) $\mathsf{PA} \vdash \forall x \forall y(x \mid y \wedge y \mid x \to x = y)$
(c) $\mathsf{PA} \vdash \forall x \forall y \forall z(x \mid y \wedge y \mid z \to x \mid z)$

Also without much effort, we can prove the following

Lemma 8.10.

(a) $\mathsf{PA} \vdash \forall x \forall y \forall z(x \mid y \wedge x \mid z \to x \mid y \pm z)$, *where the symbol* $\pm$ *stands for either* $+$ *or* $-$.
(b) $\mathsf{PA} \vdash \forall x \forall y \forall z(x \mid y \to x \mid yz)$

Proof. For (a), assume that x divides y and z. Then there are r, s such that $y = rx$ and $z = sx$. Then $y \pm z = rx \pm sx = (r \pm s)x$, thus x divides $y \pm z$. Condition (b) is obvious. ⊣

In most textbooks, one defines two numbers to be *coprime* (or *relatively prime*), if they have no common divisor greater than 1. Nevertheless, for our purpose it is more convenient to use the following equivalent definition:

$$\text{coprime}(x, y) \;:\Longleftrightarrow\; x \neq 0 \wedge y \neq 0 \wedge \forall z\big(x \mid yz \to x \mid z\big)$$

Since we will be working with this somewhat unusual definition of relative primality, we first check that it is a symmetric relation.

Lemma 8.11. $\mathsf{PA} \vdash \forall x \forall y\big(\text{coprime}(x, y) \leftrightarrow \text{coprime}(y, x)\big)$

Proof. Assume $\text{coprime}(x, y)$. We have to show that for every z we have that $y \mid xz$ implies $y \mid z$. So, let z be such that $y \mid xz$. Since $y \mid xz$, there is an r with $yr = xz$. Furthermore, since $x \mid xz$ and $xz = yr$, we get $x \mid yr$, and by $\text{coprime}(x, y)$ we have $x \mid r$. Thus, there is an s such that $xs = r$, and hence, $xsy = ry = yr = xz$. Now, by Lemma 8.4 we obtain $sy = z$, and therefore $y \mid z$ as desired. ⊣

If the binary relation coprime holds for x and y, then we say that x *and* y *are coprime*.

LEMMA 8.12. $\mathsf{PA} \vdash \forall x \forall y \forall k \big(k \,|\, x \wedge \mathrm{coprime}(x, y) \to \mathrm{coprime}(k, y)\big)$.

Proof. Assume that x and y are coprime. Let k be a divisor of x and let r be such that $rk = x$. Assume $y \,|\, kz$ for some arbitrary z. We have to show that $y \,|\, z$. First, notice that by LEMMA 8.10.(b) we have $y \,|\, rkz$, and since $rkz = xz$, we have $y \,|\, xz$. Now, since $\mathrm{coprime}(x, y)$, we obtain $y \,|\, z$ as desired. ⊣

The following result is crucial for the construction of *Gödel's β-function* (see THEOREM 9.9), which will be the key to the FIRST INCOMPLETENESS THEOREM 10.5.

LEMMA 8.13. $\mathsf{PA} \vdash \forall k \forall x \neq \mathsf{0} \forall j \Big(k \,|\, x \to \mathrm{coprime}\big(\mathsf{1} + (j + k)x, \mathsf{1} + jx\big)\Big)$

Proof. We first show $\mathsf{PA} \vdash \forall x \neq \mathsf{0} \forall j \big(\mathrm{coprime}(x, \mathsf{1} + jx)\big)$, i.e., we show that for all z,

$$x \,|\, (\mathsf{1} + jx)z \;\to\; x \,|\, z\,.$$

For this, suppose $x \,|\, (\mathsf{1} + jx)z$ for some arbitrary z. Since $(\mathsf{1} + jx)z = z + jxz$, by LEMMA 8.10.(b) we have $x \,|\, jxz$, and as a consequence of LEMMA 8.10.(a) we obtain $x \,|\, z$.

Now, let k and $x \neq \mathsf{0}$ be such that $k \,|\, x$. Notice that since $x \neq \mathsf{0}$, this implies that $k \neq \mathsf{0}$. Furthermore, let j be arbitrary but fixed. We have to show

$$\mathrm{coprime}\big(\mathsf{1} + jx, \mathsf{1} + (j + k)x\big)\,,$$

i.e., we have to show that for all z,

$$(\mathsf{1} + jx) \,|\, \big(\mathsf{1} + (j + k)x\big)z \;\to\; (\mathsf{1} + jx) \,|\, z\,.$$

First, notice that

$$\big(\mathsf{1} + (j + k)x\big)z = (\mathsf{1} + jx + kx)z = (\mathsf{1} + jx)z + kxz\,.$$

Assume now that for some z,

$$(\mathsf{1} + jx) \,|\, (\mathsf{1} + jx)z + kxz\,.$$

By LEMMA 8.10.(b) we have $(\mathsf{1} + jx) \,|\, (\mathsf{1} + jx)z$, and by LEMMA 8.10.(a) this implies $(\mathsf{1} + jx) \,|\, kxz$. Now, since $\mathrm{coprime}(x, \mathsf{1} + jx)$, as shown above, we get

$$(\mathsf{1} + jx) \,|\, x(kz) \to (\mathsf{1} + jx) \,|\, kz\,.$$

Finally, since by assumption $k \,|\, x$, by LEMMA 8.12 and $\mathrm{coprime}(x, \mathsf{1} + jx)$ we get $\mathrm{coprime}(k, \mathsf{1} + jx)$. Hence, we obtain $(\mathsf{1} + jx) \,|\, z$ as desired. ⊣

Alternative Induction Schemata

A fundamental principle in elementary number theory states that if there is a natural number fulfilling some property Ψ, then there must be a least natural number satisfying Ψ. This principle can be shown in PA; actually, every instance of this principle (i.e., by considering Ψ to be some $\mathscr{L}_{\mathsf{PA}}$-formula) is equivalent to the corresponding instance of the Induction Schema PA_6. In order to prove this, we need another induction principle which will turn out to be quite useful for further proofs in this book.

Proposition 8.14 (Strong Induction Principle). *Let $\varphi(x)$ be an $\mathscr{L}_{\mathsf{PA}}$-formula. Then in PA, φ satisfies the following* ***principle of strong induction****:*

$$\mathsf{PA} \vdash \forall x\big(\forall y < x\, \varphi(y) \to \varphi(x)\big) \to \forall x \varphi(x)$$

Proof. Suppose $\forall x\big(\forall y < x\, \varphi(y) \to \varphi(x)\big)$. Using PA_6, we first show $\forall x \psi(x)$ for

$$\psi :\equiv \forall y < x\, \varphi(y).$$

Notice that $\psi(\mathsf{0})$ vacuously holds, since there is no $y < \mathsf{0}$ with $\neg\varphi(y)$. Now, if $\psi(x)$ holds, then by our assumption we have $\varphi(x)$. So, we have $\psi(x)$ and $\varphi(x)$, which is the same as $\psi(\mathsf{s}x)$. Therefore, by PA_6 we obtain $\forall x \psi(x)$. Now, because for every x, $\psi(\mathsf{s}x)$ implies $\varphi(x)$, we finally obtain $\forall x \varphi(x)$. ⊣

Proposition 8.15 (Least Number Principle). *Let $\varphi(x)$ be an $\mathscr{L}_{\mathsf{PA}}$-formula. Then*

$$\mathsf{PA} \vdash \exists x \varphi(x) \to \exists x\big(\varphi(x) \wedge \forall y < x\, \neg\varphi(y)\big).$$

Informally, the Least Number Principle states that if there is a witness to an arithmetic statement, then there is always a least witness. This principle is often used in the following equivalent form: If a universally quantified formula does not hold, then there is a least counterexample.

Proof of Proposition 8.15. By Tautology (K) and the 3-Symbols Theorem 1.7, we have

$$\exists x \varphi(x) \to \exists x\big(\varphi(x) \wedge \forall y < x\, \neg\varphi(y)\big) \;\Leftrightarrow\; \forall x\, \neg\varphi(x) \vee \exists x\big(\varphi(x) \wedge \forall y < x\, \neg\varphi(y)\big),$$

where the latter statement is equivalent to the implication

$$\forall x\big(\neg\varphi(x) \vee \neg\forall y < x\, \neg\varphi(y)\big) \to \forall x\, \neg\varphi(x).$$

Now, by Tautology (K) this implication is equivalent to

$$\forall x\big(\forall y < x\, \neg\varphi(y) \to \neg\varphi(x)\big) \to \forall x\, \neg\varphi(x),$$

which is the Strong Induction Principle 8.14 applied to the formula $\neg\varphi(x)$. Consequently, we have $\mathsf{PA} \vdash \exists x\varphi(x) \to \exists x\big(\varphi(x) \wedge \forall y < x\, \neg\varphi(y)\big)$. ⊣

Relative Primality Revisited

We conclude this chapter by providing an alternative definition of relative primality, which shall be useful in the next chapter. First, we introduce the *Principle of Division with Remainder*:

Proposition 8.16 (Principle of Division with Remainder).

$$\mathsf{PA} \vdash \forall x\, \forall y > \mathsf{0}\, \exists q\, \exists r\big(x = qy + r \wedge r < y\big).$$

Proof. Let $\varphi(x) \equiv \forall y > \mathsf{0}\, \exists q\, \exists r\big(x = qy + r \wedge r < y\big)$. The proof is by induction on x. Obviously, we have $\varphi(\mathsf{0})$. Now, assume that we have $\varphi(x)$ for some x, i.e., for each $y > \mathsf{0}$ there are q, r such that

$$x = qy + r \wedge r < y\,.$$

If we replace x by $\mathsf{s}x$, then for each $y > \mathsf{0}$ there are q, r such that

$$\mathsf{s}x = qy + \mathsf{s}r \wedge \mathsf{s}r \leq y\,.$$

If $\mathsf{s}r < y$, let $r' := \mathsf{s}r$ and $q' := q$, and if $\mathsf{s}r = y$, let $r' := \mathsf{0}$ and $q' := \mathsf{s}q$. Now, in both cases we obtain

$$\mathsf{s}x = q'y + r' \ \wedge\ r' < y\,,$$

which shows $\varphi(\mathsf{s}x)$. ⊣

The following result gives a connection between the Principle of Division with Remainder and the relatively prime numbers:

Lemma 8.17. *For any $x, y > \mathsf{0}$ with $x = qy + r$ and $r < y$ we have*

$$\mathsf{PA} \vdash \mathrm{coprime}(y, x) \leftrightarrow \mathrm{coprime}(y, r)\,.$$

Proof. By definition we have $\mathrm{coprime}(y, x) \leftrightarrow \forall z(y \mid xz \to y \mid z)$, and since $x = qy + r$, we obtain

$$\mathrm{coprime}(y, x) \leftrightarrow \forall z(y \mid yqz + rz \to y \mid z)\,.$$

Now, by Lemma 8.10 we have $(y \mid yqz + rz) \leftrightarrow (y \mid rz)$, and therefore we obtain

$$\mathrm{coprime}(y, x) \ \leftrightarrow\ \forall z(y \mid rz \to y \mid z) \ \leftrightarrow\ \mathrm{coprime}(y, r)\,.$$

⊣

Now we are ready to give the promised alternative definition of relative primality.

Proposition 8.18.

$$\mathsf{PA} \vdash \forall x \forall y \Big(\mathrm{coprime}(x,y) \;\leftrightarrow\; x \neq \mathtt{0} \wedge y \neq \mathtt{0} \wedge \forall z \big((z \mid x \wedge z \mid y) \rightarrow z = \mathtt{1} \big) \Big).$$

Proof. The statement is obvious for $x = y$, or if at least one of x and y is equal to $\mathtt{1}$. Therefore, without loss of generality, let us assume that $x > y > \mathtt{1}$.

($\rightarrow$) The proof is by contraposition. Assume that there is a z such that $z \mid x$, $z \mid y$, and $z > \mathtt{1}$. Then, there is a $u < x$ such that $uz = x$. Now, since $z \mid y$, we obtain $x \mid yu$, and since $u < x$, we have $x \nmid u$, which implies $\neg\, \mathrm{coprime}(x,y)$.

($\leftarrow$) Assume towards a contradiction that there is a pair of numbers (x,y) with $x > y > \mathtt{0}$ such that for all z we have

$$(z \mid x \wedge z \mid y) \rightarrow z = \mathtt{1}\,,$$

but $\neg\, \mathrm{coprime}(x,y)$. By the Least Number Principle, let (x_0, y_0) be such a pair of numbers where x_0 is minimal. Let q and r be such that $x_0 = qy_0 + r$. Since $\neg\, \mathrm{coprime}(x_0, y_0)$, by Lemma 8.17 we have $\neg\, \mathrm{coprime}(y_0, r)$. On the other hand, if there is a $z_0 > \mathtt{1}$ with $z_0 \mid y_0$ and $z_0 \mid r$, then this would imply that

$$z_0 \mid qy_0 + r,$$

i.e., $z_0 \mid x_0$. But since $z_0 > \mathtt{1}$, this contradicts the fact that $(z_0 \mid x_0 \wedge z_0 \mid y_0) \rightarrow z_0 = \mathtt{1}$. Therefore, for the pair (y_0, r) we have $\neg\, \mathrm{coprime}(y_0, r)$, for all z we have

$$(z \mid y_0 \wedge z \mid r) \rightarrow z = \mathtt{1}\,,$$

and in addition we have $y_0 < x_0$, which contradicts the minimality of x_0. ⊣

As an immediate consequence of Proposition 8.18, we get the following

Corollary 8.19. *For all x and y, the following statement is provable in* PA*:*

$$\mathrm{coprime}(x,y) \;\leftrightarrow\; x \neq \mathtt{0} \wedge y \neq \mathtt{0} \wedge \forall z < (x+y) \big((z \mid x \wedge z \mid y) \rightarrow z = \mathtt{1} \big)$$

Exercises

8.0 Prove that addition is associtative, i.e., $\mathsf{PA} \vdash \forall x \forall y \forall z (x + (y + z) = (x + y) + z)$.

8.1 Introduce the unary relations $\mathrm{even}(x)$ and $\mathrm{odd}(x)$ formalising evenness and oddness, and show the statements

$$\mathsf{PA} \vdash \forall x \big(\mathrm{even}(x) \vee \mathrm{odd}(x) \big) \quad \text{and} \quad \mathsf{PA} \vdash \forall x \big(\mathrm{odd}(x) \rightarrow \mathrm{even}(\mathtt{s}x) \big)\,.$$

8.2 Show that Bézout's Lemma is provable in PA, i.e., show that

$$\mathsf{PA} \vdash \forall x \forall y \Big(\text{coprime}(x, y) \leftrightarrow \big(x \neq \mathtt{0} \wedge y \neq \mathtt{0} \wedge \exists a \leq y \, \exists b \leq x \, (ax + \mathtt{1} = by) \big) \Big) .$$

8.3 (a) Prove PA_6 from PA_0–PA_5, Lemma 8.3 and the Least Number Principle.

(b) Construct a model $\mathbf{M}$ for PA_0–PA_5 and the Least Number Principle, in which PA_6 does not hold.

8.4 Prove the following alternative induction principle:

$$\mathsf{PA} \vdash \big(\varphi(\mathtt{1}) \wedge \forall x (\varphi(x) \to \varphi(\mathtt{2}x) \wedge \varphi(x - \mathtt{1})) \big) \to \forall x \varphi(x)$$

Chapter 9
Gödelisation of Peano Arithmetic

The key ingredient for Gödel's Incompleteness Theorems is the so-called Gödelisation process which allows us to code terms, formulae and even proofs within PA. In order to achieve this, we introduce Gödel's β-function, with the help of which one can encode any FINITE sequence of natural numbers by a single natural number.

Natural Numbers in Peano Arithmetic

As we have already seen in Chapter 7, every standard natural number corresponds to a unique $\mathscr{L}_{\mathsf{PA}}$-term. More precisely, every element $\sigma\mathbf{0}$ of $\mathbb{N}$ corresponds to a term $\underline{\sigma}\mathsf{0}$. In order to simplify notations, from now on we will use variables such as $n, m, \ldots$ to denote elements of $\mathbb{N}$ and $\underline{n}, \underline{m}, \ldots$ their counterpart in the formal language $\mathscr{L}_{\mathsf{PA}}$, i.e., if n stands for $\sigma\mathbf{0}$, then $\underline{n}$ denotes $\underline{\sigma}\mathsf{0}$. Then FACT 7.0 yields

$$n \equiv m \quad \Longleftrightarrow \quad \mathsf{PA} \vdash \underline{n} = \underline{m}.$$

Moreover, by definition of $\underline{n}$ for $n \in \mathbb{N}$ we have:

$$\underline{\mathbf{0}} \equiv \mathsf{0}$$
$$\underline{\mathbf{s}n} \equiv \mathsf{s}\underline{n}$$

Furthermore, we define

$$\bigvee_{k=\mathbf{0}}^{n-1} x = \underline{k} \;:\equiv\; x = \underline{\mathbf{0}} \vee x = \underline{\mathbf{s0}} \vee \cdots \vee x = \underline{n-1}\,.$$

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_10

PROPOSITION 9.0. *Any two natural numbers $n, m \in \mathbb{N}$ satisfy the following properties:*

N_0: $\mathsf{PA} \vdash \mathtt{s}\underline{n} = \underline{\mathtt{s}n}$

N_1: $\mathsf{PA} \vdash \underline{m} + \underline{n} = \underline{m +^{\mathbb{N}} n}$

N_2: $\mathsf{PA} \vdash \underline{m} \cdot \underline{n} = \underline{m \cdot^{\mathbb{N}} n}$

N_3: *If* $m \equiv n$ *then* $\mathsf{PA} \vdash \underline{m} = \underline{n}$, *and if* $m \not\equiv n$ *then* $\mathsf{PA} \vdash \underline{m} \neq \underline{n}$.

N_4: *If* $m < n$ *then* $\mathsf{PA} \vdash \underline{m} < \underline{n}$, *and if* $m \not< n$ *then* $\mathsf{PA} \vdash \underline{m} \not< \underline{n}$.

N_5: $\mathsf{PA} \vdash \forall x \big(x < \underline{n} \leftrightarrow \bigvee_{k=\mathbf{0}}^{n-1} x = \underline{k} \big)$

Before we give a proof of PROPOSITION 9.0, let us recall the INDUCTION PRINCIPLE that we have introduced in Chapter 0: *If a statement A holds for $\mathbf{0}$ and if whenever A holds for a natural number n in $\mathbb{N}$ then it also holds for $n+1$, then the statement A holds for all natural numbers n in $\mathbb{N}$.* Note that this INDUCTION PRINCIPLE is more general than what we obtain from the induction axiom PA_6 in the standard model $\mathbb{N}$. The reason is that PA_6 is restricted to properties which can be described by an $\mathscr{L}_{\mathsf{PA}}$-formula, whereas the INDUCTION PRINCIPLE applies to any statement about standard natural numbers. In order to distinguish between the INDUCTION PRINCIPLE for standard natural numbers and induction within PA using PA_6, we shall call the former *metainduction*.

Proof of Proposition 9.0. N_0 follows directly from the definition of $\underline{n}$ for natural numbers $n \in \mathbb{N}$.

We prove N_1 by metainduction on n. The case $n \equiv \mathbf{0}$ is obviously true, since $\underline{\mathbf{0}}$ is $\mathtt{0}$. For the induction step, let us assume $\mathsf{PA} \vdash \underline{m} + \underline{n} = \underline{m +^{\mathbb{N}} n}$. Using N_0 and PA_3 both within PA and in $\mathbb{N}$ we obtain

$$\mathsf{PA} \vdash \underline{m} + \underline{\mathbf{s}n} = \underline{m} + \mathtt{s}\underline{n} = \mathtt{s}(\underline{m} + \underline{n}) = \mathtt{s}(\underline{m +^{\mathbb{N}} n}) = \underline{\mathbf{s}(m +^{\mathbb{N}} n)} = \underline{m +^{\mathbb{N}} \mathbf{s}n}.$$

The proof of N_2 is similar and is left as an exercise to the reader.

The first part of N_3 follows from FACT 7.0, and the second part is a consequence of N_4, since whenever $m \not\equiv n$, then either $m < n$ or $n < m$.

Let us now turn to N_4. If $m < n$, then there is $k \in \mathbb{N}$ such that $m +^{\mathbb{N}} k \equiv n$ and $k \not\equiv \mathbf{0}$. By N_3 and N_1 we get $\mathsf{PA} \vdash \underline{m} + \underline{k} = \underline{m +^{\mathbb{N}} k} = \underline{n}$. It remains to show that $\mathsf{PA} \vdash \underline{k} \neq \mathtt{0}$. Since $k \not\equiv \mathbf{0}$, it is of the form $\mathbf{s}k'$ for some $k' \in \mathbb{N}$. Thus by N_0 and PA_0, $\mathsf{PA} \vdash \underline{k} = \underline{\mathbf{s}k'} = \mathtt{s}\underline{k'} \neq \mathtt{0}$. The second statement of N_4 follows from the first one and N_3 by observing that if $m \not< n$, then either $m \equiv n$ or $n < m$.

In order to prove N_5, we proceed by metainduction on n. The case $n \equiv \mathbf{0}$ is trivially satisfied. Now, assume that N_5 holds for some n and let $x < \underline{\mathbf{s}n} = \mathtt{s}\underline{n}$. Then, by LEMMA 8.5 we get $x \leq \underline{n}$, i.e., either $x < \underline{n}$ or $x = \underline{n}$. Since the first case is equivalent to $\bigvee_{k=\mathbf{0}}^{n-1} x = \underline{k}$ by assumption, we obtain $\bigvee_{k=\mathbf{0}}^{n} x = \underline{k}$ as desired. The converse is a consequence of N_4. ⊣

On the one hand, by the SOUNDNESS THEOREM 3.7 we know that every statement which is provable within PA holds in every model of PA, in particular in the standard model $\mathbb{N}$. On the other hand, not every statement which is true in $\mathbb{N}$ must be provable within PA. In this respect, PROPOSITION 9.0 gives us a few statements which are true in the standard model $\mathbb{N}$ and which are provable within PA. In order to obtain more such statements, we shall introduce the notion of $\mathbb{N}$-*conformity*; but before doing so, we have to give a few preliminary notions.

We call an $\mathscr{L}_{\mathsf{PA}}$-formula φ a **strict $\exists$-formula** if it is built up from atomic formulae and negated atomic formulae using $\wedge$, $\vee$, existential quantification $\exists\nu$ and bounded universal quantification, i.e., "$\forall\nu < \tau$" for some term τ. Furthermore, φ is said to be an **$\exists$-formula** if there is a strict $\exists$-formula ψ such that $\varphi \Leftrightarrow_{\mathsf{PA}} \psi$. By exchanging the role of universal and existential quantification in the above definition, we can analogously define (**strict**) **$\forall$-formulae**. Furthermore, if a formula is both an $\exists$- and a $\forall$-formula, then we call it a ***Δ*-formula**. In particular, every formula which contains only bounded quantifiers is a Δ-formula.

EXAMPLE 9.1. The formulae "$x \leq y$" and "$x \mid y$" are Δ-formulae:

$$x \leq y \;\Leftrightarrow_{\mathsf{PA}}\; \exists r < \mathsf{s}y\,(x + r = y) \quad \text{and} \quad x \mid y \;\Leftrightarrow_{\mathsf{PA}}\; \exists r < \mathsf{s}y\,(rx = y)$$

PROPOSITION 9.2. *Let $\varphi(x_1, \ldots, x_n)$ be a formula whose free variables are among $x_1, \ldots, x_n$, and let $a_1, \ldots, a_n \in \mathbb{N}$.*

(a) *If φ is an $\exists$-formula and $\mathbb{N} \vDash \varphi(a_1, \ldots, a_n)$, then $\mathsf{PA} \vdash \varphi(\underline{a_1}, \ldots, \underline{a_n})$.*

(b) *If φ is a $\forall$-formula and $\mathbb{N} \vDash \neg\varphi(a_1, \ldots, a_n)$, then $\mathsf{PA} \vdash \neg\varphi(\underline{a_1}, \ldots, \underline{a_n})$.*

Proof. Observe first that (b) follows from (a), since the negation of a $\forall$-formula is an $\exists$-formula. Furthermore, note that it is enough to prove (a) for strict $\exists$-formulae. We proceed by induction on the construction of φ.

- If φ is an atomic formula, then it is of the form $\tau_0(x_1, \ldots, x_n) = \tau_1(x_1, \ldots, x_n)$ for some terms τ_0, τ_1 whose variables are among $x_1, \ldots, x_n$. We show by induction on the construction of terms that for every term $\tau(x_1, \ldots, x_n)$ and for all standard natural numbers $a_1, \ldots, a_n \in \mathbb{N}$, we have

$$\mathsf{PA} \vdash \underline{\tau^{\mathbb{N}}(a_1, \ldots, a_n)} = \tau(\underline{a_1}, \ldots, \underline{a_n}). \qquad (*)$$

  The statement is clear for terms τ of the form $\tau \equiv \nu$ (for a variable ν) or $\tau \equiv 0$. If τ is of the form $\mathsf{s}\tau'$ for some term τ', by the induction hypothesis we have $\mathsf{PA} \vdash \underline{a} = \tau'(\underline{a_1}, \ldots, \underline{a_n})$, where $a = \tau'^{\mathbb{N}}(a_1, \ldots, a_n) \in \mathbb{N}$. Therefore, $\tau^{\mathbb{N}}(a_1, \ldots, a_n)$ is $\mathsf{s}a$. Then by N_0 we have

$$\mathsf{PA} \vdash \underline{\mathsf{s}a} = \mathsf{s}\underline{a} \wedge \mathsf{s}\underline{a} = \mathsf{s}\tau'(\underline{a_1}, \ldots, \underline{a_n}) \wedge \mathsf{s}\tau'(\underline{a_1}, \ldots, \underline{a_n}) = \tau(\underline{a_1}, \ldots, \underline{a_n})$$

as desired. The proofs for terms of the form $\tau_0 + \tau_1$ or $\tau_0 \cdot \tau_1$ are similar using N_1 and N_2, respectively. This shows $(*)$.
Now assume $\mathbb{N} \vDash \tau_0(a_1, \ldots, a_n) = \tau_1(a_1, \ldots, a_n)$ and put $a \equiv \tau_0(a_1, \ldots, a_n)$ and $b \equiv \tau_1(a_1, \ldots, a_n)$. Then by $(*)$ and N_3 we get

$$\mathsf{PA} \vdash \tau_0(\underline{a_1}, \ldots, \underline{a_n}) = \underline{a} \wedge \underline{a} = \underline{b} \wedge \underline{b} = \tau_1(\underline{a_1}, \ldots, \underline{a_n}).$$

- If φ is a negated atomic formula, then it is of the form $\tau_0 \neq \tau_1$ for some terms τ_0 and τ_1, and since

$$\tau_0 \neq \tau_1 \Leftrightarrow_{\mathsf{PA}} \exists y\big(\tau_0 + \mathtt{s}y = \tau_1 \vee \tau_0 = \tau_1 + \mathtt{s}y\big)\,,$$

this case follows from the cases below.

- Suppose that $\varphi(x_1, \ldots, x_n) \equiv \varphi_0(x_1, \ldots, x_n) \wedge \varphi_1(x_1, \ldots, x_n)$ and $\mathbb{N} \vDash \varphi(a_1, \ldots, a_n)$. Then $\mathbb{N} \vDash \varphi_0(a_1, \ldots, a_n)$ and $\mathbb{N} \vDash \varphi_1(a_1, \ldots, a_n)$. By induction hypothesis, $\mathsf{PA} \vdash \varphi_0(\underline{a_1}, \ldots \underline{a_n})$ and $\mathsf{PA} \vdash \varphi_1(\underline{a_1}, \ldots, \underline{a_n})$. Using $(\mathsf{I}\wedge)$ this shows that $\mathsf{PA} \vdash \varphi(\underline{a_1}, \ldots, \underline{a_n})$. The disjunctive case is similar.
- Let now $\varphi(x_1, \ldots, x_n) \equiv \forall y < \tau(x_1, \ldots, x_n)\, \psi(x_1, \ldots, x_n, y)$ and suppose that $\mathbb{N} \vDash \varphi(a_1, \ldots, a_n)$. Let $a \equiv \tau^{\mathbb{N}}(a_1, \ldots, a_n)$. Then for every $b < a$ we have $\mathbb{N} \vDash \psi(a_1, \ldots, a_n, b)$. Hence, by induction hypothesis, for every $b < a$ we have $\mathsf{PA} \vdash \psi(\underline{a_1}, \ldots, \underline{a_n}, \underline{b})$, and by $(*)$, $\mathsf{PA} \vdash \underline{a} = \tau(\underline{a_1}, \ldots, \underline{a_n})$. Now using N_5 we obtain

$$\mathsf{PA} \vdash\ \varphi(\underline{a_1}, \ldots, \underline{a_n}) \leftrightarrow \forall y\Big(\bigvee_{b=0}^{a-1} y = \underline{b} \to \psi(\underline{a_1}, \ldots, \underline{a_n}, y)\Big).$$

The right-hand side can clearly be derived in PA.

- Finally, let $\varphi(x_1, \ldots, x_n) \equiv \exists y \psi(x_1, \ldots, x_n, y)$. Then $\mathbb{N} \vDash \varphi(a_1, \ldots, a_n)$ implies that there exists $b \in \mathbb{N}$ such that $\mathbb{N} \vDash \psi(a_1, \ldots, a_n, b)$. Inductively, we get $\mathbb{N} \vDash \psi(\underline{a_1}, \ldots, \underline{a_n}, \underline{b})$, which completes the proof.

⊣

Note that any constants, relations, and functions that one can define in PA in the sense of Chapter 6 can be interpreted in the standard model $\mathbb{N}$.

A relation $R(x_1, \ldots, x_n)$ defined by

$$R(x_1, \ldots, x_n) :\Longleftrightarrow \psi_R(x_1, \ldots, x_n)$$

is said to be $\mathbb{N}$**-conform** if for all $a_1, \ldots, a_n \in \mathbb{N}$ the following two properties are satisfied:

(a) If $\mathbb{N} \vDash \psi_R(a_1, \ldots, a_n)$, then $\mathsf{PA} \vdash \psi_R(\underline{a_1}, \ldots, \underline{a_n})$.
(b) If $\mathbb{N} \vDash \neg\psi_R(a_1, \ldots, a_n)$, then $\mathsf{PA} \vdash \neg\psi_R(\underline{a_1}, \ldots, \underline{a_n})$.

For the sake of simplicity, the formula ψ_R is also called $\mathbb{N}$-conform.

Now, let f be a function symbol whose defining formula is $\psi_f(x_1, \cdots, x_n, y)$, i.e., $\mathsf{PA} \vdash \forall x_1 \ldots \forall x_n \exists! y\, \psi_f(x_1, \ldots, x_n, y)$ and

$$f x_0 \cdots x_n = y :\Longleftrightarrow \psi_f(x_1, \ldots, x_n, y).$$

Then we say that f is $\mathbb{N}$**-conform** if its defining formula ψ_f is $\mathbb{N}$-conform. Let $f^{\mathbb{N}}$ be the interpretation of f in $\mathbb{N}$. If f is $\mathbb{N}$-conform, then for all $a_1, \ldots, a_n \in \mathbb{N}$

$$\mathsf{PA} \vdash \psi_f(\underline{a_1}, \ldots, \underline{a_n}, \underline{f^{\mathbb{N}}(a_1, \ldots, a_n)}),$$

and hence

$$\mathsf{PA} \vdash f(\underline{a_1}, \ldots, \underline{a_n}) = \underline{f^{\mathbb{N}}(a_1, \ldots, a_n)}.$$

To see this, suppose that f is $\mathbb{N}$-conform. For the sake of simplicity, suppose that $n \equiv 1$ and let $a \in \mathbb{N}$. Then $\mathbb{N} \vDash \psi_f(a, f^{\mathbb{N}}(a))$, hence by $\mathbb{N}$-conformity we get $\mathsf{PA} \vdash \psi_f(\underline{a}, \underline{f^{\mathbb{N}}(a)})$. On the other hand, we have $\mathsf{PA} \vdash \psi_f(\underline{a}, f(\underline{a}))$ and hence by functionality of ψ_f we get $\mathsf{PA} \vdash f(\underline{a}) = \underline{f^{\mathbb{N}}(a)}$.

Corollary 9.3.

(a) *Every relation which is defined by a Δ-formula is $\mathbb{N}$-conform.*

(b) *Every function which is defined by an $\exists$-formula is $\mathbb{N}$-conform.*

Proof. Condition (a) follows directly from Proposition 9.2. For (b), it suffices to prove that every function whose defining formula is an $\exists$-formula is already a Δ-formula. Suppose that f is defined by the $\exists$-formula ψ_f, i.e.,

$$f(x_1, \ldots, x_n) = y :\Longleftrightarrow \psi_f(x_1, \ldots, x_n, y).$$

Now note that by functionality of ψ_f we have

$$\psi_f(x_1, \ldots, x_n, y) \Leftrightarrow_{\mathsf{PA}} \forall z(\psi_f(x_1, \ldots, x_n, z) \to z = y).$$

Moreover, Tautology (K) yields

$$\forall z(\psi_f(x_1, \ldots, x_n, z) \to z = y) \Leftrightarrow \forall z(\neg\psi_f(x_1, \ldots, x_n, z) \vee z = y),$$

which is a $\forall$-formula. ⊣

Example 9.4. The binary coprimality relation "coprime" is $\mathbb{N}$-conform. To see this, first notice that by the previous example, the defining formula of the divisibility relation is a Δ-formula, and therefore, by Corollary 9.3, the symbol $\mid$ is $\mathbb{N}$-conform. Furthermore, by Corollary 8.19 the defining formula of "coprime" is equivalent to a Δ-formula, and therefore, by Corollary 9.3 the binary relation "coprime" is $\mathbb{N}$-conform.

Gödel's β-Function

The main goal of this section is to define a binary function (the so-called **β-function** introduced by Kurt Gödel) which encodes a FINITE sequence of natural numbers $c_0, \cdots, c_{n-1}$ in the standard model by a single number c such that for all $i < n$,

$$\mathsf{PA} \vdash \beta(\underline{c}, \underline{i}) = \underline{c_i}.$$

In fact, one can even do better than that and introduce a function β such that for every unary function f definable in Peano Arithmetic,

$$\mathsf{PA} \vdash \forall k \exists c \forall i < k \big(\beta(c, i) = f(i)\big).$$

The first step is to encode *ordered pairs* of numbers by introducing a binary pairing function op. We define

$$\mathrm{op}(x, y) = z :\Longleftrightarrow (x + y) \cdot (x + y) + x + \mathsf{1} = z\,.$$

Furthermore, we define the unary relation *not an ordered pair* "nop" and the two binary functions *first element* "fst" and *second element* "snd" by stipulating

$$\begin{aligned}
\mathrm{nop}(c) &:\Longleftrightarrow \neg\exists x \exists y \big(\mathrm{op}(x, y) = c\big)\,,\\
\mathrm{fst}(c) = x &:\Longleftrightarrow \exists y \big(\mathrm{op}(x, y) = c\big) \vee \big(\mathrm{nop}(c) \wedge x = \mathsf{0}\big)\,,\\
\mathrm{snd}(c) = y &:\Longleftrightarrow \exists x \big(\mathrm{op}(x, y) = c\big) \vee \big(\mathrm{nop}(c) \wedge y = \mathsf{0}\big)\,.
\end{aligned}$$

In particular, whenever $\mathrm{op}(x, y) = c$, then

$$c = \mathrm{op}\big(\mathrm{fst}(c), \mathrm{snd}(c)\big)\,.$$

Until now, we did not show that the above definitions are well-defined. This, however, follows from the following

LEMMA 9.5. $\mathsf{PA} \vdash \mathrm{op}(x, y) = \mathrm{op}(x', y') \to x = x' \wedge y = y'$.

Proof. Assume that $\mathrm{op}(x, y) = \mathrm{op}(x', y')$. We first show that this implies $x + y = x' + y'$: Suppose towards a contradiction that $x + y < x' + y'$. Then, by $\mathsf{PA_3}$ and LEMMA 8.5, we obtain $\mathsf{s}(x + y) = x + \mathsf{s}y \leq x' + y'$. Therefore,

$$\begin{aligned}
\mathrm{op}(x', y') = \mathrm{op}(x, y) &= (x + y) \cdot (x + y) + x + \mathsf{1}\\
&\leq (x + \mathsf{s}y) \cdot (x + y) + (x + \mathsf{s}y)\\
&= (x + \mathsf{s}y) \cdot (x + \mathsf{s}y)\\
&= \mathsf{s}(x + y) \cdot \mathsf{s}(x + y)\\
&\leq (x' + y') \cdot (x' + y')\\
&< \mathrm{op}(x', y'),
\end{aligned}$$

which is obviously a contradiction. By symmetry, the relation $x'+y' < x+y$ can also be ruled out, and therefore we have that $\mathrm{op}(x,y) = \mathrm{op}(x',y')$ implies $x+y = x'+y'$. Now, if $x+y = x'+y'$, then $(x+y)\cdot(x+y) = (x'+y')\cdot(x'+y')$, and since $\mathrm{op}(x,y) = \mathrm{op}(x',y')$, by Lemma 8.2 we obtain $x+\mathtt{1} = x'+\mathtt{1}$, which implies $x = x'$ and also $y = y'$. ⊣

Now we are ready to define the β-function. Let

$$\gamma(a,i,y,x) :\equiv \big(1 + (\mathrm{op}(a,i)+1)\cdot y\big) \mid x$$

and define

$$\begin{aligned}\beta(c,i) = a \;:\Longleftrightarrow\; &\big(\mathrm{nop}(c) \wedge a = \mathtt{0}\big) \vee \\ &\exists x \exists y \Big(c = \mathrm{op}(x,y) \wedge \Big(\big(\forall b < x\ (\neg\gamma(b,i,y,x)) \wedge a = \mathtt{0}\big) \vee \\ &\qquad \big(\gamma(a,i,y,x) \wedge \forall b < a\ (\neg\gamma(b,i,y,x))\big)\Big)\Big).\end{aligned}$$

Slightly less formal, we can define $\beta(c,i)$ by stipulating

$$\beta(c,i) = \begin{cases} \mathtt{0} & \text{if } \mathrm{nop}(c), \\ \mathtt{0} & \text{if } \exists x \exists y \big(c = \mathrm{op}(x,y) \wedge \neg\exists b < x\ (\gamma(b,i,y,x))\big), \\ a & \text{if } \exists x \exists y \big(c = \mathrm{op}(x,y) \wedge a = \min\big\{b : \gamma(b,i,y,x)\big\}\big). \end{cases}$$

Observe that as a consequence of the Least Number Principle, β is a binary function.

Before we can encode finite sequences with the β-function, we have to prove a few auxiliary results. The first one states that for every m there exists a y which is a multiple of $\mathrm{lcm}(1,\ldots,m)$.

Lemma 9.6. $\mathsf{PA} \vdash \forall m \exists y \forall k \big((k \neq \mathtt{0} \wedge k \leq m) \to k \mid y\big)$.

Proof. We proceed by induction on m. The case when $m = \mathtt{0}$ is clear. Assume that there is a y such that for every k with $\mathtt{0} < k \leq m$ we have $k \mid y$. Let $y' = y \cdot \mathtt{s}m$. Then, by Lemma 8.10, every k with $\mathtt{0} < k \leq \mathtt{s}m$ divides y'. ⊣

As described in Chapter 6, for any $\mathscr{L}_{\mathsf{PA}}$-formula φ which is *functional*, i.e., $\mathsf{PA} \vdash \forall x \exists! y \varphi(x,y)$, we can introduce a function symbol F_φ by stipulating

$$F_\varphi(x) = y :\Longleftrightarrow \varphi(x,y)\,.$$

If F is defined by some functional $\mathscr{L}_{\mathsf{PA}}$-formula, then we say that F is **definable** in PA.

The next result shows that for every function F which is definable in PA and for every $k > \mathtt{0}$, we can define $\max\big\{F(\mathtt{0}),\ldots,F(k-\mathtt{1})\big\}$.

LEMMA 9.7. *Let F be a function which is definable in* PA. *Then*

$$\mathsf{PA} \vdash \forall k > \mathtt{0}\, \exists! x \big(\exists i < k \big(F(i) = x\big) \wedge \forall i < k \big(F(i) \leq x\big)\big).$$

Proof. We prove the statement by induction on k starting with $\mathtt{1}$. For $k = \mathtt{1}$, one can clearly take $x = F(\mathtt{0})$. Assume that there is a unique x and there is $i_0 < k$ such that $F(i_0) = x$ and for all $i < k$, $F(i) \leq x$. Now if $F(k) \leq x$, then set $x' = x$; otherwise let $x' = F(k)$. Then for every $i < \mathtt{s}k$, we have $F(i) \leq x$ and the first condition is also satisfied since x' is either $F(i_0)$ or $F(k)$; uniqueness is trivial. ⊣

This leads to the following definition:

$$\max_{i<k} F(i) = x :\Longleftrightarrow \exists i < k \big(F(i) = x\big) \wedge \forall i < k \big(F(i) \leq x\big)$$

The next result plays an important role in the coding of finite sequences.

LEMMA 9.8. *Let G be a unary, strictly increasing function with $G(0) > 1$ which is definable in* PA *and let $\varphi(\nu)$ be an $\mathscr{L}_{\mathsf{PA}}$-formula. Then*

$$\mathsf{PA} \vdash \forall m \Big(\forall j < m\, \forall j' < m \big(j \neq j' \to \mathrm{coprime}(G(j), G(j'))\big) \\ \to \exists x\, \forall j < m \, \big(G(j) \mid x \leftrightarrow \varphi(j)\big)\Big).$$

Proof. We proceed by induction on m starting with $m = 1$. If $\varphi(\mathtt{0})$ holds, let $x := G(\mathtt{0})$, otherwise let $x := \mathtt{1}$. For the induction step, assume that for all distinct $j, j' \leq \mathtt{s}m$, $G(j)$ and $G(j')$ are coprime and that there is an x such that for all $j < m$, $G(j) \mid x \leftrightarrow \varphi(j)$. By the LEAST NUMBER PRINCIPLE, let x_0 be the least such x. Now we consider the following two cases: If $\varphi(m)$ holds, let $x_1 := G(m) \cdot x_0$, otherwise, let $x_1 := x_0$. It remains to show that for all $j \leq m$ we have $G(j) \mid x_1 \leftrightarrow \varphi(j)$.

If $\varphi(m)$ fails (i.e., $x_1 = x_0$), then, by induction hypothesis, for all $j < m$ we have $G(j) \mid x_0 \leftrightarrow \varphi(j)$ and $\mathrm{coprime}(G(j), G(m))$, where the latter implies by the choice of x_0 that $G(m) \nmid x_0$. To see this, assume that $G(m) \mid x_0$. Then there is an r such that $G(m) \cdot r = x_0$, and since $m \geq \mathtt{1}$, G is strictly increasing and $G(\mathtt{0}) \neq \mathtt{0}$ by $\mathrm{coprime}(G(\mathtt{0}), G(\mathtt{1}))$, we get that $G(m) > \mathtt{1}$ and consequently $r < x_0$. Moreover, since for all $j < m$ we have $\mathrm{coprime}(G(j), G(m))$, this implies

$$\forall j < m \big(G(j) \mid \underbrace{G(m) \cdot r}_{= x_0} \;\leftrightarrow\; G(j) \mid r\big)\,,$$

which contradicts the minimality of x_0.

If $\varphi(m)$ holds (i.e., $x_1 = G(m) \cdot x_0$), then, since $\mathrm{coprime}(G(j), G(m))$ for all $j < m$ we have

$$G(j) \mid \underbrace{G(m) \cdot x_0}_{= x_1} \;\leftrightarrow\; G(j) \mid x_0\,.$$

Furthermore, we obviously have $G(m) \mid x_1$, which completes the proof. ⊣

The following theorem states how the β-function can be used to code finite sequences.

THEOREM 9.9. *Let F be a function which is definable in* PA. *Then*

$$\mathsf{PA} \vdash \forall k\, \exists c\, \forall i < k\, \big(\beta(c,i) = F(i)\big).$$

Proof. Fix an arbitrary number k. Let $F'(i) := \mathrm{op}\big(F(i), i\big) + 1$ and let

$$m := \max\nolimits_{i<k} F'(i)\,.$$

By LEMMA 9.6 there is a y such that every $j \le m$ divides y. Furthermore, by LEMMA 8.13 we have for all u with $u \mid y$ (i.e., $1 \le u \le m$) and for all w

$$\mathrm{coprime}\big(1 + wy, 1 + (w+u)y\big)\,.$$

In particular, if $i < j < m$, then for $w := i + 1$ and $u := j - i$, we obtain

$$\mathrm{coprime}\big(1 + (i+1)y, 1 + (j+1)y\big)\,.$$

Finally, define the unary function G by

$$G(j) = z \;:\Longleftrightarrow\; z = 1 + (j+1)y\,,$$

and let

$$\varphi_0(z) \;:\equiv\; \exists i < k\big(z = \mathrm{op}(F(i), i)\big)\,.$$

Then G is a strictly increasing function and we can apply LEMMA 9.8 in order to find a number x such that for all $j < m$, where $m \ge \mathrm{op}\big(F(i), i\big)$ (for all $i < k$), we have

$$G(j) \mid x \;\leftrightarrow\; \varphi_0(j)\,,$$

in other words,

$$\forall j < m \Big(1 + (j+1)y \mid x \;\leftrightarrow\; \exists i < k\,\big(j = \mathrm{op}(F(i), i)\big)\Big)\,.$$

It remains to show that for $c = \mathrm{op}(x, y)$ we have $\beta(c, i) = F(i)$ for all $i < k$. By our assumption on x, we have $1 + (\mathrm{op}(F(i), i) + 1)y \mid x$, i.e., $\gamma\big(F(i), i, y, x\big)$. Therefore, it is enough to check that $F(i)$ is minimal with this property. Assume towards a contradiction that there is an $a < F(i)$ with $\gamma(a, i, y, x)$, i.e.,

$$1 + (\mathrm{op}(a, i) + 1)y \mid x\,.$$

Then, by the formula φ_0, there is a j with $j = \mathrm{op}(a, i) = \mathrm{op}(F(i'), i')$ for some $i' < k$. Thus, by LEMMA 9.5, we have $i = i'$ and $a = F(i') = F(i)$, which is a contradiction to the assumption $a < F(i)$. ⊣

Note that all functions — in particular the β-function — which we have introduced in this section can be defined by an $\exists$-formula and are therefore $\mathbb{N}$-conform.

Encoding Finite Sequences

This section aims at showing how the β-function can be used to encode a finite sequence of numbers; but what is meant by the words "finite" and "number"? In the standard model, this coincides with FINITENESS and the usual natural numbers. In general, however, this means that the sequence has a limited length k for some k, i.e., in non-standard models its length can actually be a non-standard number.

In a naive way, sequences of natural numbers can be viewed as functions from some $\{0, \cdots, n\}$ to the natural numbers, where $\{0, \cdots, n\}$ is the domain of the function. In PA, however, we cannot specify the domain of a definable function, which is why we will use $\beta(\cdot, 0)$ to encode the length of a sequence. Concretely, we will encode $\langle F(i) \mid i < n\rangle$ using some c (whose existence is guaranteed by THEOREM 9.9) such that

$$\begin{aligned} \beta(c, 0) &= n \\ \forall i < n(\beta(c, i+1) &= F(i)). \end{aligned}$$

This motivates us to introduce the functions

$$\begin{aligned} \mathrm{lh}(c) &:\equiv \beta(c, 0) \\ c_i &:\equiv \beta(c, i+1). \end{aligned}$$

We will also call $\mathrm{lh}(c)$ the *length* of c. Furthermore, we define s to be a *sequence*, (denoted $\mathrm{seq}(s)$), if s is the smallest code for $\langle s_i \mid i < \mathrm{lh}(s)\rangle$:

$$\mathrm{seq}(s) :\Longleftrightarrow \forall t < s\big(\mathrm{lh}(t) = \mathrm{lh}(s) \to \exists i < \mathrm{lh}(s)(t_i \neq s_i)\big).$$

Note that the definition of seq assures that codes for finite sequences are unique, i.e.,

$$\mathsf{PA} \vdash \big(\mathrm{seq}(s) \wedge \mathrm{seq}(s') \wedge \mathrm{lh}(s) = \mathrm{lh}(s') \wedge \forall i < \mathrm{lh}(s)(s_i = s'_i)\big) \to s = s'.$$

EXAMPLE 9.10. The simplest example is the empty sequence $\langle\,\rangle$ which is defined by $\langle\,\rangle = s :\Longleftrightarrow \mathrm{seq}(s) \wedge \mathrm{lh}(s) = 0$. By taking a closer look at the definition of the β-function, one can easily see that $\langle\,\rangle$ is actually 0, since it is the smallest code s with $\beta(s, 0) = 0$.

Secondly, we consider one-element sequences: The sequence just consisting of x is given by

$$\langle x\rangle = s :\Longleftrightarrow \mathrm{seq}(s) \wedge \mathrm{lh}(s) = 1 \wedge s_0 = x.$$

In the same way, one can define two-element sequences as

$$\langle x, y\rangle = s :\Longleftrightarrow \mathrm{seq}(s) \wedge \mathrm{lh}(s) = 2 \wedge s_0 = x \wedge s_1 = y.$$

More generally, if F is definable in PA, then one can define

$$\langle F(i) \mid i < k\rangle = s :\Longleftrightarrow \mathrm{seq}(s) \wedge \mathrm{lh}(s) = k \wedge \forall i < k\big(s_i = F(i)\big).$$

Theorem 9.9 assures that such a number s always exists and since it is the least code it is unique.

The functions $c, i \mapsto c_i, \mathrm{lh}$ and $\langle\cdot\rangle$ are all defined by $\exists$-formulae and are thus $\mathbb{N}$-conform as a consequence of Corollary 9.3. We will use the same notation for the corresponding function in $\mathbb{N}$, for example we write $\langle n, m\rangle$ for $\langle n, m\rangle^{\mathbb{N}}$.

Next, we show how finite sequences can be concatenated.

Proposition 9.11.

$$\mathsf{PA} \vdash \forall s \forall s' \exists t \Big(\mathrm{seq}(t) \wedge \mathrm{lh}(t) = \mathrm{lh}(s) + \mathrm{lh}(s') \wedge \forall i < \mathrm{lh}(s)(t_i = s_i) \wedge \forall i < \mathrm{lh}(s')(t_{\mathrm{lh}(s)+i} = s'_i)\Big)$$

Proof. Put $F(0) = \mathrm{lh}(s) + \mathrm{lh}(t)$, $F(i) = \beta(s, i)$ for $0 < i < \mathrm{lh}(s) + 1$, and $F(i) = \beta(t, i - \mathrm{lh}(s))$ for $i \geq \mathrm{lh}(s) + 1$. This clearly defines a function, so we can apply Theorem 9.9 and obtain a code t such that

$$\text{for all } \; i < \mathrm{lh}(s) + \mathrm{lh}(t) + 1 \; \text{ we have } \; \beta(t, i) = F(i)\,.$$

In particular, this means that $\mathrm{lh}(t) = \mathrm{lh}(s) + \mathrm{lh}(s')$, $(t)_i = \beta(t, i + 1) = \beta(s, i + 1) = s_i$ for $i < \mathrm{lh}(s)$. Similarly, we get $t_{\mathrm{lh}(s)+i} = s'_i$ for $i < \mathrm{lh}(s')$. The Least Number Principle then enables us to choose t minimal with the properties from above, i.e., such that $\mathrm{seq}(t)$. ⊣

With Proposition 9.11, we can define

$$s^\frown s' = t :\Longleftrightarrow \mathrm{seq}(t) \wedge \mathrm{lh}(t) = \mathrm{lh}(s) + \mathrm{lh}(s') \wedge \forall i < \mathrm{lh}(s)(t_i = s_i) \wedge \forall i < \mathrm{lh}(s')(t_{\mathrm{lh}(s)+i} = s'_i)\,.$$

Note that by Proposition 9.11, $s^\frown s'$ is functional. Moreover, it is easy to check that concatenation is associative, i.e.,

$$\mathsf{PA} \vdash (s^\frown s')^\frown s'' = s^\frown(s'^\frown s'').$$

Therefore, we can omit the brackets and write $s^\frown s'^\frown s''$ instead of $s^\frown(s'^\frown s'')$.

Encoding Power Functions

In the previous paragraphs, we have seen how the β-function allows us to encode finite sequences. Now we will use these insights to show how recursive functions can be defined in PA. We will not do this in general, since the only crucial function we need is the power function. Further examples of recursive functions can be found in the exercises.

The definability of the power function is remarkable, since it means that we can define exponentiation from addition and multiplication; however, as we will see in Chapter 12, multiplication cannot be defined from addition. The idea is to interpret the power x^k as the sequence $\langle 1, x, \cdots, x^{k-1}, x^k\rangle$ of length $k+1$.

We introduce the function x^k by stipulating

$$x^k = y :\Longleftrightarrow \exists t\big(\mathrm{seq}(t) \wedge \mathrm{lh}(t) = \mathsf{s}k \wedge t_0 = 1 \wedge \forall i < k(t_{\mathsf{s}i} = x \cdot t_i) \wedge t_k = y\big).$$

Why is x^k functional? Clearly, the function x^k has (if defined) a unique value. In order to see that it is always defined, we can use induction: For $k = 0$ it is clear. Now assume that there is a sequence s of length $k+1$ such that $s_0 = 1$ and for all $i < k$ we have $s_{\mathsf{s}i} = x \cdot s_i$. Consider $t = s^\frown\langle x \cdot s_k\rangle$. Then t is a sequence of length $k+2$ which satisfies the desired properties.

Note that the power function is defined by an $\exists$-formula and therefore $\mathbb{N}$-conform by Corollary 9.3. Furthermore, observe that the power function fulfils the usual recursive definition, i.e.,

$$\mathsf{PA} \vdash \forall x(x^0 = 1)$$
$$\mathsf{PA} \vdash \forall x \forall k(x^{\mathsf{s}k} = x \cdot x^k).$$

Our next aim is to encode terms, formulae and proofs by making use of unique prime decomposition. This can be shown in PA. However, for us it suffices to show that the function mapping x, y, z to $2^x \cdot 3^y \cdot 5^z$ is injective. The general result is left as an exercise to the interested reader. We define primality by

$$\mathrm{prime}(x) :\Longleftrightarrow x > 1 \wedge \forall z\big(z \mid x \to (z = x \vee z = 1)\big).$$

If $\mathrm{prime}(x)$, we say that x is *prime*. Note that prime can be defined by an $\exists$-formula, since $\forall z$ can be replaced by $\forall z \leq x$. By using the fact that $2, 3$ and 5 are prime in the standard model $\mathbb{N}$ and by Proposition 9.2, we obtain

$$\mathsf{PA} \vdash \mathrm{prime}(2) \wedge \mathrm{prime}(3) \wedge \mathrm{prime}(5).$$

Moreover, prime decomposition up to $\mathsf{5}$ is easily seen to be unique: One just has to show that

$$\mathsf{PA} \vdash 2^x \cdot 3^y \cdot 5^z = 2^{x'} \cdot 3^{y'} \cdot 5^{z'} \;\rightarrow\; x = x' \wedge y = y' \wedge z = z'.$$

This is usually proved by induction on $x + y + z$. Note that the simplest way to achieve this is to use the following characterisation of primality (see Exercise 9.1):

$$\mathsf{PA} \vdash \mathrm{prime}(x) \leftrightarrow \forall y \forall z (x \mid yz \rightarrow x \mid y \vee x \mid z)$$

Encoding Terms and Formulae

In a first step, every logical and every non-logical symbol ζ of Peano Arithmetic is assigned a natural number $\#\zeta$ in $\mathbb{N}$, called **Gödel number** of ζ. Since from now on, we will often switch between the meta-level and the formal level, we will always explicitly mention whenever we are reasoning formally, i.e., within PA. Otherwise the proofs will be on the meta-level.

Symbol ζ	Gödel number $\#\zeta$
$\mathsf{0}$	$\mathbf{0}$
s	$\mathbf{2}$
$+$	$\mathbf{4}$
$\cdot$	$\mathbf{6}$
$=$	$\mathbf{8}$
$\neg$	$\mathbf{10}$
$\wedge$	$\mathbf{12}$
$\vee$	$\mathbf{14}$
$\rightarrow$	$\mathbf{16}$
$\exists$	$\mathbf{18}$
$\forall$	$\mathbf{20}$
v_0	$\mathbf{1}$
v_1	$\mathbf{3}$
$\vdots$	$\vdots$
v_n	$\mathbf{2} \cdot n + \mathbf{1}$

In the previous section, we introduced power functions in PA. Since $\mathbb{N} \vDash \mathsf{PA}$, such functions also exist in $\mathbb{N}$, and we will use the same notation n^k as in PA. By $\mathbb{N}$-conformity we have $\mathsf{PA} \vdash \underline{n^k} = \underline{n}^k$ for all $n, k \in \mathbb{N}$. Note that by Theorem 1.7 it would already suffice to just gödelize the logical operators $\neg, \wedge$ and $\exists$. Next we encode terms and formulae.

Term τ	Gödel number $\#\tau$
$\mathsf{0}$	$\mathbf{0}$
v_n	$\mathbf{2} \cdot n + \mathbf{1}$
$\mathsf{s}t$	$\mathbf{2}^{\#\mathsf{s}} \cdot \mathbf{3}^{\#t}$
$t_0 + t_1$	$\mathbf{2}^{\#+} \cdot \mathbf{3}^{\#t_0} \cdot \mathbf{5}^{\#t_1}$
$t_0 \cdot t_1$	$\mathbf{2}^{\#\cdot} \cdot \mathbf{3}^{\#t_0} \cdot \mathbf{5}^{\#t_1}$

Formula φ	Gödel number $\#\varphi$
$\tau_0 = \tau_1$	$\mathbf{2}^{\#=} \cdot \mathbf{3}^{\#\tau_0} \cdot \mathbf{5}^{\#\tau_1}$
$\neg\psi$	$\mathbf{2}^{\#\neg} \cdot \mathbf{3}^{\#\psi}$
$\psi_0 \wedge \psi_1$	$\mathbf{2}^{\#\wedge} \cdot \mathbf{3}^{\#\psi_0} \cdot \mathbf{5}^{\#\psi_1}$
$\psi_0 \vee \psi_1$	$\mathbf{2}^{\#\vee} \cdot \mathbf{3}^{\#\psi_0} \cdot \mathbf{5}^{\#\psi_1}$
$\psi_0 \to \psi_1$	$\mathbf{2}^{\#\to} \cdot \mathbf{3}^{\#\psi_0} \cdot \mathbf{5}^{\#\psi_1}$
$\exists x\psi$	$\mathbf{2}^{\#\exists} \cdot \mathbf{3}^{\#x} \cdot \mathbf{5}^{\#\psi}$
$\forall x\psi$	$\mathbf{2}^{\#\forall} \cdot \mathbf{3}^{\#x} \cdot \mathbf{5}^{\#\psi}$

Observe that by the uniqueness of the prime decomposition up to $\mathsf{5}$, every natural number encodes at most one variable, term or formula. So far, we have only assigned a natural number in the standard model to each symbol, term, and formula. However, we want to do this within Peano Arithmetic. This can be achieved by stipulating

$$\ulcorner\zeta\urcorner :\equiv \underline{\#\zeta}$$

for an arbitrary symbol, term or formula ζ. This allows us to express in PA that some number is the code of a variable, term or formula. However, we can easily formalize this so-called **Gödelisation** process, where $\mathsf{2} :\equiv \mathtt{ss0}$, $\mathsf{3} :\equiv \mathtt{sss0}$, and $\mathsf{5} :\equiv \mathtt{sssss0}$.

$$\begin{aligned}
\mathrm{succ}(n) &:\equiv \mathsf{2}^{\ulcorner\mathsf{s}\urcorner} \cdot \mathsf{3}^{n} & \mathrm{add}(n,m) &:\equiv \mathsf{2}^{\ulcorner+\urcorner} \cdot \mathsf{3}^{n} \cdot \mathsf{5}^{m} \\
\mathrm{mult}(n,m) &:\equiv \mathsf{2}^{\ulcorner\cdot\urcorner} \cdot \mathsf{3}^{n} \cdot \mathsf{5}^{m} & \mathrm{eq}(t,t') &:\equiv \mathsf{2}^{\ulcorner=\urcorner} \cdot \mathsf{3}^{t} \cdot \mathsf{5}^{t'} \\
\mathrm{not}(f) &:\equiv \mathsf{2}^{\ulcorner\neg\urcorner} \cdot \mathsf{3}^{f} & \mathrm{and}(f,f') &:\equiv \mathsf{2}^{\ulcorner\wedge\urcorner} \cdot \mathsf{3}^{f} \cdot \mathsf{5}^{f'} \\
\mathrm{or}(f,f') &:\equiv \mathsf{2}^{\ulcorner\vee\urcorner} \cdot \mathsf{3}^{f} \cdot \mathsf{5}^{f'} & \mathrm{imp}(f,f') &:\equiv \mathsf{2}^{\ulcorner\to\urcorner} \cdot \mathsf{3}^{f} \cdot \mathsf{5}^{f'} \\
\mathrm{ex}(v,f) &:\equiv \mathsf{2}^{\ulcorner\exists\urcorner} \cdot \mathsf{3}^{v} \cdot \mathsf{5}^{f} & \mathrm{all}(v,f) &:\equiv \mathsf{2}^{\ulcorner\forall\urcorner} \cdot \mathsf{3}^{v} \cdot \mathsf{5}^{f}
\end{aligned}$$

In order to simplify the notation, for terms $\tau, \tau_0, \ldots, \tau_n$, we define

$$\tau \in \{\tau_0, \ldots, \tau_n\} \;:\equiv\; \bigvee_{i=0}^{n} \tau = \tau_i \,.$$

Now we are ready to provide a formalised version of construction of terms and formulae:

$$\mathrm{var}(v) :\Longleftrightarrow \exists n(v = \mathsf{2} \cdot n + \mathsf{1})$$

$$\begin{aligned}
\mathrm{c_term}(c,t) :\Longleftrightarrow\ & \mathrm{seq}(c) \wedge c_{\mathrm{lh}(c)-1} = t \wedge \\
& \forall k < \mathrm{lh}(c)\Big(\mathrm{var}(c_k) \vee c_k = \mathsf{0} \vee \\
& \quad \exists i < k\, \exists j < k\big(c_k \in \big\{\mathrm{succ}(c_i), \mathrm{add}(c_i,c_j), \mathrm{mult}(c_i,c_j)\big\}\big)\Big)
\end{aligned}$$

$$\mathrm{term}(t) :\Longleftrightarrow \exists c\big(\mathrm{c\_term}(c,t)\big)$$

$$\begin{aligned}\mathrm{c\_fml}(c,f) :\Longleftrightarrow\ & \mathrm{seq}(c) \wedge c_{\mathrm{lh}(c)-1} = f \wedge \\ & \forall k < \mathrm{lh}(c)\Big(\exists t\,\exists t'\big(\mathrm{term}(t) \wedge \mathrm{term}(t') \wedge c_k = \mathrm{eq}(t,t')\big) \vee \\ & \quad \exists i,j < k\big(c_k \in \big\{\mathrm{not}(c_i), \mathrm{and}(c_i,c_j), \mathrm{or}(c_i,c_j), \mathrm{imp}(c_i,c_j)\big\}\big) \vee \\ & \quad\quad \exists i < k\,\exists v\big(\mathrm{var}(v) \wedge c_k \in \big\{\mathrm{ex}(v,c_i), \mathrm{all}(v,c_i)\big\}\big)\Big)\end{aligned}$$

$$\mathrm{fml}(f) :\Longleftrightarrow \exists c\big(\mathrm{c\_fml}(c,f)\big)$$

Note that all the above relations are defined by $\exists$-formulae.

Example 9.12. Let us consider the term $\tau \equiv \mathtt{s}v_n + \mathtt{0}$. In the standard model $\mathbb{N}$, the sequence $c \equiv \langle \# v_n, \# \mathtt{s}v_n, \# \mathtt{0}, \# \mathtt{s}v_n + \mathtt{0}\rangle$ encodes τ, i.e., $\mathbb{N} \vDash \mathrm{c\_term}(c, \# \tau)$. By Proposition 9.2 this implies $\mathsf{PA} \vdash \mathrm{c\_term}(\underline{c}, \ulcorner \tau \urcorner)$.

Lemma 9.13. *For $n \in \mathbb{N}$ we have*

(a) $\mathbb{N} \vDash \mathrm{var}(n)$ *if and only if* $n \equiv \# \nu$ *for some variable* ν.

(b) $\mathbb{N} \vDash \mathrm{term}(n)$ *if and only if* $n \equiv \# \tau$ *for some* $\mathscr{L}_{\mathsf{PA}}$*-term* τ.

(c) $\mathbb{N} \vDash \mathrm{fml}(n)$ *if and only if* $n \equiv \# \varphi$ *for some* $\mathscr{L}_{\mathsf{PA}}$*-formula* φ.

Proof. Condition (a) is obvious. For (b), we first prove that $\mathbb{N} \vDash \mathrm{term}(\# \tau)$ for every term τ. We proceed by induction on the term construction of τ. If $\tau \equiv \mathtt{0}$ or τ is a variable, then clearly $\mathbb{N} \vDash \mathrm{c\_term}(\langle \# \tau \rangle, \# \tau)$ and hence the claim follows. Now, if $\tau \equiv \mathtt{s}\tau'$ for some term τ' with $\mathbb{N} \vDash \mathrm{term}(\# \tau')$, and $c \in \mathbb{N}$ is a code with $\mathbb{N} \vDash \mathrm{c\_term}(c, \# \tau')$, then $\mathbb{N} \vDash \mathrm{succ}(\# \tau')$, and hence, $\mathbb{N} \vDash \mathrm{c\_term}(c^\frown \langle \# \tau \rangle, \# \tau)$. The other cases are similar. For the converse, we use the principle of strong induction in $\mathbb{N}$. Suppose that the claim holds for all $m < n$ in $\mathbb{N}$ and let $\mathbb{N} \vDash \mathrm{term}(n)$. If $n \equiv \mathbf{0}$ then $n \equiv \# \mathtt{0}$, and if $n \equiv 2m + 1$ for some m, then $n \equiv \# v_m$. Let $\mathbb{N} \vDash \mathrm{c\_term}(c, n)$ for some $c \in \mathbb{N}$ with $\mathrm{lh}(c) > 1$. Now in $\mathbb{N}$ we have either $n \equiv \mathrm{succ}^{\mathbb{N}}(c_i)$ for some $i < \mathrm{lh}(c)$, $n \equiv \mathrm{add}^{\mathbb{N}}(c_i, c_j)$ or $n \equiv \mathrm{mult}^{\mathbb{N}}(c_i, c_j)$ for $i, j < \mathrm{lh}(c)$. In the first case, note that $\mathbb{N} \vDash \mathrm{c\_term}(\langle c_k \mid k < \mathbf{s}i \rangle, c_i)$. By our induction hypothesis, we can take a term τ such that $c_i \equiv \# \tau$, and by But then, by $\mathbb{N}$-conformity we have $n \equiv \mathrm{succ}^{\mathbb{N}}(c_i) \equiv (2^{\ulcorner \mathtt{s} \urcorner} \cdot 3^{\underline{c_i}})^{\mathbb{N}} \equiv 2^{\# \mathtt{s}} \cdot 3^{c_i} \equiv \# \mathtt{s}\tau$, and hence, n encodes $\mathtt{s}\tau$. The other cases are similar.

The corresponding statement for formulae is proved in the same way and is therefore left as an exercise. ⊣

Note that the relations var, term and formula are $\exists$-formulae. Therefore, by combining Lemma 9.13 and Proposition 9.2 we obtain: If ν is a variable, τ is a term and φ is a formula, then

$$\mathsf{PA} \vdash \mathrm{var}(\ulcorner \nu \urcorner), \quad \mathsf{PA} \vdash \mathrm{term}(\ulcorner \tau \urcorner), \quad \text{and} \quad \mathsf{PA} \vdash \mathrm{formula}(\ulcorner \varphi \urcorner).$$

Before we proceed to gödelise logical axioms, the axioms of Peano Arithmetic and formal proofs, we have to deal with substitution: First, we introduce new relations which check wether a code for a variable appears in the code of a term or formula, respectively.

$$\begin{aligned}\mathrm{var\_in\_term}(v,t) \;:\Longleftrightarrow\; &\mathrm{var}(v) \wedge \exists c\Big(\mathrm{c\_term}(c,t) \wedge \\ &\forall c' < c \neg\, \mathrm{c\_term}(c',t) \wedge \exists i < \mathrm{lh}(c)(c_i = v)\Big)\end{aligned}$$

Note that the minimality of c is necessary since otherwise, any code for a variable could appear in the sequence of codes of the term construction. The same holds for the following relation var_in_fml:

$$\begin{aligned}\mathrm{var\_in\_fml}(v,f) :\Longleftrightarrow\; &\exists c\Big(\mathrm{c\_fml}(c,f) \wedge \forall c' < c \neg\, \mathrm{c\_fml}(c',f) \wedge \\ &\exists i < \mathrm{lh}(c)\, \exists t_0\, \exists t_1\big(\mathrm{term}(t_0) \wedge \mathrm{term}(t_1) \wedge c_i = \mathrm{eq}(t_0,t_1) \wedge \\ &\big(\mathrm{var\_in\_term}(v,t_0) \vee \mathrm{var\_in\_term}(v,t_1)\big)\big)\Big)\end{aligned}$$

$$\begin{aligned}\mathrm{free}(v,f) :\Longleftrightarrow\; &\exists c\Big(\mathrm{c\_fml}(c,f) \wedge \mathrm{var\_in\_fml}(v,f) \wedge \\ &\forall i < \mathrm{lh}(c)\, \forall j < i\big(c_i \neq \mathrm{ex}(v,c_j) \wedge c_i \neq \mathrm{all}(v,c_j)\big)\Big)\end{aligned}$$

For the sake of simplicity, we permit the substitution $\varphi(x/\tau)$ only if it is admissible and x as well as all variables in τ appear only free in φ. This does not impose a restriction, since by renaming of variables, every formula is equivalent to one in which no variable occurs both bound and free. We can thus define

$$\begin{aligned}\mathrm{sb\_adm}(v,t,f) :\Longleftrightarrow\; &\mathrm{var\_in\_fml}(v,f) \wedge \mathrm{free}(v,f) \wedge \\ &\forall v' < t\big(\mathrm{var\_in\_term}(v',t) \to \mathrm{free}(v',f)\big).\end{aligned}$$

Note that the relations var_in_term, var_in_fml, free, and sb_adm are all $\exists$-formulae: The only unbounded universal quantifier appears in the relation sb_adm, where var_in_term occurs as negated — recall that $\mathrm{var\_in\_term}(v',t) \to \mathrm{free}(v',f)$ is equivalent to $\neg\, \mathrm{var\_in\_term}(v',t) \vee \mathrm{free}(v',f)$. However, the existential quantifier in the definition of $\mathrm{var\_in\_term}(v',t)$ can be replaced by a bounded one, since the code of t has to be smaller than the code of f.

The next relation expresses that c' encodes the construction of the term obtained from the term with code t by replacing every occurrence of the code v of a variable by the code t_0.

$$\begin{aligned}\mathrm{c\_sb\_term}(c,c',c'',v,t_0,t,t') :\Longleftrightarrow\; &\mathrm{var}(v) \wedge \mathrm{c\_term}(c,t) \wedge \mathrm{c\_term}(c',t') \wedge \\ &\mathrm{c\_term}(c'',t_0) \;\wedge \mathrm{lh}(c') = \mathrm{lh}(c'') + \mathrm{lh}(c)\,,\end{aligned}$$

and for all $k < \mathrm{lh}(c)$ we have

$$(\mathrm{var}(c_k) \ \wedge \ c_k \neq v \to c'_{\mathrm{lh}(c'')+k} = c_k) \wedge (c_k = \mathsf{0} \to c'_{\mathrm{lh}(c'')+k} = \mathsf{0}) \\ \wedge (c_k = v \to c'_{\mathrm{lh}(c'')+k} = t_0)$$

and

$$\forall i < k \, \forall j < k \, \big(\ c_k = \mathrm{succ}(c_i) \quad \to c'_{\mathrm{lh}(c'')+k} = \mathrm{succ}(c'_{\mathrm{lh}(c'')+i}) \quad \wedge \\ c_k = \mathrm{add}(c_i, c_j) \ \to c_{\mathrm{lh}(c'')+k} = \mathrm{add}(c'_{\mathrm{lh}(c'')+i}, c'_{\mathrm{lh}(c'')+j}) \ \wedge \\ c_k = \mathrm{mult}(c_i, c_j) \to c_{\mathrm{lh}(c'')+k} = \mathrm{mult}(c'_{\mathrm{lh}(c'')+i}, c'_{\mathrm{lh}(c'')+j}) \quad \big) \, .$$

By omitting the codes c, c', c'' we can describe term substitution by

$$\mathrm{sb\_term}(v, t_0, t, t') :\Longleftrightarrow \exists c \, \exists c' \, \exists c'' \big(\mathrm{c\_sb\_term}(c, c', c'', v, t_0, t, t') \big).$$

Informally, if t encodes the term τ, v the variable ν, t_0 the term τ_0, and t' encodes $\tau(\nu/\tau_0)$, then the relation $\mathrm{sb\_term}(v, t_0, t, t')$ holds.

For formulae, we proceed similarly, except that we first have to make sure that the substitution is admissible.

$\mathrm{c\_sb\_fml}(c, c', v, t_0, f, f') \ :\Longleftrightarrow$

$\mathrm{c\_fml}(c, f) \wedge \mathrm{c\_fml}(c', f') \wedge \mathrm{sb\_adm}(v, t_0, f) \wedge \mathrm{lh}(c') = \mathrm{lh}(c)$, and for all $k < \mathrm{lh}(c)$ we have:

$$\forall t \, \forall t' \, \forall s \, \forall s' \, \Big(\big(c_k = \mathrm{eq}(t, t') \wedge \mathrm{sb\_term}(v, t_0, t, s) \wedge \\ \mathrm{sb\_term}(v, t_0, t', s') \big) \ \to c'_k = \mathrm{eq}(s, s') \Big)$$

and

$$\forall i < k \, \forall j < k \, \big(\ c_k = \mathrm{not}(c_i) \quad \to c'_k = \mathrm{not}(c'_i) \quad \wedge \\ c_k = \mathrm{or}(c_i, c_j) \quad \to c'_k = \mathrm{or}(c'_i, c'_j) \quad \wedge \\ c_k = \mathrm{and}(c_i, c_j) \ \to c'_k = \mathrm{and}(c'_i, c'_j) \ \wedge \\ c_k = \mathrm{imp}(c_i, c_j) \ \to c'_k = \mathrm{imp}(c'_i, c'_j) \quad \big)$$

and

$$\forall i < k \, \forall v \, \big(\ c_k = \mathrm{all}(v, c_i) \to c'_k = \mathrm{ex}(v, c'_i) \wedge \\ c_k = \mathrm{ex}(v, c_i) \ \to c'_k = \mathrm{ex}(v, c'_i) \quad \big)$$

Again, by leaving out the sequence codes, we define

$$\mathrm{sb\_fml}(v, t_0, f, f') :\Longleftrightarrow \exists c \exists c' \big(\mathrm{c\_sb\_fml}(c, c', v, t_0, f, f') \big) \, .$$

Informally, if f encodes the formula φ, v the variable ν, t_0 the term τ, and if the substitution $\varphi(\nu/\tau)$ is admissible and f' encodes $\varphi(\tau)$, then the relation $\mathrm{sb\_fml}(v, t_0, f, f')$ holds.

LEMMA 9.14. *Let τ and τ_0 be two terms, φ a formula and ν a variable such that the substitution $\varphi(\nu/\tau_0)$ is admissible. Then we have:*

(a) $\mathsf{PA} \vdash \mathrm{sb\_term}(\ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\tau\urcorner, t) \leftrightarrow t = \ulcorner\tau(\nu/\tau_0)\urcorner$

(b) $\mathsf{PA} \vdash \mathrm{sb\_fml}(\ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\varphi\urcorner, f) \leftrightarrow f = \ulcorner\varphi(\nu/\tau_0)\urcorner$

Proof. This follows from the definition of the relations sb_term and sb_fml using induction on the term construction of τ and the formula construction of φ. ⊣

EXAMPLE 9.15. Let $\tau \equiv \mathtt{s}x{+}y$ and $\tau_0 \equiv \mathtt{0}$. Then the sequence $\langle \ulcorner x\urcorner, \ulcorner \mathtt{s}x\urcorner, \ulcorner y\urcorner, \ulcorner \mathtt{s}x + y\urcorner\rangle$ encodes $\ulcorner\tau\urcorner$ and $\langle\ulcorner\mathtt{0}\urcorner\rangle$ encodes τ_0. Now, when coding $\tau(x/\tau_0)$, we first take the code of τ_0 and then replace every occurrence of x in the code of τ by τ_0. This gives

$$\langle \ulcorner\mathtt{0}\urcorner, \ulcorner\mathtt{s0}\urcorner, \ulcorner y\urcorner, \ulcorner \mathtt{s0} + y\urcorner\rangle$$

which encodes $\tau(x/\tau_0)$.

Encoding Formal Proofs

Finally, we can use the machinery as developed in the previous section to encode axioms and formal proofs. We first show how to achieve this in $\mathbb{N}$. For this, recall that a formal proof of some $\mathscr{L}_{\mathsf{PA}}$-formula φ is a FINITE sequence $\varphi_0, \ldots, \varphi_n$ with $\varphi_n \equiv \varphi$ such that each φ_i is an instance of a logical axiom, an axiom of PA or is obtained from preceding elements of the sequence by using Modus Ponens or Generalisation. Hence we can code a formal proof of φ by $\langle \#\varphi_0, \ldots, \#\varphi_n\rangle$. Conversely, from such a code we can recover the sequence $\varphi_0, \ldots, \varphi_n$ and hence reconstruct a formal proof of φ.

As for terms and formulae, we proceed to code formal proofs in PA. The goal is to define a relation prv with the property that $\mathbb{N} \vDash \mathrm{prv}(\ulcorner\varphi\urcorner)$ for some formula φ if and only if there is a formal proof of φ, i.e., $\mathsf{PA} \vdash \varphi$. The following examples illustrate how axioms can be formalised in PA:

$$\mathrm{ax\_L_1}(f) :\Longleftrightarrow \exists f'\exists f''\big(\mathrm{fml}(f') \wedge \mathrm{fml}(f'') \wedge f = \mathrm{imp}(f', \mathrm{imp}(f'', f'))\big)$$

$$\mathrm{ax\_L_{10}}(f) :\Longleftrightarrow \exists f'\exists f''\exists v\exists t\big(\mathrm{sb\_fml}(v, t, f', f'') \wedge f = \mathrm{imp}(\mathrm{all}(v, f'), f'')\big)$$

$$\mathrm{ax\_L_{14}}(f) :\Longleftrightarrow \exists t\big(\mathrm{term}(t) \wedge f = \mathrm{eq}(t, t)\big)$$

$\mathsf{PA_0}$ and the Induction Schema are gödelised as follows:

$$\mathrm{ax\_PA_0}(f) :\Longleftrightarrow f = \ulcorner\forall v_0\neg(\mathtt{s}v_0 = \mathtt{0})\urcorner$$

$$\begin{aligned}\mathrm{ax\_PA_6}(f) :\Longleftrightarrow \exists f'\exists f''\exists f'''\exists v\exists g\big(&\mathrm{free}(v,f')\wedge \mathrm{sb\_fml}(v,\ulcorner 0\urcorner,f',f'')\\ &\wedge \mathrm{sb\_fml}(v,\mathrm{succ}(v),f',f''')\wedge g=\mathrm{all}(v,\mathrm{imp}(f',f'''))\\ &\wedge f=\mathrm{imp}(\mathrm{and}(f'',g),\mathrm{all}(v,f'))\big)\end{aligned}$$

We leave it to the reader to formalise the other axioms. Similarly, we define axioms:

$$\begin{aligned}\mathrm{log\_ax}(f) &:\Longleftrightarrow \mathrm{ax\_L_0}(f)\vee\cdots\vee\mathrm{ax\_L_{16}}(f)\\ \mathrm{peano\_ax}(f) &:\Longleftrightarrow \mathrm{ax\_PA_0}(f)\vee\cdots\vee\mathrm{ax\_PA_6}(f)\\ \mathrm{ax}(f) &:\Longleftrightarrow \mathrm{log\_ax}(f)\vee\mathrm{peano\_ax}(f)\end{aligned}$$

Next, we formalise the inference rules Modus Ponens and Generalisation:

$$\begin{aligned}\mathrm{mp}(f',f'',f) &:\Longleftrightarrow \mathrm{fml}(f')\wedge\mathrm{fml}(f)\wedge f''=\mathrm{imp}(f',f)\\ \mathrm{gen}(v,f',f) &:\Longleftrightarrow \mathrm{var}(v)\wedge\mathrm{fml}(f')\wedge f=\mathrm{all}(v,f')\end{aligned}$$

Finally, we encode formal proofs as sequences of codes of formulae which are either axioms or produced by one of the inference rules. Therefore, we define the predicates $\mathrm{c\_prv}(c,f)$ in order to specify that c encodes a proof of the formula coded by f and prv to express provability.

$$\begin{aligned}\mathrm{c\_prv}(c,f) :\Longleftrightarrow\ &\mathrm{seq}(c)\wedge c_{\mathrm{lh}(c)-1}=f\wedge\forall k<\mathrm{lh}(c)\big(\mathrm{ax}(c_k)\vee\\ &\qquad\exists i<k\exists j<k\big(\mathrm{mp}(c_i,c_j,c_k)\vee\exists v(\mathrm{gen}(v,c_i,c_k))\big)\big)\\ \mathrm{prv}(f) :\Longleftrightarrow\ &\exists c(\mathrm{c\_prv}(c,f)).\end{aligned}$$

Note that in the standard model $\mathbb{N}$ we have $\mathbb{N}\vDash \mathrm{c\_prv}(c,\#\varphi)$ if and only if c encodes a sequence $\langle\#\varphi_0,\ldots,\#\varphi_n\rangle$, where $\varphi_0,\ldots,\varphi_n$ is a formal proof of φ.

Lemma 9.16. *Let $n,c\in\mathbb{N}$ be natural numbers. Then $\mathbb{N}\vDash\mathrm{c\_prv}(c,n)$ if and only if c encodes a formal proof of some $\mathscr{L}_{\mathsf{PA}}$-formula φ with $\#\varphi=n$. In particular, $\mathbb{N}\vDash\mathrm{prv}(\#\varphi)$ if and only if $\mathsf{PA}\vdash\varphi$.*

Proof. Note first that $\mathbb{N}\vDash\mathrm{ax}(m)$ for some $m\in\mathbb{N}$ if and only if $m=\#\psi$ encodes an instance of a logical axiom or an axiom of PA. The proof is the same as the proof of Lemma 9.13, where in the forward direction, one proceeds by induction on $\mathrm{lh}(c)$, and for the converse by induction on the length of the formal proof of φ.

For the second part, suppose that $\mathbb{N}\vDash\mathrm{prv}(\#\varphi)$. Then there is $c\in\mathbb{N}$ which encodes a formal proof of φ, and hence, $\mathsf{PA}\vdash\varphi$. ⊣

Note that Lemma 9.16 does not hold if we replace $\mathbb{N}$ by a non-standard model $\mathbf{M}$ of PA: If $\mathbf{M}\vDash\mathrm{c\_prv}(c,\ulcorner\varphi\urcorner)$ and $c^{\mathbf{M}}$ is a non-standard number,

then the "proof" encoded by $c^{\mathbf{M}}$ is of non-standard length and therefore, we cannot conclude that $\mathsf{PA} \vdash \varphi$.

COROLLARY 9.17. *Let φ be an $\mathscr{L}_{\mathsf{PA}}$-formula. If $\mathsf{PA} \vdash \varphi$ then there is $n \in \mathbb{N}$ such that $\mathsf{PA} \vdash \mathrm{c\_prv}(\underline{n}, \ulcorner\varphi\urcorner)$. In particular, if $\mathsf{PA} \vdash \varphi$ then $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner\varphi\urcorner)$.*

Proof. Suppose that $\mathsf{PA} \vdash \varphi$. Then by LEMMA 9.16 we have $\mathbb{N} \vDash \mathrm{c\_prv}(c, \#\varphi)$ for some $c \in \mathbb{N}$. Observe that the relations c_prv and prv are defined by an $\exists$-formula. Hence, it follows from PROPOSITION 9.2 that $\mathsf{PA} \vdash \mathrm{c\_prv}(\underline{c}, \ulcorner\varphi\urcorner)$, and in particular, $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner\varphi\urcorner)$. ⊣

NOTES

In his proof of the incompleteness theorems [16], Gödel used the β-function and unique prime decomposition in order to encode finite sequences by a single number. There are, however, other ways to achieve this (e.g., the coding provided by Smullyan in [53]). In our presentation, we mainly followed Shoenfield [48].

EXERCISES

9.0 Prove the statement N_2.

9.1 (a) Show that $\mathsf{PA} \vdash \forall x \geq 2\, \exists y \big(\mathrm{prime}(y) \wedge y \mid x\big)$, i.e., every $x \geq 2$ has a prime divisor.

(b) Show that $\mathsf{PA} \vdash \forall x \geq 2 \big(\mathrm{prime}(x) \leftrightarrow \forall y \forall z (x \mid yz \rightarrow x \mid y \vee x \mid z)\big)$.

Hint: Use BÉZOUT'S LEMMA (see EXERCISE 8.2).

9.2 Introduce a factorial function "!" such that $n! = 1 \cdot \ldots \cdot n$, and show that it is $\mathbb{N}$-conform.

9.3 Introduce a function $\mathrm{lcm}_{i<k} F(i)$ for a function F definable in PA such that $\mathrm{lcm}_{i<k} F(i)$ is the least common multiple of $F(\mathsf{0}), \ldots, F(k - \mathsf{1})$ and show that it is $\mathbb{N}$-conform.

9.4 Let $\mathbf{M}$ be an arbitrary non-standard model of PA. Prove the following statements:

(a) $\mathbf{M}$ contains a non-standard prime number.

(b) $\mathbf{M}$ contains a number that is divisible by every standard prime number.

Hint: Use EXERCISE 7.4.

9.5 State and prove in PA that every number has a unique prime decomposition, i.e., prove that every number is a product of primes, and show that this is unique up to permutations of the factors.

9.6 (a) Show that the encoding of terms and formulae by Gödel numbers is one-to-one, i.e., show that there are no two valid terms or formulae with the same Gödel number.

Remark: The Gödel number $\mathbf{4}$ might stand for $+$ as well as for $\mathtt{s0}$, but $+$ is neither a valid term nor a valid formula.

(b) Give the sequence encoding the construction of the formula $\forall v_0 \neg = \mathsf{0}\, v_0$, which corresponds to $\forall v_0 (\mathsf{0} \neq v_0)$ in "infix notation".

9.7 An alternative way to utilise the uniqueness of the prime number decomposition for Gödel coding is to use the existence and uniqueness of base b notation for any $b \geq 2$.

(a) Show that it suffices to gödelise only b symbols for some $b \in \mathbb{N}$.

(b) Prove that for every number n there are numbers $n_0, \ldots, n_k$ with $0 \leq n_i < b$ for all $0 \leq i \leq k$ such that

$$n = n_k b^k + \ldots + n_1 b + n_0 \equiv: (n_k, \ldots, n_0)_b.$$

(c) Show that there is a function $*$ definable in PA such that

$$(n_k, \ldots, n_0)_b * (m_l, \ldots, m_0)_b = (n_k, \ldots, n_0, m_l, \ldots, m_0)_b \,.$$

(d) Use (a) and (b) to give an alternative of Gödel coding using base b notation rather than the unique prime decomposition.

9.8 Show that there is a function which truncates sequences, i.e., introduce a binary function $\restriction$ such that

$$\mathsf{PA} \vdash \big(\mathrm{seq}(s) \wedge k < \mathrm{lh}(s)\big) \to \big(\mathrm{seq}(s \restriction k) \wedge \mathrm{lh}(s \restriction k) = k \wedge \forall i < k((s \restriction k)_i = s_i)\big) \,.$$

9.9 Complete the proof of Lemma 9.14.

Chapter 10
The First Incompleteness Theorem

In 1931, Gödel proved his First Incompleteness Theorem which states that if PA is consistent, then it is incomplete, i.e., there is a $\mathscr{L}_{\mathsf{PA}}$-sentence σ such that $\mathsf{PA} \nvdash \sigma$ and $\mathsf{PA} \nvdash \neg\sigma$. In this chapter, we prove the First Incompleteness Theorem not only for PA but also for weaker and stronger theories.

The Provability Predicate

In this section, we state some properties of the provability predicate that we introduced in Chapter 9.

Lemma 10.0.

(a) $\mathsf{PA} \vdash \mathrm{prv}(x) \wedge \mathrm{prv}\big(\mathrm{imp}(x,y)\big) \to \mathrm{prv}(y)$

(b) $\mathsf{PA} \vdash \mathrm{prv}(x) \wedge \mathrm{prv}(y) \to \mathrm{prv}\big(\mathrm{and}(x,y)\big)$.

Proof. For (a), note that $\mathrm{prv}(x)$ and $\mathrm{prv}(\mathrm{imp}(x,y))$ imply $\mathrm{mp}(x,\mathrm{imp}(x,y),y)$. Now, if c, c' satisfy $\mathrm{c\_prv}(c,x)$ and $\mathrm{c\_prv}(c',\mathrm{imp}(x,y))$, respectively, then the concatenation of the codes yields $\mathrm{c\_prv}(c^\frown c'^\frown\langle y\rangle, y)$ and hence $\mathrm{prv}(y)$ as desired.

For (b), assume $\mathrm{prv}(x)$ and $\mathrm{prv}(y)$. In particular, this implies $\mathrm{fml}(x)$ and $\mathrm{fml}(y)$. Note that using the formalised version of the axiom L_5, we obtain

$$\mathsf{PA} \vdash \mathrm{prv}\Big(\mathrm{imp}\big(y, \mathrm{imp}\big(x, \mathrm{and}(x,y)\big)\big)\Big).$$

Using $\mathrm{prv}(y)$ and (a), we get $\mathrm{prv}\big(\mathrm{imp}\big(x, \mathrm{and}(x,y)\big)\big)$, and a further application of (a) yields $\mathrm{prv}(\mathrm{and}(x,y))$. ⊣

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_11

As an immediate consequence of Lemma 10.0, we obtain the following

Corollary 10.1. *Let φ and ψ be $\mathscr{L}_{\mathsf{PA}}$-formulae. Then we have*

(a) $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \varphi \to \psi \urcorner) \to (\mathrm{prv}(\ulcorner \varphi \urcorner) \to \mathrm{prv}(\ulcorner \psi \urcorner))$,
(b) $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \varphi \urcorner) \wedge \mathrm{prv}(\ulcorner \psi \urcorner) \to \mathrm{prv}(\ulcorner \varphi \wedge \psi \urcorner)$.

Note that (a) corresponds to a formalised version of the inference rule (MP).

Corollary 10.2. *Let φ and ψ be $\mathscr{L}_{\mathsf{PA}}$-formulae. Then the following statements hold:*

(a) *If $\varphi \Leftrightarrow_{\mathsf{PA}} \psi$, then $\mathrm{prv}(\ulcorner \varphi \urcorner) \Leftrightarrow_{\mathsf{PA}} \mathrm{prv}(\ulcorner \psi \urcorner)$.*
(b) $\mathrm{prv}(\ulcorner \varphi \urcorner) \wedge \mathrm{prv}(\ulcorner \psi \urcorner) \Leftrightarrow_{\mathsf{PA}} \mathrm{prv}(\ulcorner \varphi \wedge \psi \urcorner)$.

Proof. For (a), assume that $\varphi \Leftrightarrow_{\mathsf{PA}} \psi$. By symmetry, it suffices to verify that $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \varphi \urcorner) \to \mathrm{prv}(\ulcorner \psi \urcorner)$. Since $\mathsf{PA} \vdash \varphi \to \psi$, Corollary 9.17 yields $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \varphi \to \psi \urcorner)$. The assertion then follows from Corollary 10.1 using Modus Ponens.

For (b), note that by Corollary 10.1.(b), it suffices to prove $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \varphi \wedge \psi \urcorner) \to \mathrm{prv}(\ulcorner \varphi \urcorner) \wedge \mathrm{prv}(\ulcorner \psi \urcorner)$. But this is a direct consequence of Corollary 10.1.(a) using L_3 and L_4. ⊣

The Diagonalisation Lemma

We already know that every standard natural number is either $\mathbf{0}$ or the successor of a standard natural number. Hence, we can introduce a binary relation which states that x is the code of a standard natural number:

$$\mathrm{c\_nat}(c, n, x) :\Longleftrightarrow \mathrm{seq}(c) \wedge \mathrm{lh}(c) = \mathtt{s}n \wedge c_0 = \ulcorner \mathtt{0} \urcorner \wedge \forall i < n \left(c_{\mathtt{s}i} = \mathrm{succ}(c_i) \wedge c_n = x \right)$$

$$\mathrm{nat}(n, x) :\Longleftrightarrow \exists c \left(\mathrm{c\_nat}(c, n, x) \right)$$

Clearly, it follows from the definition that

$$\mathsf{PA} \vdash \mathrm{c\_nat}(c, n, x) \to \mathrm{c\_nat}\big(c^\frown \langle \mathrm{succ}(x) \rangle, \mathtt{s}n, \mathrm{succ}(x)\big) .$$

Lemma 10.3. *For any natural number $n \in \mathbb{N}$ we have $\mathsf{PA} \vdash \mathrm{nat}(\underline{n}, \ulcorner \underline{n} \urcorner)$. In particular, if φ is an $\mathscr{L}_{\mathsf{PA}}$-formula, then $\mathsf{PA} \vdash \mathrm{nat}(\ulcorner \varphi \urcorner, \ulcorner \ulcorner \varphi \urcorner \urcorner)$.*

Proof. We proceed by metainduction on n. For $n \equiv \mathbf{0}$, the term $\underline{\mathbf{0}}$ is the same as $\mathtt{0}$, and clearly, the code c of the singleton sequence $\langle \ulcorner \mathtt{0} \urcorner \rangle$ witnesses $\mathrm{c\_nat}(c, \mathtt{0}, \ulcorner \mathtt{0} \urcorner)$. Now, suppose that for some c and $n \in \mathbb{N}$ we have $\mathrm{c\_nat}(c, \underline{n}, \ulcorner \underline{n} \urcorner)$. Let c'' be the code for $\langle \ulcorner \underline{\mathtt{s}n} \urcorner \rangle$ and let $c' := c^\frown c''$. Notice

that since $\mathrm{lh}(c) = \mathtt{s}\underline{n}$, we have $\mathrm{lh}(c') = \mathtt{ss}\underline{n}$, and by definition of c' we have $(c')_{\mathtt{s}\underline{n}} = \ulcorner \mathtt{s}\underline{n} \urcorner = \mathrm{succ}(\ulcorner \underline{n} \urcorner)$. Using the induction hypothesis and the above observation, we obtain $\mathrm{c\_nat}(c', \mathtt{s}\underline{n}, \ulcorner \mathtt{s}\underline{n} \urcorner)$ as desired. ⊣

Finally, we define the Gödel number of a number by stipulating

$$\mathrm{gn}(n) = x :\Longleftrightarrow \mathrm{nat}(n, x) \vee \big(\neg\exists y(\mathrm{nat}(n, y)) \wedge x = \mathtt{0}\big).$$

This indeed defines a function, since using the definition of the predicate seq, one can easily show that

$$\mathsf{PA} \vdash \mathrm{nat}(n, x) \wedge \mathrm{nat}(n, y) \to x = y\,.$$

In particular, by Lemma 10.3 we have

$$\mathsf{PA} \vdash \mathrm{gn}(\ulcorner \varphi \urcorner) = \ulcorner\ulcorner \varphi \urcorner\urcorner\,. \tag{$*$}$$

We have now assembled all the ingredients to prove the Diagonalisation Lemma, which will be crucial in the proof of the First Incompleteness Theorem.

Theorem 10.4 (Diagonalisation Lemma). *Let $\varphi(\nu)$ be an $\mathscr{L}_{\mathsf{PA}}$-formula with one free variable ν which does not occur bound in φ. Then there exists an $\mathscr{L}_{\mathsf{PA}}$-sentence σ_φ such that*

$$\mathsf{PA} \vdash \sigma_\varphi \leftrightarrow \varphi(\nu/\ulcorner \sigma_\varphi \urcorner)\,, \quad \textit{i.e.,}\ \ \sigma_\varphi \Leftrightarrow_{\mathsf{PA}} \varphi(\ulcorner \sigma_\varphi \urcorner).$$

Proof. We define

$$\psi(v_0) :\equiv \forall v_1 \Big(\mathrm{sb\_fml}\big(\ulcorner v_0 \urcorner, \mathrm{gn}(v_0), v_0, v_1\big) \to \varphi(\nu/v_1)\Big)$$

and

$$\sigma_\varphi :\equiv \psi(v_0/\ulcorner \psi \urcorner).$$

In other words, $\sigma_\varphi \equiv \psi(\ulcorner \psi \urcorner)$ and $\ulcorner \sigma_\varphi \urcorner = \ulcorner \psi(\ulcorner \psi \urcorner) \urcorner$. Since $\ulcorner v_0 \urcorner = \mathtt{s0}$, we have

$$\begin{aligned}
\sigma_\varphi &\equiv \forall v_1 \big(\mathrm{sb_fml}(\mathtt{s0}, \mathrm{gn}(\ulcorner \psi \urcorner), \ulcorner \psi(v_0) \urcorner, v_1) \to \varphi(\nu/v_1)\big) \\
&\Leftrightarrow_{\mathsf{PA}} \forall v_1 \big(\mathrm{sb_fml}(\mathtt{s0}, \ulcorner\ulcorner \psi \urcorner\urcorner, \ulcorner \psi(v_0) \urcorner, v_1) \to \varphi(v_1)\big) \\
&\Leftrightarrow_{\mathsf{PA}} \forall v_1 \big(v_1 = \ulcorner \psi(v_0/\ulcorner \psi \urcorner) \urcorner \to \varphi(v_1)\big) \\
&\Leftrightarrow_{\mathsf{PA}} \varphi(\ulcorner \psi(v_0/\ulcorner \psi \urcorner) \urcorner) \\
&\equiv \varphi(\ulcorner \psi(\ulcorner \psi \urcorner) \urcorner) \\
&\equiv \varphi(\ulcorner \sigma_\varphi \urcorner),
\end{aligned}$$

where the first equivalence follows from $(*)$, the second equivalence follows from Lemma 9.14, and the third equivalence follows from L_{10} and L_{14}. ⊣

The Diagonalisation Lemma is often called Fixpoint Lemma, since the sentence σ_φ can be conceived as a fixed point of φ. It is a powerful tool, since it allows us to make self-referential statements, i.e., for a formula φ with one free variable it provides a sentence σ_φ which states "I have the property φ".

The First Incompleteness Theorem

Now we are ready to prove the First Incompleteness Theorem:

Theorem 10.5 (First Incompleteness Theorem for PA). PA *is incomplete.*

Proof. By the Diagonalisation Lemma there is an $\mathscr{L}_{\mathsf{PA}}$-sentence σ such that

$$\sigma \Leftrightarrow_{\mathsf{PA}} \neg \operatorname{prv}(\ulcorner \sigma \urcorner).$$

To see this, let $\varphi(v_0) :\equiv \neg \operatorname{prv}(v_0)$. Then $\sigma_\varphi \Leftrightarrow_{\mathsf{PA}} \varphi(\ulcorner \sigma_\varphi \urcorner)$ and

$$\varphi(\ulcorner \sigma_\varphi \urcorner) \equiv \varphi(v_0/\ulcorner \sigma_\varphi \urcorner) \equiv \neg \operatorname{prv}(v_0/\ulcorner \sigma_\varphi \urcorner) \equiv \neg \operatorname{prv}(\ulcorner \sigma_\varphi \urcorner)\,.$$

Assume towards a contradiction that PA is complete. We have to consider the following two cases:

Case 1: $\mathsf{PA} \vdash \sigma$. On the one hand, by Corollary 9.17 we have that $\mathsf{PA} \vdash \operatorname{prv}(\ulcorner \sigma \urcorner)$. On the other hand, since $\sigma \Leftrightarrow_{\mathsf{PA}} \neg \operatorname{prv}(\ulcorner \sigma \urcorner)$, we have $\mathsf{PA} \vdash \neg \operatorname{prv}(\ulcorner \sigma \urcorner)$. Therefore, $\mathsf{PA} \vdash \lightning$, i.e., PA is inconsistent.

Case 2: $\mathsf{PA} \vdash \neg\sigma$. From

$$\neg\sigma \Leftrightarrow_{\mathsf{PA}} \neg\neg \operatorname{prv}(\ulcorner \sigma \urcorner) \Leftrightarrow_{\mathsf{PA}} \operatorname{prv}(\ulcorner \sigma \urcorner)$$

we obtain $\mathsf{PA} \vdash \operatorname{prv}(\ulcorner \sigma \urcorner)$. In particular, $\mathbb{N} \vDash \operatorname{prv}(\#\sigma)$, so there exists an $n \in \mathbb{N}$ with $\mathbb{N} \vDash \mathrm{c\_prv}(n, \#\sigma)$. Thus, by Lemma 9.16, n encodes a formal proof of σ, which implies $\mathsf{PA} \vdash \sigma$. Therefore, by our assumption, we have $\mathsf{PA} \vdash \lightning$, i.e., PA is inconsistent.

Summing up, we have

$$\mathsf{PA} \vdash \sigma \quad \text{or} \quad \mathsf{PA} \vdash \neg\sigma \quad \lightning \quad \neg\,\mathrm{Con}(\mathsf{PA}),$$

which shows that PA is incomplete. ⊣

Remarks

- In the above proof of Theorem 10.5 we proved that a sentence σ with the property $\sigma \Leftrightarrow_{\mathsf{PA}} \neg \operatorname{prv}(\ulcorner \sigma \urcorner)$ witnesses the incompleteness of PA. In $\mathbb{N}$, however, σ is true: Note that if $\mathbb{N} \vDash \neg\sigma$, then $\mathbb{N} \vDash \operatorname{prv}(\#\sigma)$. On the

other hand, Lemma 9.16 implies $\mathsf{PA} \vdash \sigma$, and hence, $\mathbb{N} \vDash \sigma$, which is a contradiction. Observe that in $\mathbb{N}$ the sentence σ expresses "I am not provable" — where the expression "provable" is meant with respect to prv — which is, of course, true.

- With respect to the model $\mathbb{N}$, we have

$$\mathbb{N} \vDash \mathrm{prv}(\ulcorner \sigma \urcorner) \quad \Longleftrightarrow \quad \mathsf{PA} \vdash \sigma$$

for every $\mathscr{L}_{\mathsf{PA}}$-sentence σ. However, by the First Incompleteness Theorem we know that this is generally not the case for arbitrary models $\mathbf{M} \vDash \mathsf{PA}$.

- The existence of a sentence σ such that $\mathsf{PA} \nvdash \sigma$ as well as $\mathsf{PA} \nvdash \neg\sigma$ implies that both theories $\mathsf{PA} + \neg\sigma$ and $\mathsf{PA} + \sigma$ are consistent. Hence, by Gödel's Completeness Theorem 5.5, there are models $\mathbf{M}_{\neg\sigma}$ and $\mathbf{M}_{\sigma}$ for $\mathsf{PA} + \neg\sigma$ and $\mathsf{PA} + \sigma$, respectively. Notice that the models $\mathbf{M}_{\neg\sigma}$ and $\mathbf{M}_{\sigma}$ are not elementarily equivalent, and therefore, they are also not isomorphic.

- Let σ_0 be such that neither $\mathsf{PA} \vdash \sigma_0$ nor $\mathsf{PA} \vdash \neg\sigma_0$. Since $\mathsf{PA} \nvdash \sigma_0$, we find that the $\mathscr{L}_{\mathsf{PA}}$-theory $\mathsf{PA} + \neg\sigma_0$ is consistent. Now, by adding the sentence σ_0 as a new axiom to PA, in the same way as above we can construct an $\mathscr{L}_{\mathsf{PA}}$-sentence σ_1 such that neither $\mathsf{PA} + \neg\sigma_0 \vdash \sigma_1$ nor $\mathsf{PA} + \neg\sigma_0 \vdash \neg\sigma_1$ (see below). Proceeding this way, we see that we cannot complete PA by just adding finitely many axioms (for a stronger result see Theorem 10.10 & 10.11).

Completeness and Incompleteness of Arithmetics

A first attempt to deal with the incompleteness phenomenon might be to replace PA with $\mathsf{T} \equiv \mathsf{PA} + \sigma$, since $\mathbb{N} \vDash \mathsf{T}$. Moreover, the gödelisation process could be done in the same way, where one would just need to code an additional axiom, namely σ. This, however, would lead to a modified provability predicate $\mathrm{prv}_{\mathsf{T}}$ which additionally allows formal proofs to be initialised by σ. One could then prove a version of the Diagonalisation Lemma which allows us to define a version σ_{T} of σ with the property

$$\mathsf{T} \vdash \sigma_{\mathsf{T}} \leftrightarrow \neg\, \mathrm{prv}_{\mathsf{T}}(\ulcorner \sigma_{\mathsf{T}} \urcorner)\,,$$

and, since $\mathsf{T} \nvdash \sigma_{\mathsf{T}}$ and $\mathsf{T} \nvdash \neg\sigma_{\mathsf{T}}$, we obtain a version of the First Incompleteness Theorem. This suggests that Theorem 10.5 can be generalised. Such a generalisation is the very goal of this section, whereby we consider both theories which are weaker and stronger than PA. We investigate how

much of PA is really needed for the proof of the First Incompleteness Theorem. As we have seen, exponentiation can be expressed using addition and multiplication. Therefore, one idea might be to leave out multiplication and thus delete PA_4 and PA_5. However, the resulting theory, called **Presburger Arithmetic**, will turn out to be complete (see Chapter 12).

Robinson Arithmetic

The most critical axiom is certainly the Induction Schema PA_6, so we might consider the theory with PA_6 deleted. This is still not strong enough, but we will see that one instance of PA_6 actually suffices: **Robinson Arithmetic** RA is the axiom system consisting of PA_0-PA_5 and the additional axiom

$$\forall x\big(x = \mathsf{0} \vee \exists y(x = \mathsf{s}y)\big)\,.$$

The language of RA is also $\mathscr{L}_{\mathsf{PA}}$, so we can express the same statements as in PA, which implies that RA must be incomplete. On the other hand, we can prove much less in RA than in PA. For example, RA is so weak that we cannot even prove $\forall x(\mathsf{0} + x = x)$.

Example 10.6. We show that $\mathsf{RA} \nvdash \forall x(\mathsf{0} + x = x)$, which implies that we cannot prove within RA that addition is commutative. In order to achieve this, we provide a model $\mathbf{M}$ of RA in which $\forall x(\mathsf{0} + x = x)$ is false. The domain of the model is $M = \mathbb{N} \cup \{a, b\}$, where a and b are two distinct objects which do not belong to $\mathbb{N}$. Furthermore, let $\bar{a} \equiv b$ and $\bar{b} \equiv a$. Then we can interpret $\mathsf{0}^{\mathbf{M}}$ by $\mathbf{0}$ and define the functions $\mathsf{s}^{\mathbf{M}}$, $+^{\mathbf{M}}$, and $\cdot^{\mathbf{M}}$ as follows:

$$\mathsf{s}^{\mathbf{M}}(x) \equiv \begin{cases} \mathsf{s}^{\mathbb{N}}(x) & x \in \mathbb{N} \\ x & x \in \{a, b\} \end{cases}$$

$$x +^{\mathbf{M}} y \equiv \begin{cases} x +^{\mathbb{N}} y & x, y \in \mathbb{N} \\ x & y \in \mathbb{N} \text{ and } x \notin \mathbb{N} \\ \bar{y} & y \notin \mathbb{N} \end{cases}$$

$$x \cdot^{\mathbf{M}} y \equiv \begin{cases} x \cdot^{\mathbb{N}} y & x, y \in \mathbb{N} \\ y & y \in \{0, a, b\} \\ \bar{x} & y \not\equiv 0 \text{ and } x \in \{a, b\} \end{cases}$$

It is easy to check that $\mathbf{M}$ is a model of RA, and that $\mathbf{0} +^{\mathbf{M}} b \equiv a \not\equiv b \equiv b +^{\mathbf{M}} \mathbf{0}$.

Note that N_0–N_5 in Proposition 9.0 are also provable in RA, since the proof uses metainduction rather than induction in PA and the only non-trivial argument uses Lemma 8.5, which can easily be seen to hold in RA.

$\mathbb{N}$-*Conformity Revisited*

In what follows, we prove that all relations and functions that were introduced in Chapters 8 and 9 are $\mathbb{N}$-conform — even with respect to RA. For this purpose, we prove that each such relation and function can be defined both by an $\exists$-formula and a $\forall$-formula. The representations with an $\exists$-formula are already given, and by the proof of Corollary 9.3.(b), functions defined by an $\exists$-formula always have an equivalent definition by a $\forall$-formula. The only relations whose representation by a $\forall$-formula is non-trivial, are term, fml, as well as all relations used to formalise substitution and formal proofs. Note that if we are able to show that the existential quantifiers in term and fml can be replaced by bounded existential quantifiers, then the same can be achieved for all subsequent relations.

Lemma 10.7. *If ψ is a formula of the form $\psi \equiv \exists c(\mathrm{seq}(c) \wedge \varphi(c))$ for some Δ-formula φ, and if there is a term τ whose variables are among* $\mathrm{free}(\psi)$ *such that*

$$\mathsf{PA} \vdash \mathrm{seq}(c) \wedge \varphi(c) \to \big(\mathrm{lh}(c) < \tau \wedge \forall i < \mathrm{lh}(c)(c_i < \tau)\big),$$

then ψ is a Δ-formula as well.

Proof. We go once more through the proof of Theorem 9.9 and show that the quantifier $\exists c$ can be replaced by a bounded quantifier. For this purpose, suppose that $F(i)$ is a function defined by a Δ-formula. Let $F'(i) = \mathrm{op}(\tau, i)+1$ and $m = \max_{i<\tau} F'(i)$. Moreover, note that by Exercise 9.2 we can define factorials in PA; so, let $y := m!$. Furthermore, put $G(j) = 1 + (j+1)y$. By Lemma 8.13, we have that $G(i)$ and $G(j)$ are coprime for all $i, j < m$. Now, Lemma 9.8 allows us to pick x with $\chi(x)$, where

$$\chi(x) \equiv \forall j < m\Big(G(j) \mid x \leftrightarrow \exists i < \tau\big(j = \mathrm{op}(\tau, i)\big)\Big).$$

We check that if $F(i) < \tau$ for every $i < \tau$ then we can find an upper bound τ' whose variables coincide with the variables of τ such that there is $c < \tau'$ with $\beta(c, i) = F(i)$ for all $i < \tau$. If this can be accomplished, then we have

$$\psi \Leftrightarrow_{\mathsf{PA}} \exists c < \tau'(\mathrm{seq}(c) \wedge \varphi(c)).$$

To see this, suppose that $\mathrm{seq}(c) \wedge \varphi(c)$ with $c \geq \tau'$. Now take $F(i) := \beta(c, i) < \tau$. By our assumption, there is $c' < \tau' \leq c$ with $\beta(c', i) = F(i) = \beta(c, i)$ for all $i < \tau$. Moreover, note that $\mathrm{lh}(c') = \beta(c', 0) = \beta(c, 0) = \mathrm{lh}(c)$ and $\mathrm{lh}(c') = F(0) < \tau$, and hence $c'_i = c_i$ for all $i < \mathrm{lh}(c)$, contradicting $\mathrm{seq}(c)$.

It remains to find τ'. Note that we clearly have $m \leq \tau_1$ with $\tau_1 \equiv \mathrm{op}(\tau, \tau)+1$ and hence $y \leq \tau_1!$. Furthermore, we have $G(j) < 1 + (\tau_1 + 1)!$ for each $j < m$. Therefore, since $G(i)$ and $G(j)$ are coprime for all $i, j < m$, we can find x which satisfies $\chi(x)$ such that $x < \tau_2$ with $\tau_2 \equiv (1 + (\tau_1 + 1)!)^{\tau_1}$. In particular, there is $c = \mathrm{op}(x, y)$ with $\mathrm{seq}(c) \wedge \varphi(c)$ and $c < \mathrm{op}(\tau_1, \tau_2)$. ⊣

LEMMA 10.8. *The relations* term *and* fml *are* $\mathbb{N}$*-conform.*

Proof. We want to apply LEMMA 10.7 to the defining formulae of term and formula, respectively. Since both cases are similar, we only consider term. We prove that $\exists c\big(\mathrm{c\_term}(c,t)\big)$ is equivalent to the formula

$$\varphi(t) \equiv \exists c(\mathrm{c\_term}(c,t) \wedge \forall i < \mathrm{lh}(c) \forall j < i(c_j < c_i)).$$

Then LEMMA 10.7 for $\tau \equiv t+1$ concludes the proof. We proceed by strong induction on $\mathrm{lh}(c)$. If $\mathrm{lh}(c) = \mathtt{1}$, then there is nothing to prove. Suppose now that $\mathrm{term}(t) \to \varphi(t)$ holds for all $t' < t$ and assume $\mathrm{c\_term}(c,t)$. If $t = \mathtt{0}$ or $\mathrm{var}(t)$, then $\mathrm{c\_term}(\langle t\rangle, t)$ and hence $\varphi(t)$ holds. Therefore, we have either $t = \mathrm{succ}(c_i)$ or $t = \mathrm{add}(c_i, c_j)$ or $t = \mathrm{mult}(c_i, c_j)$ for $i, j < \mathrm{lh}(c)$. We only focus on the first case, since the others can be handled in the same way. Note that by Exercise 9.8 we can restrict c to $\langle c_j \mid j \leq i\rangle$, which we denote by $c \restriction \mathtt{s}c_i$. Clearly, $c_i < c$ and $\mathrm{c\_term}(c \restriction \mathtt{s}i, c_i)$. Hence, by our induction hypothesis, there is d with $\mathrm{c\_term}(d, c_i)$ and $d_k < d_j$ for all $j < \mathrm{lh}(d)$ and $k < j$. But then, $d^\frown\langle t\rangle$ witnesses $\varphi(t)$. ⊣

Let us now turn back to RA: LEMMA 10.8 implies that if $n \in \mathbb{N}$ is a natural number which is not the Gödel number of a term or formula, then

$$\mathsf{RA} \vdash \neg\, \mathrm{term}(\underline{n}) \qquad \text{and} \qquad \mathsf{RA} \vdash \neg\, \mathrm{fml}(\underline{n})\,.$$

Moreover, since the relation c_prv is also a Δ-formula, we have

$$\mathsf{RA} \vdash \neg\, \mathrm{c\_prv}(\underline{n}, \ulcorner \sigma \urcorner)$$

whenever n does not encode a formal proof of σ. However, the existential quantifier in the definition of the provability relation prv cannot be bounded, since otherwise, $\mathsf{RA} \nvdash \sigma$ would imply $\mathsf{RA} \vdash \neg\, \mathrm{prv}(\ulcorner \sigma \urcorner)$, which contradicts the incompleteness of RA.

Generalising the First Incompleteness Theorem

There are two ways to generalise the FIRST INCOMPLETENESS THEOREM: Firstly, one can modify the underlying language, and secondly, one can use a different axiom system. If the language satisfies $\mathscr{L} \supseteq \mathscr{L}_{\mathsf{PA}}$ and we have $\mathbb{N}$-conformity for all relevant relations, then, as we shall see, the proof can easily be transferred to the new setting. However, there are two issues at stake, namely the gödelisation of the language and the gödelisation of the axioms. The coding of terms, formulae and proofs can then be realised in the same way as in Chapter 9.

A language $\mathscr{L} \supseteq \mathscr{L}_{\mathsf{PA}}$ is said to be **gödelisable** if it is countable. Note that if $\mathscr{L}$ is gödelisable, then its constant symbols, relation and function symbols admit a Gödel coding as described in Chapter 9. A theory T in some

gödelisable language $\mathscr{L} \supseteq \mathscr{L}_{\mathsf{PA}}$ is **gödelisable**, if there is a Δ-formula ax_{T} in the language $\mathscr{L}_{\mathsf{PA}}$ with the property that

$$\mathbb{N} \vDash \mathrm{ax}_{\mathsf{T}}(\#\varphi) \quad \text{if and only if} \quad \varphi \in \mathsf{T}\,,$$

where $\#\varphi$ is the Gödel code of φ. As in the case of PA, we introduce Gödel codes on the formal level by stipulating $\ulcorner\varphi\urcorner :\equiv \underline{\#\varphi}$. Note that if T is gödelisable and satisfies N_0–N_5, then by Corollary 9.3, every Δ-formula φ in the language $\mathscr{L}_{\mathsf{PA}}$ is $\mathbb{N}$-conform. In particular, by Lemma 10.8 it is possible to define Δ-formulae $\mathrm{term}_{\mathsf{T}}$ and $\mathrm{fml}_{\mathsf{T}}$ such that

$$\begin{aligned} \mathbb{N} \vDash \mathrm{term}_{\mathsf{T}}(n) \quad &\Longleftrightarrow \quad n \equiv \#\tau \text{ for some } \mathscr{L}\text{-term } \tau\,, \\ \mathbb{N} \vDash \mathrm{fml}_{\mathsf{T}}(n) \quad &\Longleftrightarrow \quad n \equiv \#\varphi \text{ for some } \mathscr{L}\text{-formula } \varphi\,. \end{aligned}$$

Moreover, by the gödelisability of T, the axioms can be coded by some Δ-formula ax_{T}. One can then proceed to define a Δ-formula $\mathrm{c\_prv}_{\mathsf{T}}$ and an $\exists$-formula $\mathrm{prv}_{\mathsf{T}}$ such that

$$\begin{aligned} \mathbb{N} \vDash \mathrm{c\_prv}_{\mathsf{T}}(n, \#\varphi) \quad &\Longleftrightarrow \quad n \text{ codes a formal proof of } \varphi\,, \\ \mathbb{N} \vDash \mathrm{prv}_{\mathsf{T}}(\#\varphi) \quad &\Longleftrightarrow \quad \mathsf{T} \vdash \varphi \end{aligned}$$

for every $n \in \mathbb{N}$ and $\mathscr{L}$-formula φ. Notice that it is crucial that $\mathrm{c\_prv}_{\mathsf{T}}$ and $\mathrm{prv}_{\mathsf{T}}$ are $\mathscr{L}_{\mathsf{PA}}$-formulae, since otherwise we would have to specify how to interpret them in the standard model $\mathbb{N}$. Moreover, using Corollary 9.3, we obtain

$$\begin{aligned} &\mathsf{P}_0\text{:} && \mathbb{N} \vDash \mathrm{c\_prv}_{\mathsf{T}}(n, \#\varphi) \quad &&\Longrightarrow \quad \mathsf{T} \vdash \mathrm{c\_prv}_{\mathsf{T}}(\underline{n}, \ulcorner\varphi\urcorner)\,, \\ &\mathsf{P}_1\text{:} && \mathbb{N} \vDash \neg\,\mathrm{c\_prv}_{\mathsf{T}}(n, \#\varphi) \quad &&\Longrightarrow \quad \mathsf{T} \vdash \neg\,\mathrm{c\_prv}_{\mathsf{T}}(\underline{n}, \ulcorner\varphi\urcorner)\,. \end{aligned}$$

In the following, we present two proofs of the First Incompleteness Theorem for gödelisable theories $\mathsf{T} \supseteq \mathsf{RA}$. The restriction to extensions of RA ensures that N_0–N_5, and hence also Corollary 9.3, hold.

Gödel's original proof uses the assumption of a slightly stronger property than just consistency: An $\mathscr{L}_{\mathsf{PA}}$-theory T is said to be **ω-consistent** if whenever $\mathsf{T} \vdash \exists x\varphi(x)$ for some $\mathscr{L}_{\mathsf{PA}}$-formula $\varphi(x)$, then there exists $n \in \mathbb{N}$ such that $\mathsf{T} \nvdash \neg\varphi(\underline{n})$.

Fact 10.9. *If T is an $\mathscr{L}_{\mathsf{PA}}$-theory with $\mathbb{N} \vDash \mathsf{T}$, then T is ω-consistent. In particular,* PA *and* RA *are ω-consistent.*

Proof. If $\mathsf{T} \vdash \exists x\varphi(x)$, then $\mathbb{N} \vDash \exists x\varphi(x)$. Hence, there is an $n \in \mathbb{N}$ with $\mathbb{N} \vDash \varphi(n)$, which shows that $\mathsf{T} + \varphi(\underline{n})$ is consistent and implies $\mathsf{T} \nvdash \neg\varphi(\underline{n})$. ⊣

Theorem 10.10 (First Incompleteness Theorem, Gödel's version). *Let $\mathsf{T} \supseteq \mathsf{RA}$ be a gödelisable $\mathscr{L}_{\mathsf{PA}}$-theory. If T is ω-consistent, then T is incomplete.*

Proof. Observe that the proof of the Diagonalisation Lemma still works if we replace PA by T. Take a sentence σ such that

$$\sigma \Leftrightarrow_{\mathsf{PA}} \neg \mathrm{prv}_{\mathsf{T}}(\ulcorner \sigma \urcorner).$$

Assume towards a contradiction that T is complete. Then we have that either $\mathsf{T} \vdash \sigma$ or $\mathsf{T} \vdash \neg\sigma$.

Case 1: $\mathsf{T} \vdash \sigma$. In this case, the argument is the same as in Theorem 10.5.

Case 2: $\mathsf{T} \vdash \neg\sigma$. Since we can encode the proof of $\neg\sigma$ within T, we have

$$\mathsf{T} \vdash \mathrm{prv}_{\mathsf{T}}(\ulcorner \neg\sigma \urcorner)\,.$$

On the other hand, we have $\neg\sigma \Leftrightarrow_{\mathsf{T}} \neg\neg \mathrm{prv}_{\mathsf{T}}(\ulcorner \sigma \urcorner) \Leftrightarrow_{\mathsf{T}} \mathrm{prv}_{\mathsf{T}}(\ulcorner \sigma \urcorner)$, and therefore we also have

$$\mathsf{T} \vdash \mathrm{prv}_{\mathsf{T}}(\ulcorner \sigma \urcorner)\,.$$

So, by Corollary 10.1, we have $\mathsf{T} \vdash \mathrm{prv}_{\mathsf{T}}(\ulcorner \sigma \wedge \neg\sigma \urcorner)$, and by the ω-consistency of T, there is an $n \in \mathbb{N}$ such that

$$\mathsf{T} \nvdash \neg\, \mathrm{c\_prv}_{\mathsf{T}}(\underline{n}, \ulcorner \sigma \wedge \neg\sigma \urcorner)\,.$$

However, since T is consistent, we have $\mathsf{T} \nvdash \sigma \wedge \neg\sigma$, which implies

$$\mathbb{N} \vDash \neg\, \mathrm{c\_prv}_{\mathsf{T}}(n, \#(\sigma \wedge \neg\sigma))\,,$$

and by P_1 we obtain

$$\mathsf{T} \vdash \neg\, \mathrm{c\_prv}_{\mathsf{T}}(\underline{n}, \ulcorner \sigma \wedge \neg\sigma \urcorner)\,,$$

which is obviously a contradiction. ⊣

In [47], Rosser showed how to get rid of this dependency on ω-consistency by slightly modifying the provability predicate:

$$\mathrm{c\_prv}^{\mathrm{R}}_{\mathsf{T}}(c, x) :\Longleftrightarrow \mathrm{c\_prv}_{\mathsf{T}}(c, x) \wedge \neg\exists c' < c(\mathrm{c\_prv}_{\mathsf{T}}(c', \mathrm{not}(x)))$$
$$\mathrm{prv}^{\mathrm{R}}_{\mathsf{T}}(x) :\Longleftrightarrow \exists c(\mathrm{c\_prv}^{\mathrm{R}}_{\mathsf{T}}(c, x))$$

Theorem 10.11 (First Incompleteness Theorem, using Rosser's Trick). *Let $\mathscr{L} \supseteq \mathscr{L}_{\mathsf{PA}}$ be a gödelisable language and let T be a gödelisable $\mathscr{L}$-theory. If T is consistent, then it is incomplete.*

Proof. As before, we apply the Diagonalisation Lemma, this time to the formula $\neg\, \mathrm{prv}^{\mathrm{R}}_{\mathsf{T}}(x)$. Thus we obtain an $\mathscr{L}$-sentence σ with

$$\sigma \Leftrightarrow_{\mathsf{PA}} \neg\, \mathrm{prv}^{\mathrm{R}}_{\mathsf{T}}(\ulcorner \sigma \urcorner).$$

Again, we prove that neither σ nor $\neg\sigma$ is provable from T. Observe first that our assumption on σ implies

$$\sigma \Leftrightarrow_{\mathsf{PA}} \forall c(\mathrm{c\_prv}(c, \ulcorner \sigma \urcorner) \to \exists c' < c\,(\mathrm{c\_prv}(c', \ulcorner \neg\sigma \urcorner)))$$

since $\mathrm{not}(\ulcorner \sigma \urcorner) \equiv \ulcorner \neg\sigma \urcorner$. Assume towards a contradiction that T is complete. As before, we have two cases:

Case 1: $\mathsf{T} \vdash \sigma$. By $\mathsf{P_0}$, there is an $n \in \mathbb{N}$ such that

$$\mathsf{T} \vdash \mathrm{c\_prv}_{\mathsf{T}}(\underline{n}, \ulcorner \sigma \urcorner)\,,$$

and by our above observation we have

$$\mathsf{T} \vdash \exists c' < \underline{n}(\mathrm{c\_prv}_{\mathsf{T}}(c', \ulcorner \neg\sigma \urcorner))\,.$$

Since T satisfies $\mathsf{N_5}$, this means that there exists $k < n$ in $\mathbb{N}$ such that

$$\mathsf{T} \vdash \mathrm{c\_prv}_{\mathsf{T}}(\underline{k}, \ulcorner \neg\sigma \urcorner)\,,$$

and therefore, there is an $m \in \mathbb{N}$ with

$$\mathsf{T} \vdash \mathrm{c\_prv}_{\mathsf{T}}(\underline{m}, \ulcorner \sigma \wedge \neg\sigma \urcorner)\,.$$

Hence, by $\mathbb{N}$-conformity of $\mathrm{c\_prv}_{\mathsf{T}}$, we have

$$\mathbb{N} \vDash \mathrm{c\_prv}_{\mathsf{T}}(m, \#(\sigma \wedge \neg\sigma))\,,$$

which implies

$$\mathsf{T} \vdash \sigma \wedge \neg\sigma.$$

This contradicts our assumption that T is consistent.

Case 2: $\mathsf{T} \vdash \neg\sigma$. In this case, there is a $c' \in \mathbb{N}$ such that

$$\mathsf{T} \vdash \mathrm{c\_prv}_{\mathsf{T}}(\underline{c'}, \ulcorner \neg\sigma \urcorner)\,.$$

On the other hand, we have $\mathsf{T} \vdash \mathrm{prv}^{\mathrm{R}}_{\mathsf{T}}(\ulcorner \sigma \urcorner)$, and hence, there is a $c \in \mathbb{N}$ with

$$\mathsf{T} \vdash \mathrm{c\_prv}^{\mathrm{R}}_{\mathsf{T}}(\underline{c}, \ulcorner \sigma \urcorner)\,.$$

By definition of $\mathrm{c\_prv}^{\mathrm{R}}_{\mathsf{T}}$, we must have $c \leq c'$. Now, we can use $\mathsf{N_5}$ to reach the same contradiction as in the first case. ⊣

Tarski's Theorem

The Diagonalisation Lemma allows us to make self-referential statements such as the Gödel sentence which formalises to some extent the sentence "This sentence is not provable". Recall that we call an $\mathscr{L}_{\mathsf{PA}}$-sentence φ **true** in $\mathbb{N}$, if $\mathbb{N} \vDash \varphi$. Is it possible to express truth in the standard model $\mathbb{N}$ by a

formula, i.e., is there a formula $\mathrm{truth}(x)$ with one free variable x such that for every $\mathscr{L}_{\mathsf{PA}}$-sentence φ,

$$\mathbb{N} \vDash \mathrm{truth}(\#\varphi) \quad \Longleftrightarrow \quad \mathbb{N} \vDash \varphi\,?$$

Or equivalently, is there a formula $\mathrm{truth}(x)$ such that for every $\mathscr{L}_{\mathsf{PA}}$-sentence φ,

$$\mathbb{N} \vDash \mathrm{truth}(\#\varphi) \leftrightarrow \varphi\,?$$

Using the Diagonalisation Lemma, we provide a negative answer to this question.

Theorem 10.12 (Tarski's Theorem). *There is no $\mathscr{L}_{\mathsf{PA}}$-formula $\mathrm{truth}(x)$ with one free variable x such that $\mathbb{N} \vDash \mathrm{truth}(\#\varphi) \leftrightarrow \varphi$.*

Proof. Assume towards a contradiction that such a formula truth exists. By the Diagonalisation Lemma there exists an $\mathscr{L}_{\mathsf{PA}}$-sentence σ such that

$$\mathsf{PA} \vdash \sigma \leftrightarrow \neg\,\mathrm{truth}(\ulcorner\sigma\urcorner)\,.$$

But then

$$\begin{aligned} \mathbb{N} \vDash \mathrm{truth}(\#\sigma) \quad &\Longleftrightarrow \quad \mathbb{N} \vDash \sigma \\ &\Longleftrightarrow \quad \mathbb{N} \vDash \neg\,\mathrm{truth}(\#\sigma), \end{aligned}$$

which is impossible. ⊣

Note that we have just solved the so-called **Liar Paradox** concerned with the sentence

"This sentence is false.",

which is true in some model if and only if it is false in that model. If we work in the model $\mathbb{N}$, then the above sentence corresponds to the sentence σ in the proof of Tarski's Theorem. Hence, in order to express it in PA, one would need to be able to define truth in $\mathbb{N}$, which is impossible by Tarski's Theorem.

Notes

The First Incompleteness Theorem was first proven by Gödel [16] in 1931. Rather than using Peano Arithmetic in first-order logic, as we did, he based his proof on Type Theory in the system of Principia Mathematica [58] introduced by Russell and Whitehead. Gödel's original proof makes use of the stronger assumption of ω-consistency, which Rosser [47] showed to be negligible. The observation that all proof steps of the First Incompleteness Theorem can in fact be carried out in Robinson Arithmetic was made by Robinson [46] in 1950. Although Tarski's Theorem is usually attributed to Tarski and was first published by him in [55], Gödel already mentioned this result in 1931 in a letter

to Bernays; previously he had been trying to come up with a definition of a truth predicate (see [37]). Usually, gödelisable theories are called *recursive*, which means that there exists an algorithm terminating after finitely many steps that can decide whether $\varphi \in \mathsf{T}$ or $\varphi \notin \mathsf{T}$. More generally, a property $P(n)$ of natural numbers is said to be recursive, if there is an algorithm which decides in finitely many steps whether a given number n has the property P (i.e., whether or not $P(n)$ holds). With the so-called *Recursion Theory*, one can analyse the strength of various theories of Arithmetic very precisely.

Exercises

10.0 Let φ and ψ be $\mathscr{L}_{\mathsf{PA}}$-formulae.

(a) Show that $\mathsf{PA} \vdash \big(\mathrm{prv}(\ulcorner\varphi\urcorner) \vee \mathrm{prv}(\ulcorner\psi\urcorner)\big) \to \mathrm{prv}(\ulcorner\varphi \vee \psi\urcorner)$.

(b) Does the converse also hold?

10.1 A theory T with signature $\mathscr{L}_{\mathsf{PA}}$ is said to be **ω-incomplete** if there is an $\mathscr{L}_{\mathsf{PA}}$-formula φ such that $\mathsf{T} \vdash \varphi(\underline{n})$ for every $n \in \mathbb{N}$ but $\mathsf{T} \nvdash \forall x \varphi(x)$.

Show that PA is ω-incomplete.

10.2 Let $\varphi_1(x, y)$ and $\varphi_2(x, y)$ be $\mathscr{L}_{\mathsf{PA}}$-formulae with at most two free variables. Show that there are $\mathscr{L}_{\mathsf{PA}}$-sentences σ_1 and σ_2 such that

$$\sigma_1 \Leftrightarrow_{\mathsf{PA}} \varphi_1(\ulcorner\sigma_1\urcorner, \ulcorner\sigma_2\urcorner) \quad \text{and} \quad \sigma_2 \Leftrightarrow_{\mathsf{PA}} \varphi_2(\ulcorner\sigma_1\urcorner, \ulcorner\sigma_2\urcorner).$$

Note that this is a generalisation of the Diagonalisation Lemma.

10.3 **Goldbach's Conjecture** (GC) states that every positive even integer can be written as the sum of two primes.

Show that if GC is independent of PA, i.e., $\mathsf{PA} \nvdash \mathsf{GC}$ and $\mathsf{PA} \nvdash \neg\mathsf{GC}$, then it is true in the model $\mathbb{N}$, i.e., $\mathbb{N} \vDash \mathsf{GC}$.

10.4 Show that there is a consistent extension of PA which is not ω-consistent.

Chapter 11
The Second Incompleteness Theorem

It follows from Gödel's Completeness Theorem that a theory is consistent if and only if it has a model. In particular, the consistency of Peano Arithmetic follows from $\mathbb{N} \vDash \mathsf{PA}$. With the help of the provability relation prv, we are even able to express consistency of an arithmetical theory on the formal level, i.e., we can introduce a sentence $\mathrm{con}_{\mathsf{PA}}$ which expresses in $\mathbb{N}$ the consistency of PA. The Second Incompleteness Theorem which we shall prove in this chapter states that $\mathsf{PA} \nvdash \mathrm{con}_{\mathsf{PA}}$, i.e., PA cannot prove its own consistency.

Outline of the Proof

Recall that a theory is consistent, if it cannot prove contradictions. In the case of PA, a simple contradiction is the sentence $0 = 1$. Thus, we have

$$\mathrm{Con}(\mathsf{PA}) \quad \Lleftarrow\Rrightarrow \quad \mathsf{PA} \nvdash 0 = 1\,.$$

As a formalised version of this statement, we define the $\mathscr{L}_{\mathsf{PA}}$-sentence $\mathrm{con}_{\mathsf{PA}}$ by stipulating

$$\mathrm{con}_{\mathsf{PA}} :\Longleftrightarrow \neg\,\mathrm{prv}(\ulcorner 0 = 1 \urcorner).$$

Since $\mathbb{N} \vDash \mathrm{prv}(\ulcorner \varphi \urcorner)$ if and only if $\mathsf{PA} \vdash \varphi$, the consistency of PA implies that $\mathbb{N} \vDash \mathrm{con}_{\mathsf{PA}}$. In particular, this shows that $\mathsf{PA} \nvdash \neg\,\mathrm{con}_{\mathsf{PA}}$. The Second Incompleteness Theorem states that $\mathrm{con}_{\mathsf{PA}}$ is independent of the axioms of PA.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_12

THEOREM 11.0 (SECOND INCOMPLETENESS THEOREM). $\mathsf{PA} \nvdash \mathrm{con}_{\mathsf{PA}}$.

As a matter of fact, we would like to mention that by the COMPLETENESS THEOREM, $\mathsf{PA} \nvdash \mathrm{con}_{\mathsf{PA}}$ implies that there exists a model $\mathbf{M} \models \mathsf{PA}$ in which $\mathrm{con}_{\mathsf{PA}}$ fails (i.e., $\mathbf{M} \models \neg\,\mathrm{con}_{\mathsf{PA}}$), which shows that with respect to $\mathbf{M}$, the sentence $\mathrm{con}_{\mathsf{PA}}$ is *not* equivalent to the statement $\mathrm{Con}(\mathsf{PA})$.

The proof of the SECOND INCOMPLETENESS THEOREM hinges on the following properties of the provability predicate, also called the *Hilbert-Bernays-Löb derivability conditions*, which state that for every $\mathscr{L}_{\mathsf{PA}}$-formula φ the following conditions hold:

D_0: If $\mathsf{PA} \vdash \varphi$ then $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner\varphi\urcorner)$,

D_1: $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner\varphi \to \psi\urcorner) \to (\mathrm{prv}(\ulcorner\varphi\urcorner) \to \mathrm{prv}(\ulcorner\psi\urcorner))$,

D_2: $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner\varphi\urcorner) \to \mathrm{prv}(\ulcorner\mathrm{prv}(\ulcorner\varphi\urcorner)\urcorner)$.

Note that D_0 follows from COROLLARY 9.17 and D_1 is exactly the statement of Corollary 10.1.(a). Assuming D_0–D_2, the proof of the SECOND INCOMPLETENESS THEOREM becomes quite simple:

Proof of Theorem 11.0. Assume towards a contradiction that $\mathsf{PA} \vdash \mathrm{con}_{\mathsf{PA}}$, in other words, assume that $\mathsf{PA} \vdash \neg\,\mathrm{prv}(\ulcorner\mathsf{0} = \mathsf{1}\urcorner)$. Using the DIAGONALISATION LEMMA we can find an $\mathscr{L}_{\mathsf{PA}}$-sentence σ such that

$$\sigma \Leftrightarrow_{\mathsf{PA}} \neg\,\mathrm{prv}(\ulcorner\sigma\urcorner).$$

Now, observe that by COROLLARY 10.2 we have

$$\mathrm{prv}(\ulcorner\mathsf{0} = \mathsf{1}\urcorner) \;\Leftrightarrow_{\mathsf{PA}}\; \mathrm{prv}(\ulcorner\sigma \wedge \neg\sigma\urcorner) \;\Leftrightarrow_{\mathsf{PA}}\; \mathrm{prv}(\ulcorner\sigma\urcorner) \wedge \mathrm{prv}(\ulcorner\neg\sigma\urcorner).$$

Another application of COROLLARY 10.2 yields

$$\mathrm{prv}(\ulcorner\neg\sigma\urcorner) \Leftrightarrow_{\mathsf{PA}} \mathrm{prv}(\ulcorner\mathrm{prv}(\ulcorner\sigma\urcorner)\urcorner)\,,$$

and therefore we have

$$\mathrm{prv}(\ulcorner\mathsf{0} = \mathsf{1}\urcorner) \Leftrightarrow_{\mathsf{PA}} \mathrm{prv}(\ulcorner\sigma\urcorner) \wedge \mathrm{prv}(\ulcorner\mathrm{prv}(\ulcorner\sigma\urcorner)\urcorner).$$

Furthermore, by D_2 we have $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner\varphi\urcorner) \to \mathrm{prv}(\ulcorner\mathrm{prv}(\ulcorner\varphi\urcorner)\urcorner)$, and hence, by TAUTOLOGY (D.2) we obtain

$$\mathsf{PA} \vdash \mathrm{prv}(\ulcorner\sigma\urcorner) \to \big(\mathrm{prv}(\ulcorner\sigma\urcorner) \wedge \mathrm{prv}(\ulcorner\mathrm{prv}(\ulcorner\sigma\urcorner)\urcorner)\big)\,,$$

and by L_3, this implies

$$\mathrm{prv}(\ulcorner\sigma\urcorner) \wedge \mathrm{prv}(\ulcorner\mathrm{prv}(\ulcorner\sigma\urcorner)\urcorner) \Leftrightarrow_{\mathsf{PA}} \mathrm{prv}(\ulcorner\sigma\urcorner)\,.$$

Therefore, we obtain

$$\mathrm{prv}(\ulcorner 0 = 1 \urcorner) \Leftrightarrow_{\mathsf{PA}} \mathrm{prv}(\ulcorner \sigma \urcorner)\,,$$

and consequently, we have

$$\mathrm{con}_{\mathsf{PA}} \Leftrightarrow_{\mathsf{PA}} \neg\, \mathrm{prv}(\ulcorner \sigma \urcorner) \Leftrightarrow_{\mathsf{PA}} \sigma,$$

which is a contradiction to Theorem 10.5 which states that $\mathsf{PA} \nvdash \sigma$. ⊣

Proving the Derivability Condition D_2

In order to complete our proof of Gödel's Second Incompleteness Theorem, it remains to prove D_2. At a first glance, it looks very similar to the statement D_0. There is, however, a subtle difference between the two statements: While the implication in D_0 is just a meta-implication, i.e., an implication in the meta-logic, the implication in D_2 is a formal one. Note that it follows from D_0 that

$$\text{if } \ \mathsf{PA} \vdash \mathrm{prv}(\ulcorner \varphi \urcorner) \ \text{ then } \ \mathsf{PA} \vdash \mathrm{prv}(\ulcorner \mathrm{prv}(\ulcorner \varphi \urcorner) \urcorner)\,,$$

which, however, is weaker than D_2. A first attempt would be to try to prove

$$\mathsf{PA} \vdash \alpha \to \mathrm{prv}(\ulcorner \alpha \urcorner)$$

for every $\mathscr{L}_{\mathsf{PA}}$-formula α. However, this is false in general, as the following example shows:

Example 11.1. Let σ denote the formula from the proof of the First Incompleteness Theorem, i.e., σ satisfies

$$\sigma \Leftrightarrow_{\mathsf{PA}} \neg\, \mathrm{prv}(\ulcorner \sigma \urcorner).$$

As a consequence of the proof of First Incompleteness Theorem, we have $\mathbb{N} \vDash \sigma$ but $\mathsf{PA} \nvdash \sigma$. Now, if $\mathsf{PA} \vdash \sigma \to \mathrm{prv}(\ulcorner \sigma \urcorner)$, then $\mathbb{N} \vDash \sigma \to \mathrm{prv}(\ulcorner \sigma \urcorner)$ and hence $\mathbb{N} \vDash \mathrm{prv}(\ulcorner \sigma \urcorner)$. By construction of the provability predicate, this would imply $\mathsf{PA} \vdash \sigma$, which is not the case, as we have seen.

This means that we have to slightly modify our approach. For this purpose, recall that we proved in Proposition 9.2.(a) that every $\exists$-sentence which is true in the standard model $\mathbb{N}$ has a formal proof in PA. If we can transfer this result to PA, this would mean that we have

D_3: $\mathsf{PA} \vdash \alpha \to \mathrm{prv}(\ulcorner \alpha \urcorner)$ for every $\exists$-sentence α.

Clearly, once we have established D_3 we obtain D_2 by taking α to be the $\exists$-sentence $\mathrm{prv}(\ulcorner \varphi \urcorner)$ for some $\mathscr{L}_{\mathsf{PA}}$-formula φ. The most natural way to prove

D_3 is by induction on the construction of the $\exists$-sentence α. This, however, turns out to be problematic, since in the formula construction of α there are also subformulae which are not sentences:

Example 11.2. Assume that we can prove $\mathsf{PA} \vdash \alpha \to \mathrm{prv}(\ulcorner \alpha \urcorner)$ for some $\mathscr{L}_{\mathsf{PA}}$-formula $\alpha \equiv (v_0 = v_1)$, i.e.,

$$\mathsf{PA} \vdash v_0 = v_1 \to \mathrm{prv}(\ulcorner v_0 = v_1 \urcorner)\,.$$

Observe that $\mathrm{prv}(\ulcorner v_0 = v_1 \urcorner)$ does not contain any free variables. Now, since v_0 and v_1 are free variables in α, we obtain by substitution $\mathsf{PA} \vdash \mathtt{0} = \mathtt{0} \to \mathrm{prv}(\ulcorner v_0 = v_1 \urcorner)$. Therefore, using Modus Ponens, we get $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner v_0 = v_1 \urcorner)$, which is clearly false in the standard model $\mathbb{N}$, since there is no formal proof of $v_0 = v_1$.

This problem can be solved by slightly modifying our provability predicate in such a way that free variables are permitted. Thus, we first want to adjust our Gödel coding such that (some) free variables can be preserved. The way to do this is by defining for some set V of variables

$$\lceil \nu \rceil_V :\equiv \begin{cases} \nu & \text{if } \nu \in V, \\ \ulcorner \nu \urcorner & \text{otherwise.} \end{cases}$$

Roughly speaking, variables $\nu \in V$ remain variables and all other variables become natural numbers, namely $\ulcorner \nu \urcorner$. Now, as in the case of gödelisation, we can inductively extend this definition to terms by stipulating:

$$\begin{aligned} \lceil \mathtt{0} \rceil_V &:\equiv \ulcorner \mathtt{0} \urcorner \\ \lceil \mathtt{s}\tau \rceil_V &:\equiv \mathrm{succ}(\lceil \tau \rceil_V) \\ \lceil \tau_1 + \tau_2 \rceil_V &:\equiv \mathrm{add}(\lceil \tau_1 \rceil_V, \lceil \tau_2 \rceil_V) \\ \lceil \tau_1 \cdot \tau_2 \rceil_V &:\equiv \mathrm{mult}(\lceil \tau_1 \rceil_V, \lceil \tau_2 \rceil_V) \end{aligned}$$

For formulae, one proceeds similarly. The only noteworthy cases are those of quantification:

$$\begin{aligned} \lceil \exists \nu \varphi \rceil_V &:\equiv \mathrm{ex}(\ulcorner \nu \urcorner, \lceil \varphi \rceil_{V \setminus \{\nu\}}), \\ \lceil \forall \nu \varphi \rceil_V &:\equiv \mathrm{all}(\ulcorner \nu \urcorner, \lceil \varphi \rceil_{V \setminus \{\nu\}}). \end{aligned}$$

Thus, the set V contains all variables which remain free in $\lceil \varphi \rceil_V$. In particular, if $V \cap \mathrm{free}(\varphi)$ is the empty-set, then $\lceil \varphi \rceil_V$ is the same as $\ulcorner \varphi \urcorner$. The other special case is when V contains all free variables in φ. In that case, we write $\lceil \varphi \rceil$ for $\lceil \varphi \rceil_V$ and say that $\lceil \varphi \rceil$ is the **pseudo-code** of φ.

Pseudo-coding is intended to mimic the usual process of Gödel coding. Hence, we will often substitute the free variables ν of $\lceil \varphi \rceil$ by the term $\mathrm{gn}(\nu)$

with free variable ν. For terms τ and formulae φ whose free variables are among $\{x_1, \ldots, x_n\}$, we will henceforth use the notation

$$\lceil \tau \rceil_V^{\mathrm{gn}} :\equiv \lceil \tau \rceil_V \big(x_1/\operatorname{gn}(x_1), \ldots, x_n/\operatorname{gn}(x_n)\big)\,,$$
$$\lceil \varphi \rceil_V^{\mathrm{gn}} :\equiv \lceil \varphi \rceil_V \big(x_1/\operatorname{gn}(x_1), \ldots, x_n/\operatorname{gn}(x_n)\big)\,.$$

Note that $\lceil \varphi \rceil_V^{\mathrm{gn}}$ has the same free variables as $\lceil \varphi \rceil_V$. Again, if V contains all free variables of φ then we simply write $\lceil \varphi \rceil^{\mathrm{gn}}$ for $\lceil \varphi \rceil_V^{\mathrm{gn}}$. Recall that for natural numbers $n \in \mathbb{N}$, Lemma 10.3 implies $\mathsf{PA} \vdash \ulcorner \underline{n} \urcorner = \operatorname{gn}(\underline{n})$. In particular, if we substitute each variable x_i by some natural number m_i, then $\ulcorner \varphi \urcorner$ and $\lceil \varphi \rceil^{\mathrm{gn}}$ coincide, i.e.,

$$\ulcorner \varphi(x_1/\underline{m_1}, \ldots, x_n/\underline{m_n}) \urcorner \quad \text{is equal to} \quad \lceil \varphi \rceil^{\mathrm{gn}}(x_1/\underline{m_1}, \ldots, x_n/\underline{m_n})\,.$$

To see this, notice that on the left hand side, the variable x_i is first replaced by $\underline{m_i}$, and when computing $\ulcorner \varphi \urcorner$, $\underline{m_i}$ is replaced by $\ulcorner \underline{m_i} \urcorner$. On the right hand side, when computing $\lceil \varphi \rceil^{\mathrm{gn}}$, the variable x_i is replaced by $\operatorname{gn}(x_i)$, and then x_i — which is a free variable in $\operatorname{gn}(x_i)$ — is replaced by $\underline{m_i}$. Thus, on the left hand side, x_i is replaced by $\ulcorner \underline{m_i} \urcorner$, and on the right hand side, x_i is replaced by $\operatorname{gn}(\underline{m_i})$, and as mentioned above, $\ulcorner \underline{m_i} \urcorner$ is equal to $\operatorname{gn}(\underline{m_i})$. A slightly stronger result is given by the following

Fact 11.3. *For terms τ and formulae φ whose free variables are among $\{x_1, \ldots, x_n\}$, we have*

$$\mathsf{PA} \vdash \ulcorner \tau(x_1/\underline{m_1}, \ldots, x_n/\underline{m_n}) \urcorner = \lceil \tau \rceil^{\mathrm{gn}}(x_1/\underline{m_1}, \ldots, x_n/\underline{m_n})\,,$$
$$\mathsf{PA} \vdash \ulcorner \varphi(x_1/\underline{m_1}, \ldots, x_n/\underline{m_n}) \urcorner = \lceil \varphi \rceil^{\mathrm{gn}}(x_1/\underline{m_1}, \ldots, x_n/\underline{m_n})\,.$$

For a proof see the Solution to Exercise 11.1.

Our next goal is to prove the following

Theorem 11.4. *If φ is an $\exists$-formula, then*

$$\mathsf{PA} \vdash \varphi \to \operatorname{prv}(\lceil \varphi \rceil^{\mathrm{gn}}). \tag{$*$}$$

Notice that for $\exists$-sentences φ, Theorem 11.4 implies D_3. In order to prove Theorem 11.4, we first need some auxiliary results whose proofs turn out to be quite technical. The following lemma essentially states that removing a variable x from V amounts to substituting in $\lceil \varphi \rceil_{V \setminus \{x\}}^{\mathrm{gn}}$ each occurrence of $\ulcorner x \urcorner$ by the term $\operatorname{gn}(x)$, thus obtaining $\lceil \varphi \rceil_V^{\mathrm{gn}}$. While this seems to be completely obvious, its proof is highly non-trivial, since it requires us to unravel all the details of the formalised substitution function. Before we prove the lemma, we first prove the following

Fact 11.5. $\mathsf{PA} \vdash \forall x \operatorname{term}(\operatorname{gn}(x))$.

Proof. The proof is by induction. Notice first that $\mathrm{gn}(\mathtt{0}) = \mathtt{0}$ and observe that $\mathsf{PA} \vdash \mathrm{term}(\mathtt{0})$ by taking the sequence $c_0 = \mathtt{0}$ with $\mathrm{lh}(c) = \mathtt{1}$. Now, by the definition of gn we may assume that $\mathrm{nat}(x, y)$ for some x and y. Thus, there exists c such that $\mathrm{c\_nat}(c, x, y)$, i.e., $\mathrm{seq}(c)$, $c_0 = \ulcorner \mathtt{0} \urcorner$, $c_x = y$ and $c_{\mathtt{s}i} = \mathrm{succ}(c_i)$ for every $i < x$. With this sequence, $\mathrm{c\_term}(c, y)$ holds, and hence we have $\mathrm{term}(y)$. On the other hand, $\mathrm{nat}(x, y)$ implies $\mathrm{gn}(x) = y$, and therefore we have $\mathrm{term}(\mathrm{gn}(x))$, and applying $(\forall)$ yields $\mathsf{PA} \vdash \forall x\big(\mathrm{term}(\mathrm{gn}(x))\big)$. ⊣

Lemma 11.6. *Let V be a finite set of variables, let $x \in V$, and suppose that τ is an $\mathscr{L}_{\mathsf{PA}}$-term and that φ is an $\mathscr{L}_{\mathsf{PA}}$-formula with $x \in \mathrm{free}(\varphi)$. Then we have:*

(a) $\mathsf{PA} \vdash \mathrm{sb\_term}(\ulcorner x \urcorner, \mathrm{gn}(x), \lceil \tau \rceil^{\mathrm{gn}}_{V \setminus \{x\}}, \lceil \tau \rceil^{\mathrm{gn}}_{V})$

(b) $\mathsf{PA} \vdash \mathrm{sb\_fml}(\ulcorner x \urcorner, \mathrm{gn}(x), \lceil \varphi \rceil^{\mathrm{gn}}_{V \setminus \{x\}}, \lceil \varphi \rceil^{\mathrm{gn}}_{V})$

Proof. We give a detailed proof of (a). Note that (b) is very similar, and since the proof is quite lengthy, we omit the proof of (b). A complete proof of both statements, in a slightly different context, is given in [54, Lem. 7.4–Lem. 7.6].

In order to prove (a), we proceed by induction on the construction of τ. The case when $\tau \equiv \mathtt{0}$ is trivial, since in that case, τ does not have any free variables. The other atomic case is when $\tau \equiv y$ is a variable. In that case, we need to distinguish between three possibilities: either $y \equiv x$ or, if $y \not\equiv x$, then either $y \in V$ or $y \notin V$.

Case 1. If $y \equiv x$, then $\lceil y \rceil_{V \setminus \{x\}} = \lceil y \rceil_{V \setminus \{y\}} = \ulcorner y \urcorner$, and, since $x \in V$, $\lceil y \rceil_V = y$, which implies $\lceil y \rceil^{\mathrm{gn}}_V = \mathrm{gn}(y)$. Then the claim follows, since by Fact 11.5 we have $\mathsf{PA} \vdash \mathrm{sb\_term}(\ulcorner y \urcorner, \mathrm{gn}(y), \ulcorner y \urcorner, \mathrm{gn}(y))$.

Case 2. If $y \not\equiv x$ and $y \in V$, then $\lceil y \rceil_{V \setminus \{x\}} = \lceil y \rceil_V = y$ and therefore $\lceil y \rceil^{\mathrm{gn}}_{V \setminus \{x\}} = \lceil y \rceil^{\mathrm{gn}}_V = \mathrm{gn}(y)$. Since the variable x does not appear in $\tau \equiv y$, there is nothing to substitute. More precisely, we have

$$\mathsf{PA} \vdash \mathrm{sb\_term}(\ulcorner x \urcorner, \mathrm{gn}(x), \mathrm{gn}(y), \mathrm{gn}(y))$$

as desired.

Case 3. Suppose that $y \not\equiv x$ and $y \notin V$. Then $\lceil y \rceil_{V \setminus \{x\}} = \lceil y \rceil_V = \ulcorner y \urcorner$, and therefore $\lceil y \rceil^{\mathrm{gn}}_{V \setminus \{x\}} = \lceil y \rceil^{\mathrm{gn}}_V = \ulcorner y \urcorner$, and since $\ulcorner y \urcorner$ does not have any free variables, the claim trivially holds.

Let us now consider the cases when τ is not atomic. Suppose that $\tau \equiv \mathtt{s}\tau'$ for some term τ'. Clearly, all variables in τ' are also among V. By our inductive assumption, we have

$$\mathsf{PA} \vdash \mathrm{sb\_term}(\ulcorner x \urcorner, \mathrm{gn}(x), \lceil \tau' \rceil^{\mathrm{gn}}_{V \setminus \{x\}}, \lceil \tau' \rceil^{\mathrm{gn}}_{V}).$$

Note that by definition of sb_term we have in general

$$\mathsf{PA} \vdash \mathrm{sb\_term}(v, t_0, t, t') \to \mathrm{sb\_term}(v, t_0, \mathrm{succ}(t), \mathrm{succ}(t')).$$

In our case, this means that if we set

$$v :\equiv \ulcorner x \urcorner\,, \quad t_0 :\equiv \mathrm{gn}(x)\,, \quad t :\equiv \lceil \tau' \rceil^{\mathrm{gn}}_{V\setminus\{x\}}\,, \quad t' :\equiv \lceil \tau' \rceil^{\mathrm{gn}}_{V}$$

in the above formula, then we have $\lceil \tau \rceil^{\mathrm{gn}}_{V\setminus\{x\}} \equiv \mathrm{succ}(t)$ and $\lceil \tau \rceil^{\mathrm{gn}}_{V} \equiv \mathrm{succ}(t')$, and therefore, τ satisfies (a).

The cases when $\tau \equiv \tau_1 + \tau_2$ or $\tau \equiv \tau_1 \cdot \tau_2$ are shown similarly. ⊣

Theorem 11.7. *For every $\mathscr{L}_{\mathsf{PA}}$-formula φ, we have:*

$$\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \varphi \urcorner) \to \mathrm{prv}(\lceil \varphi \rceil^{\mathrm{gn}})$$

Note that for sentences φ, Theorem 11.7 becomes trivial. On the other hand, if φ has free variables, then the statement still seems obvious, since it should not matter whether the free variables are gödelized at the same time as φ — as in the case of $\ulcorner \varphi \urcorner$ — or whether one gödelizes the formula such that the variables remain free, and afterwards substitutes the Gödel code of the variables — as in the case of $\lceil \varphi \rceil^{\mathrm{gn}}$. However, the proof is trickier than it might be expected, since one needs to use the properties of the formalised substitution function, which is, unfortunately, a very complicated function.

Proof of Theorem 11.7. Recall that by Fact 11.5 we have $\mathsf{PA} \vdash \forall x\, \mathrm{term}(\mathrm{gn}(x))$. Moreover, by induction on x we can also show

$$\mathsf{PA} \vdash \forall v \forall x \neg\, \mathrm{var\_in\_term}(v, \mathrm{gn}(x))\,,$$

which proves that the formalised substitution $v/\,\mathrm{gn}(x)$ is always admissible.

Note that if we have $\mathsf{PA} \vdash \varphi(\nu)$ for some variable ν, then we have $\mathsf{PA} \vdash \varphi(\nu/\tau)$ whenever τ is a term such that the substitution ν/τ is admissible:

φ_0:	$\varphi(\nu)$	by assumption
φ_1:	$\forall \nu \varphi(\nu)$	from φ_0 using $(\forall)$
φ_2:	$\forall \nu \varphi(\nu) \to \varphi(\nu/\tau)$	instance of L_{10}
φ_3:	$\varphi(\tau)$	from φ_2 and φ_1 using (MP)

Now, if we transfer this proof to the formalised level, by using Lemma 9.16, for the standard model $\mathbb{N}$ and some $c \in \mathbb{N}$ we have:

$$\text{If } \mathbb{N} \vDash \mathrm{c\_prv}(\underline{c}, \ulcorner \varphi \urcorner) \wedge \mathrm{sb\_fml}(\ulcorner \nu \urcorner, \ulcorner \tau \urcorner, \ulcorner \varphi \urcorner, \ulcorner \varphi(\tau) \urcorner) \text{ then } \mathbb{N} \vDash \mathrm{c\_prv}(\underline{c'}, \ulcorner \varphi(\tau) \urcorner)$$

where

$$c' :\equiv c^{\frown}\Big\langle \mathrm{all}(\ulcorner \nu \urcorner, \ulcorner \varphi \urcorner),\ \mathrm{imp}\big(\mathrm{all}(\ulcorner \nu \urcorner, \ulcorner \varphi \urcorner), \ulcorner \varphi(\tau) \urcorner\big),\ \mathrm{mp}\Big(\mathrm{all}(\ulcorner \nu \urcorner, \ulcorner \varphi \urcorner), \mathrm{imp}\big(\mathrm{all}(\ulcorner \nu \urcorner, \ulcorner \varphi \urcorner), \ulcorner \varphi(\tau) \urcorner\big)\Big)\Big\rangle\,.$$

By rewriting this implication syntactically, we obtain

$$\mathbb{N} \vDash \mathrm{c\_prv}(\underline{c}, \ulcorner \varphi \urcorner) \to \mathrm{c\_prv}(\underline{c}', \ulcorner \varphi(\tau) \urcorner)\,,$$

and therefore, by Corollary 9.3 we have

$$\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \varphi \urcorner) \to \mathrm{prv}(\ulcorner \varphi(\tau) \urcorner)\,.$$

Now, let φ be an arbitrary $\mathscr{L}_{\mathsf{PA}}$-formula. We may assume that all free variables of φ are among $v_0, \ldots, v_n$ for some $n \in \mathbb{N}$. Using the above observation together with Lemma 11.6, we obtain that $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \varphi \urcorner) \to \mathrm{prv}(\lceil \varphi \rceil^{\mathrm{gn}}_{\{v_0\}})$, and for each $k \in \{\mathbf{1}, \ldots, n\}$,

$$\mathsf{PA} \vdash \mathrm{prv}(\lceil \varphi \rceil^{\mathrm{gn}}_{\{v_0,\ldots,v_{k-1}\}}) \to \mathrm{prv}(\lceil \varphi \rceil^{\mathrm{gn}}_{\{v_0,\ldots,v_k\}}).$$

After finitely many applications of Tautology (D.0), we obtain

$$\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \varphi \urcorner) \to \mathrm{prv}\big(\lceil \varphi \rceil^{\mathrm{gn}}_{\{v_0,\ldots,v_n\}}\big)\,,$$

and since $\lceil \varphi \rceil^{\mathrm{gn}}_{\{v_0,\ldots,v_n\}} \equiv \lceil \varphi \rceil^{\mathrm{gn}}$, this completes the proof. ⊣

The following results are easy consequences of Theorem 11.7.

Corollary 11.8. *Let φ be an $\mathscr{L}_{\mathsf{PA}}$-formula. If $\mathsf{PA} \vdash \varphi$ then $\mathsf{PA} \vdash \mathrm{prv}(\lceil \varphi \rceil^{\mathrm{gn}})$.*

Proof. Note that from $\mathsf{PA} \vdash \varphi$ and $\mathsf{D_0}$ we obtain $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \varphi \urcorner)$, and by Theorem 11.7 we have $\mathsf{PA} \vdash \mathrm{prv}(\lceil \varphi \rceil^{\mathrm{gn}})$. ⊣

Corollary 11.9. *Let φ and ψ be arbitrary $\mathscr{L}_{\mathsf{PA}}$-formulae. Then $\mathsf{PA} \vdash \varphi \to \psi$ implies $\mathsf{PA} \vdash \mathrm{prv}(\lceil \varphi \rceil^{\mathrm{gn}}) \to \mathrm{prv}(\lceil \psi \rceil^{\mathrm{gn}})$. In particular, if $\varphi \Leftrightarrow_{\mathsf{PA}} \psi$ then $\mathrm{prv}(\lceil \varphi \rceil^{\mathrm{gn}}) \Leftrightarrow_{\mathsf{PA}} \mathrm{prv}(\lceil \psi \rceil^{\mathrm{gn}})$.*

Proof. Suppose that $\mathsf{PA} \vdash \varphi \to \psi$. An application of Corollary 11.8 yields

$$\mathsf{PA} \vdash \mathrm{prv}(\lceil \varphi \to \psi \rceil^{\mathrm{gn}}).$$

Recall that $\lceil \varphi \to \psi \rceil$ equals $\mathrm{imp}(\lceil \varphi \rceil, \lceil \psi \rceil)$, and therefore, $\lceil \varphi \to \psi \rceil^{\mathrm{gn}}$ equals $\mathrm{imp}(\lceil \varphi \rceil^{\mathrm{gn}}, \lceil \psi \rceil^{\mathrm{gn}})$. Moreover, by definition of formalised Modus Ponens we have

$$\mathsf{PA} \vdash \mathrm{mp}\big(\lceil \varphi \rceil^{\mathrm{gn}}, \lceil \varphi \to \psi \rceil^{\mathrm{gn}}, \lceil \psi \rceil^{\mathrm{gn}}\big)\,.$$

Hence, by Lemma 10.0.(a) we obtain

$$\mathsf{PA} \vdash \big(\mathrm{prv}(\lceil \varphi \rceil^{\mathrm{gn}}) \wedge \mathrm{prv}(\lceil \varphi \to \psi \rceil^{\mathrm{gn}})\big) \to \mathrm{prv}(\lceil \psi \rceil^{\mathrm{gn}})\,,$$

which completes the proof. ⊣

Now we have assembled all ingredients for the proof of Theorem 11.4:

Proof of Theorem 11.4. We have to show that for every $\exists$-formula φ_0,

$$\mathsf{PA} \vdash \varphi_0 \to \mathrm{prv}(\ulcorner \varphi_0 \urcorner^{\mathrm{gn}}). \tag{$*$}$$

Notice that by Corollary 11.9 it suffices to check $(*)$ for strict $\exists$-formulae φ_0, i.e., for formulae built up from atomic formulae and negated atomic formulae using $\wedge$, $\vee$, existential quantification $\exists\nu$ and bounded universal quantification $\forall\nu < \tau$ (for some term τ). We proceed by induction on the construction of formulae φ_0.

We start with atomic formulae. Since $\mathscr{L}_{\mathsf{PA}}$ does not contain relation symbols, we only have to consider atomic formulae of the form $\tau_i = \tau_j$ for some $\mathscr{L}_{\mathsf{PA}}$-terms τ_i and τ_j. Moreover, by substitution it suffices to show that each atomic formula of the form $v_i = v_j$, where v_i and v_j are variables, satisfies $(*)$. For example, if $\varphi_0 \equiv \mathsf{s0} + v_0 = \mathsf{s0} \cdot \mathsf{s0}$, then, for $\varphi \equiv v_1 = v_2$, we have $\varphi_0 \equiv \varphi\big(v_1/\mathsf{s0} + v_0,\, v_2/\mathsf{s0} \cdot \mathsf{s0}\big)$. Therefore, let us consider the formula $v_i = v_j$. First, note that we obviously have $\mathsf{PA} \vdash v_i = v_i$, and hence, by Corollary 11.8 we have

$$\mathsf{PA} \vdash \mathrm{prv}(\ulcorner v_i = v_i \urcorner^{\mathrm{gn}})\,.$$

Furthermore, since v_i and v_j are free variables in $\mathrm{prv}(\ulcorner v_i = v_j \urcorner^{\mathrm{gn}})$, we can use Exercise 2.7 to obtain

$$\mathsf{PA} \vdash \big(v_i = v_i \wedge v_i = v_j\big) \to \big(\mathrm{prv}(\ulcorner v_i = v_i \urcorner^{\mathrm{gn}}) \to \mathrm{prv}(\ulcorner v_i = v_j \urcorner^{\mathrm{gn}})\big)\,.$$

Putting these facts together and using logical axioms, tautologies and twice Modus Ponens, we obtain the following formal proof:

$$\begin{aligned}
\mathsf{PA} + v_i = v_j \;&\vdash v_i = v_i \\
&\vdash v_i = v_j \\
&\vdash v_i = v_i \wedge v_i = v_j \\
&\vdash \big(v_i = v_i \wedge v_i = v_j\big) \to \big(\mathrm{prv}(\ulcorner v_i = v_i \urcorner^{\mathrm{gn}}) \to \mathrm{prv}(\ulcorner v_i = v_j \urcorner^{\mathrm{gn}})\big) \\
&\vdash \mathrm{prv}(\ulcorner v_i = v_i \urcorner^{\mathrm{gn}}) \to \mathrm{prv}(\ulcorner v_i = v_j \urcorner^{\mathrm{gn}}) \\
&\vdash \mathrm{prv}(\ulcorner v_i = v_i \urcorner^{\mathrm{gn}}) \\
&\vdash \mathrm{prv}(\ulcorner v_i = v_j \urcorner^{\mathrm{gn}})
\end{aligned}$$

Therefore, by the Deduction Theorem we obtain

$$\mathsf{PA} \vdash v_i = v_j \to \mathrm{prv}(\ulcorner v_i = v_j \urcorner^{\mathrm{gn}})$$

as desired.

For negated atomic formulae, we only have to show that each formula of the form $v_i \neq v_j$ satisfies $(*)$. Now, we obviously have

$$v_i \neq v_j \;\Leftrightarrow_{\mathsf{PA}}\; (v_i < v_j) \vee (v_j < v_i)\,,$$

where

$$(v_i < v_j) \vee (v_j < v_i) \;\Leftrightarrow_{\mathsf{PA}}\; \exists v_k\big((v_k < v_j \wedge v_k = v_i) \vee (v_k < v_i \wedge v_k = v_j)\big)\,.$$

Hence, the case of negated atomic formulae follows from the fact that formulae of the form $\exists v_i \varphi$ satisfy $(*)$, which will be shown below.

Suppose now that φ satisfies $(*)$. We have to verify that $\varphi(v_i/\tau)$ (where the substitution v_i/τ is admissible), and that $\exists v_i \varphi$ and $\forall v_i < v_j\, \varphi$ satisfy $(*)$.

- Suppose that $v_i \in \mathrm{free}(\varphi)$ and that τ is an $\mathscr{L}_{\mathsf{PA}}$-term such that the substitution v_i/τ is admissible. We have to show that $\varphi(v_i/\tau)$ satisfies $(*)$. For the sake of simplicity, we assume that v_i is the only free variable of φ. By assumption we have $\mathsf{PA} \vdash \varphi \to \mathrm{prv}(\ulcorner\varphi\urcorner^{\mathrm{gn}})$. Now, using Generalisation we obtain

  $$\mathsf{PA} \vdash \forall v_i \left(\varphi \to \mathrm{prv}(\ulcorner\varphi\urcorner^{\mathrm{gn}})\right),$$

  and hence, by L_{10} and Modus Ponens we get

  $$\mathsf{PA} \vdash \varphi(\tau) \to \mathrm{prv}(\ulcorner\varphi\urcorner^{\mathrm{gn}})(v_i/\tau).$$

  Thus, it is enough to verify that $\mathsf{PA} \vdash \ulcorner\varphi\urcorner^{\mathrm{gn}}(v_i/\tau) = \ulcorner\varphi(\tau)\urcorner^{\mathrm{gn}}$. For this, we first prove that $\mathsf{PA} \vdash \mathrm{gn}(\tau) = \ulcorner\tau\urcorner^{\mathrm{gn}}$ by induction on the construction of the term τ: If $\tau \equiv \mathsf{0}$ then $\mathsf{PA} \vdash \ulcorner\mathsf{0}\urcorner^{\mathrm{gn}} = \ulcorner\mathsf{0}\urcorner = \ulcorner\mathsf{0}\urcorner = \mathsf{0} = \mathrm{gn}(\mathsf{0})$. The case when τ is a variable is similar. We now verify our claim for $\tau \equiv \mathtt{s}\tau'$ and leave the other cases as an exercise to the reader. By induction, we may assume that $\mathsf{PA} \vdash \mathrm{gn}(\tau') = \ulcorner\tau'\urcorner^{\mathrm{gn}}$. Then

  $$\begin{aligned}\mathsf{PA} &\vdash \ulcorner\mathtt{s}\tau'\urcorner^{\mathrm{gn}} = \mathrm{succ}(\ulcorner\tau'\urcorner^{\mathrm{gn}})\\ &\vdash \mathrm{succ}(\ulcorner\tau'\urcorner^{\mathrm{gn}}) = \mathrm{succ}(\mathrm{gn}(\tau'))\\ &\vdash \mathrm{succ}(\mathrm{gn}(\tau')) = \mathrm{gn}(\mathtt{s}\tau')\,,\end{aligned}$$

  and by transitivity of the relation $=$ we have $\mathsf{PA} \vdash \ulcorner\mathtt{s}\tau'\urcorner^{\mathrm{gn}} = \mathrm{gn}(\mathtt{s}\tau')$ as desired.

  As a consequence of $\mathsf{PA} \vdash \mathrm{gn}(\tau) = \ulcorner\tau\urcorner^{\mathrm{gn}}$, we obtain

  $$\begin{aligned}\mathsf{PA} &\vdash \ulcorner\varphi\urcorner^{\mathrm{gn}}(v_i/\tau) = \ulcorner\varphi\urcorner(v_i/\mathrm{gn}(v_i))(v_i/\tau)\\ &\vdash \ulcorner\varphi\urcorner(v_i/\mathrm{gn}(v_i))(v_i/\tau) = \ulcorner\varphi\urcorner(v_i/\mathrm{gn}(\tau))\\ &\vdash \ulcorner\varphi\urcorner(v_i/\mathrm{gn}(\tau)) = \ulcorner\varphi\urcorner(v_i/\ulcorner\tau\urcorner^{\mathrm{gn}})\\ &\vdash \ulcorner\varphi\urcorner(v_i/\ulcorner\tau\urcorner^{\mathrm{gn}}) = \ulcorner\varphi(\tau)\urcorner^{\mathrm{gn}}\,,\end{aligned}$$

  and by transitivity of $=$ we obtain $\mathsf{PA} \vdash \ulcorner\varphi\urcorner^{\mathrm{gn}}(v_i/\tau) = \ulcorner\varphi(\tau)\urcorner^{\mathrm{gn}}$ as desired.

- Under the assumption

$$\mathsf{PA} \vdash \varphi \to \mathrm{prv}(\ulcorner \varphi \urcorner^{\mathrm{gn}})$$

we show that $\forall v_i < v_j\, \varphi$ (where $i \not\equiv j$) satisfies $(*)$. Let v'_j be a variable which does not occur in φ. Since

$$\forall v_i < v_j\, \varphi \;\Leftrightarrow_{\mathsf{PA}}\; \exists v'_j \left(v'_j = v_j \wedge \forall v_i < v'_j\, \varphi\right),$$

we may assume without loss of generality that v_j does not occur in φ. Furthermore, let $\psi(v_j)$ denote the formula

$$\forall v_i < v_j\, \varphi \to \mathrm{prv}(\ulcorner \forall v_i < v_j\, \varphi \urcorner^{\mathrm{gn}}).$$

It suffices to show that $\mathsf{PA} \vdash \forall v_j\, \psi(v_j)$. So, by PA_6 it is enough to show that $\mathsf{PA} \vdash \psi(\mathsf{0})$, and for all v_j, $\mathsf{PA} \vdash \psi(v_j) \to \psi(\mathsf{s}v_j)$.
Notice that, since $\forall v_i < \mathsf{0}\, \varphi$ is a tautology, by Corollary 11.8 we have $\mathsf{PA} \vdash \psi(\mathsf{0})$. For the induction step, assume that $\mathsf{PA} \vdash \psi(v_j)$. Recall that by Lemma 8.5 we have

$$v_i < \mathsf{s}v_j \;\Leftrightarrow_{\mathsf{PA}}\; v_i < v_j \vee v_i = v_j\,,$$

and hence,

$$\forall v_i < \mathsf{s}v_j\, \varphi \;\Leftrightarrow_{\mathsf{PA}}\; \forall v_i < v_j\, \varphi \wedge \varphi(v_i/v_j)\,,$$

where the substitution v_i/v_j is admissible because v_j does not occur in φ. Since $\mathsf{PA} \vdash \psi(v_j)$, by Lemma 10.0.(b) it suffices to show that

$$\mathsf{PA} \vdash \varphi(v_i/v_j) \to \mathrm{prv}\big(\ulcorner \varphi(v_i/v_j) \urcorner^{\mathrm{gn}}\big)\,,$$

which follows from the previous case, using our assumption that φ satisfies $(*)$.

- Now, we show that $\exists v_i \varphi$ satisfies $(*)$. Since by L_{11}, $\mathsf{PA} \vdash \varphi \to \exists v_i \varphi$, we can apply Corollary 11.9 and obtain

$$\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \varphi \urcorner^{\mathrm{gn}}) \to \mathrm{prv}(\ulcorner \exists v_i \varphi \urcorner^{\mathrm{gn}}).$$

Therefore, by Tautology (D.0) we have

$$\mathsf{PA} \vdash \varphi \to \mathrm{prv}(\ulcorner \exists v_i \varphi \urcorner^{\mathrm{gn}}),$$

and by Generalisation we obtain

$$\mathsf{PA} \vdash \forall v_i \left(\varphi \to \mathrm{prv}(\ulcorner \exists v_i \varphi \urcorner^{\mathrm{gn}})\right).$$

Now, since v_i does not occur as a free variable in $\mathrm{prv}(\ulcorner \exists v_i \varphi \urcorner^{\mathrm{gn}})$, by L_{13} and Modus Ponens we finally obtain

$$\mathsf{PA} \vdash \exists v_i \varphi \to \mathrm{prv}(\ulcorner \exists v_i \varphi \urcorner^{\mathrm{gn}})$$

as desired.

Finally, suppose that φ and ψ both satisfy $(*)$. We have to show that $\varphi \wedge \psi$ and $\varphi \vee \psi$ also satisfy $(*)$.

- In order to see that $\varphi \wedge \psi$ satisfies $(*)$, notice first that

$$\mathrm{and}\big(\ulcorner \varphi \urcorner^{\mathrm{gn}}, \ulcorner \psi \urcorner^{\mathrm{gn}}\big) = \ulcorner \varphi \wedge \psi \urcorner^{\mathrm{gn}}.$$

  Therefore, by Lemma 10.0 we have

$$\mathsf{PA} \vdash \mathrm{prv}\big(\ulcorner \varphi \urcorner^{\mathrm{gn}}\big) \wedge \mathrm{prv}\big(\ulcorner \psi \urcorner^{\mathrm{gn}}\big) \to \mathrm{prv}\big(\ulcorner \varphi \wedge \psi \urcorner^{\mathrm{gn}}\big).$$

  Using our assumption that φ and ψ both satisfy $(*)$, it follows that $\varphi \wedge \psi$ also satisfies $(*)$.
- The case $\varphi \vee \psi$ is similar and thus left as an exercise to the reader.

⊣

Concluding Remarks

To summarise, we have shown that $\mathsf{PA} \nvdash \mathrm{con}_{\mathsf{PA}}$, where

$$\mathrm{con}_{\mathsf{PA}} \equiv \neg\, \mathrm{prv}(\ulcorner \mathsf{0} = \mathsf{1} \urcorner).$$

In other words, we have shown that

$$\mathsf{PA} \nvdash \neg\, \mathrm{prv}(\ulcorner \mathsf{0} = \mathsf{1} \urcorner).$$

Now, by the Completeness Theorem we know that there exists a model $\mathbf{M} \models \mathsf{PA}$ such that

$$\mathbf{M} \models \mathrm{prv}(\ulcorner \mathsf{0} = \mathsf{1} \urcorner).$$

Since $\mathsf{PA} \vdash \neg(\mathsf{0} = \mathsf{1})$, we have $\mathbf{M} \models \neg(\mathsf{0} = \mathsf{1})$, which shows that

$$\mathbf{M} \models \neg(\mathsf{0} = \mathsf{1}) \wedge \mathrm{prv}(\ulcorner \mathsf{0} = \mathsf{1} \urcorner).$$

A slightly more general result can be obtained from Löb's Theorem.

Löb's Theorem

Recall that by Lemma 9.16 we have, for every $\mathscr{L}_{\mathsf{PA}}$-formula φ, $\mathbb{N} \models \mathrm{prv}(\#\varphi)$ if and only if $\mathsf{PA} \vdash \varphi$. In particular, this implies that $\mathbb{N} \models \mathrm{prv}(\#\varphi) \to \varphi$. In other words, in the standard model $\mathbb{N}$, each "provable" formula is true, where

"provable" in $\mathbb{N}$ is meant with respect to the provability predicate prv. This applies because in $\mathbb{N}$ one can retrieve the proof of φ from its code. Another consequence of $\mathbb{N} \vDash \mathrm{prv}(\#\varphi) \to \varphi$ is that for every $\mathscr{L}_{\mathsf{PA}}$-formula φ,

$$\mathbb{N} \nvDash \neg\varphi \wedge \mathrm{prv}(\#\varphi)\,.$$

This leads to the natural question whether this is true in any model of PA. Löb's Theorem gives a negative answer to this question.

Theorem 11.10 (Löb's Theorem). *Suppose that φ is an $\mathscr{L}_{\mathsf{PA}}$-sentence. Then* $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner\varphi\urcorner) \to \varphi$ *implies* $\mathsf{PA} \vdash \varphi$.

Proof. Assume that $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner\varphi\urcorner) \to \varphi$ and let σ be an $\mathscr{L}_{\mathsf{PA}}$-sentence such that

$$\sigma \Leftrightarrow_{\mathsf{PA}} \mathrm{prv}(\ulcorner\sigma\urcorner) \to \varphi\,.$$

In order to see that such a sentence σ exists, let $\psi(v_0) :\equiv \mathrm{prv}(v_0) \to \varphi$. Then by the Diagonalisation Lemma, there exists an $\mathscr{L}_{\mathsf{PA}}$-sentence σ such that $\sigma \Leftrightarrow_{\mathsf{PA}} \psi(\ulcorner\sigma\urcorner)$.

Claim. $\mathsf{PA} \vdash \sigma$, *or equivalently,* $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner\sigma\urcorner) \to \varphi$.

Proof of Claim. By our assumption we have $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner\varphi\urcorner) \to \varphi$. Therefore, it suffices to check that $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner\sigma\urcorner) \to \mathrm{prv}(\ulcorner\varphi\urcorner)$. Note that by Corollary 10.2.(a) we have $\mathrm{prv}(\ulcorner\sigma\urcorner) \Leftrightarrow_{\mathsf{PA}} \mathrm{prv}(\ulcorner\mathrm{prv}(\ulcorner\sigma\urcorner) \to \varphi\urcorner)$. Moreover, D_1 implies

$$\mathsf{PA} \vdash \mathrm{prv}\big(\ulcorner\mathrm{prv}(\ulcorner\sigma\urcorner) \to \varphi\urcorner\big) \to \Big(\mathrm{prv}\big(\ulcorner\mathrm{prv}(\ulcorner\sigma\urcorner)\urcorner\big) \to \mathrm{prv}(\ulcorner\varphi\urcorner)\Big)\,.$$

Now, if we assume $\mathrm{prv}(\ulcorner\sigma\urcorner)$, then by D_2 we obtain $\mathrm{prv}(\ulcorner\mathrm{prv}(\ulcorner\sigma\urcorner)\urcorner)$, and by the preceding observations, Modus Ponens, and the Deduction Theorem, we finally obtain $\mathrm{prv}(\ulcorner\varphi\urcorner)$. $\dashv_{\text{Claim}}$

Using the above claim, we have $\mathsf{PA} \vdash \sigma$ and therefore, we get $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner\sigma\urcorner)$ by D_0. So, by $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner\sigma\urcorner) \to \varphi$ and Modus Ponens we get $\mathsf{PA} \vdash \varphi$ as desired. $\dashv$

Löb's Theorem has some remarkable consequences. For example, if we use the Diagonalisation Lemma to obtain a sentence σ such that $\sigma \Leftrightarrow_{\mathsf{PA}} \mathrm{prv}(\ulcorner\sigma\urcorner)$, then it follows that $\mathsf{PA} \vdash \sigma$. Hence, if we replace the sentence stating "I am unprovable" by "I am provable" — the so-called *truth-teller sentence* — then this does not yield an undecidable statement.

Löb's Theorem also implies that if $\mathsf{PA} \nvdash \varphi$ for some $\mathscr{L}_{\mathsf{PA}}$-formula φ, then $\mathsf{PA} \nvdash \mathrm{prv}(\ulcorner\varphi\urcorner) \to \varphi$, i.e., $\mathsf{PA} \nvdash \varphi \vee \neg\,\mathrm{prv}(\ulcorner\varphi\urcorner)$. This illustrates the difference between truth and provability in non-standard models of PA: For every formula φ with $\mathsf{PA} \nvdash \varphi$, there are models $\mathbf{M}$ such that $\mathbf{M} \vDash \neg\varphi \wedge \mathrm{prv}(\ulcorner\varphi\urcorner)$.

In **M**, the code of the "proof" of φ is of non-standard length and does henceforth not code an actual proof of φ. In this sense, Gödel's First Incompleteness Theorem can be viewed as a special case of Löb's Theorem, by taking φ to be the formula σ with $\sigma \Leftrightarrow_{\mathsf{PA}} \neg \mathrm{prv}(\ulcorner \sigma \urcorner)$.

Notes

The question of whether arithmetic can be shown to be consistent was the second of Hilbert's famous list [26] of 23 open problems of mathematics. While Gödel's Second Incompleteness Theorem published in [16] in 1931 gives a negative answer to Hilbert's second problem, Gentzen [13] provided in 1936 a consistency proof of PA in primitive recursive arithmetic with the additional principle of quantifier-free transfinite induction up to the ordinal number ε_0. The proof of the Second Incompleteness Theorem using pseudo-coding which we presented here follows Świerczkowski [54]. However, Świerczkowski worked in the theory of hereditarily finite sets, which is equivalent to PA. His proof was actually formalised and proof-checked using the interactive theorem prover Isabelle by Paulson [41] in 2013. The derivability conditions D_0–D_2, although already used by Gödel, were first introduced by Hilbert and Bernays [28] and re-formulated in its current form by Löb [32]. In the same paper, Löb also proved his theorem as an answer to a question posed by Henkin [24] in 1952.

Let us also say a few words about the impact of the Second Incompleteness Theorem on the so-called Hilbert's programme. In the early 1920s, Hilbert presented a new approach to the foundations of classical mathematics, which became known as Hilbert's programme. The aim of this programme was to find a formalisation of all mathematics in an axiomatic form and to prove that this axiomatisation of mathematics is consistent. The proof of consistency itself should only be carried out with what Hilbert called *finitary* methods. Even though the the work on this programme made considerable progress in the 1920s with contributions from numerous logicians, Gödel's Second Incompleteness Theorem shows that Hilbert's programme must fail. The crucial point is that we cannot formally define the notion of finiteness, and in particular of *finitary* methods, and as a consequence we get that we cannot even show the consistency of PA within PA — unless PA is inconsistent. However, despite the failure of Hilbert's programme, the framework of Hilbert's programme with notions such as *actual mathematics*, *formal mathematics*, and *metamathematics*, was very fruitfull and still influences the philosophy of mathematics (see Halbeisen [21]).

Exercises

11.0 Give an alternative proof of the Second Incompleteness Theorem by using Löb's Theorem.

11.1 Prove Fact 11.3.

Hint: Use induction on term and formula construction, respectively.

11.2 Prove that all formulae of the form $v_i + v_j = v_k$ and $v_i \cdot v_j = v_k$ satisfy $(*)$ in Theorem 11.4.

11.3 Prove that if φ and ψ satisfy $(*)$ in Theorem 11.4, then so does the disjunction $\varphi \vee \psi$.

11.4 Let φ be an $\mathscr{L}_{\mathsf{PA}}$-formula.

(a) Show that the formalisation of Löb's Theorem is provable within PA, i.e., show that

$$\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \mathrm{prv}(\ulcorner \varphi \urcorner) \to \varphi \urcorner) \to \mathrm{prv}(\ulcorner \varphi \urcorner)\,.$$

Hint: Set $\psi :\equiv \mathrm{prv}(\ulcorner \mathrm{prv}(\ulcorner \varphi \urcorner) \to \varphi \urcorner) \to \mathrm{prv}(\ulcorner \varphi \urcorner)$ and prove $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \psi \urcorner) \to \psi$.

(b) Prove $\mathsf{PA} \vdash \neg\, \mathrm{prv}(\ulcorner \varphi \urcorner) \to \neg\, \mathrm{prv}(\ulcorner \neg\, \mathrm{prv}(\ulcorner \varphi \urcorner) \urcorner)$.

(c) Conclude that the Second Incompleteness Theorem is provable within PA.

11.5 Use Exercise 11.4 to prove the following generalisation of Löb's Theorem: For all $\mathscr{L}_{\mathsf{PA}}$-formulae φ and ψ,

$$\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \varphi \urcorner) \wedge \big(\mathrm{prv}(\ulcorner \psi \urcorner) \to \psi\big) \quad \Longrightarrow\!\!\!\!\Longrightarrow \quad \mathsf{PA} \vdash \mathrm{prv}(\ulcorner \varphi \urcorner) \to \psi.$$

11.6 Prove that for every $\mathscr{L}_{\mathsf{PA}}$-formula φ,

$$\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \varphi \leftrightarrow \mathrm{con}_{\mathsf{PA}} \urcorner) \to \big(\mathrm{prv}(\ulcorner \varphi \urcorner) \leftrightarrow \neg\, \mathrm{con}_{\mathsf{PA}}\big)\,,$$

and interpret this result in the standard model.

11.7 Let $\mathrm{con}^{R}_{\mathsf{PA}}$ denote the formula $\neg\, \mathrm{prv}^{\mathrm{R}}(\ulcorner \mathsf{0} = \mathsf{1} \urcorner)$, i.e., the formula obtained from $\mathrm{con}_{\mathsf{PA}}$ by replacing the provability predicate prv by $\mathrm{prv}^{\mathrm{R}}$.

Show that $\mathsf{PA} \vdash \mathrm{con}^{R}_{\mathsf{PA}}$. Note that this implies that $\mathrm{prv}^{\mathrm{R}}$ satisfies neither D_1 nor D_2.

Chapter 12
Completeness of Presburger Arithmetic

In Chapter 10, we have seen that Peano Arithmetic PA is incomplete. Moreover, if we omit the Induction Schema and replace it by the axiom $\forall x(x = 0 \vee \exists y(x = \mathsf{s}y))$, stating that every number is either 0 or has a predecessor, then the resulting theory, called Robinson Arithmetic, is also incomplete. There are, however, other natural ways to weaken the axioms of PA: One could, for example, drop one of the function symbols $+$ or $\cdot$ as well as the corresponding axioms. In the former case, this leads to **Skolem Arithmetic**, and in the latter case to **Presburger Arithmetic**. In this chapter, we will only consider Presburger Arithmetic, denoted by PrA.

In the proof of the First Incompleteness Theorem, we introduced the β-function which allows to express exponentiation in terms of addition and multiplication. A natural question that arises in this context is whether multiplication might already be expressible in terms of the successor function and addition. If this is the case, then we can carry out the proof of the First Incompleteness Theorem in PrA, and obtain that PrA is incomplete. However, we will prove below that PrA is complete, which implies that we cannot express multiplication in terms of addition and successors.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_13

Basic Arithmetic in Presburger Arithmetic

As already mentioned above, the language of Presburger Arithmetic PrA is given by $\mathscr{L}_{\mathsf{PrA}} = \{0, \mathtt{s}, +\}$, where, as in $\mathscr{L}_{\mathsf{PA}}$, 0 is a constant symbol, $\mathtt{s}$ is a unary function symbol, and $+$ is a binary function symbol. The axioms of PrA are simply given by the axioms of PA except PA_4 and PA_5. More precisely, the axioms of PrA are

PA_0: $\neg\exists x(\mathtt{s}x = 0)$
PA_1: $\forall x\forall y(\mathtt{s}x = \mathtt{s}y \to x = y)$
PA_2: $\forall x(x + 0 = x)$
PA_3: $\forall x\forall y(x + \mathtt{s}y = \mathtt{s}(x + y))$

together with the Induction Schema, i.e., if φ is an $\mathscr{L}_{\mathsf{PrA}}$-formula such that $x \in \mathrm{free}(\varphi)$, then

PA_6: $\big(\varphi(0) \wedge \forall x(\varphi(x) \to \varphi(\mathtt{s}(x)))\big) \to \forall x\varphi(x)$.

Presburger, who first investigated PrA and proved its completeness, originally axiomatised the theory in a distinct manner: For example, he did not use the Induction Schema, but also postulated the existence of negative numbers and hence subtraction. In particular, he included the axiom $\forall x\forall y\exists z(x + z = y)$.

Clearly, in PrA one can prove all standard results of arithmetic which do not involve multiplication. In particular, we can define the relations $<$ and $\leq$ in the same way as in PA. Furthermore, as in Peano Arithmetic, we are able to define the terms $\underline{n}$ for all $n \in \mathbb{N}$, where the terms $\underline{n}$ are called *natural numbers*. Moreover, we can prove the properties N_0-N_5 stated in Proposition 9.0.

In the subsequent sections, it will become clear that it is impossible to define multiplication using addition and the successor function. However, it is possible to define the multiplication with a natural number of the form $\underline{n}$ for $n \in \mathbb{N}$. Using the Recursion Principle defined in Chapter 0, we define

$$\underline{\mathbf{0}} \cdot x :\equiv 0, \text{ and}$$
$$\underline{n + \mathbf{1}} \cdot x :\equiv \underline{n} \cdot x + x.$$

Note that for the sake of simplicity, we usually write $\underline{n}x$ rather than $\underline{n} \cdot x$.

Lemma 12.0. *Let $n, m \in \mathbb{N}$. Multiplication with natural numbers satisfies the associativity and distributivity laws:*

$$\mathsf{PrA} \vdash \forall x\forall y(\underline{n}(x + y) = \underline{n}x + \underline{n}y)$$
$$\mathsf{PrA} \vdash \forall x(\underline{m + n}x = \underline{m}x + \underline{n}x)$$
$$\mathsf{PrA} \vdash \forall x(\underline{mn}x = \underline{m} \cdot (\underline{n}x))$$

Furthermore, multiplication with natural numbers respects the two binary relations $=$ and $<$:

$$\mathsf{PrA} \vdash \forall x \forall y (\underline{n}x = \underline{n}y \leftrightarrow x = y)$$
$$\mathsf{PrA} \vdash \forall x \forall y (\underline{n}x < \underline{n}y \leftrightarrow x < y)$$

Proof. The proof is similar to the proof of Proposition 9.0 and uses metainduction in $\mathbb{N}$. We only prove the first statement, since all proofs are similar. For $n = \mathbf{0}$, we obviously have $\mathsf{0}(x + y) = \mathsf{0} = \mathsf{0} \cdot x + \mathsf{0} \cdot y$. Suppose now that the claim holds for some $n \in \mathbb{N}$. Then

$$\begin{aligned} \underline{n+\mathbf{1}}(x+y) &= \underline{n}(x+y) + (x+y) = (\underline{n}x + \underline{n}y) + (x+y) \\ &= (\underline{n}x + x) + (\underline{n}y + y) = \underline{n+\mathbf{1}} \cdot x + \underline{n+\mathbf{1}} \cdot y, \end{aligned}$$

where the second equality follows from our induction hypothesis. ⊣

For $n \in \mathbb{N}$ with $n \geq \mathbf{2}$, we define

$$x \equiv_n y :\Longleftrightarrow \exists z \big(\underline{n}z + x = y \ \vee \ \underline{n}z + y = x\big).$$

Furthermore, we abbreviate $\neg(x \equiv_n y)$ by $x \not\equiv_n y$. Formulae of the form $x \equiv_n y$ are called **congruences**. It is straightforward to check that congruences are — on the formal level — equivalence relations, i.e.,

$$\mathsf{PrA} \vdash \forall x (x \equiv_n x)\,,$$
$$\mathsf{PrA} \vdash \forall x \forall y (x \equiv_n y \leftrightarrow y \equiv_n x)\,,$$
$$\mathsf{PrA} \vdash \forall x \forall y \forall x \big(x \equiv_n y \wedge y \equiv_n z \ \rightarrow \ x \equiv_n z\big)\,.$$

In fact, as the name already suggests, they define congruence relations with respect to $+$:

$$\mathsf{PrA} \vdash \forall x \forall y \forall z \big(x \equiv_n y \ \leftrightarrow \ x + z \equiv_n y + z\big)\,.$$

The following result is a version of division with remainder, where the divisor is in $\mathbb{N}$.

Lemma 12.1. *For every natural number $n \geq \mathbf{2}$, we have*

$$\mathsf{PrA} \vdash \forall x \exists y \Big(\bigvee_{k=\mathbf{0}}^{n-\mathbf{1}} \underline{n}y + \underline{k} = x\Big).$$

In particular,

$$\mathsf{PrA} \vdash \forall x \Big(\bigvee_{k=\mathbf{0}}^{n-\mathbf{1}} x \equiv_n \underline{k}\Big).$$

Proof. We proceed by induction on x. For $x = \mathtt{0}$, the statement is trivial. Suppose that $x = \underline{n}y + \underline{k}$ for some $k < n$ and some y. Then $\mathtt{s}x = \mathtt{s}(\underline{n}y + \underline{k}) = \underline{n}y + \underline{k} + \mathbf{1}$. Now, if $k + \mathbf{1} < n$, we are done. Otherwise, $k + \mathbf{1} = n$ and hence $x = \underline{n}y + \underline{n} = \underline{n}(y + \mathtt{1})$. ⊣

Quantifier Elimination

The idea for proving that PrA is complete, consists of proving, in a language extension of $\mathscr{L}_{\mathsf{PrA}}$, that every sentence is logically equivalent to a quantifier-free one. Such sentences can easily be shown to be $\mathbb{N}$-conform, and hence, they can be either proven or disproven, depending on whether they are satisfied in $\mathbb{N}$. In this section, we will prove a more general result, which we will then apply to PrA. We say that a theory $\mathbf{\Phi}$ in some language $\mathscr{L}$ **admits quantifier elimination**, if for every $\mathscr{L}$-formula φ there is a quantifier-free formula ψ such that

$$\mathbf{\Phi} \vdash \varphi \leftrightarrow \psi\,.$$

The key point is to note that in order to prove that a theory admits quantifier elimination, it suffices to check that a single existential quantifier can be eliminated.

Theorem 12.2 (Quantifier Elimination Theorem). *Let $\mathbf{\Phi}$ be a theory in some language $\mathscr{L}$ such that the following holds:*

(a) *If φ is an atomic $\mathscr{L}$-formula, then $\neg\varphi$ is logically equivalent to a disjunction of conjunctions of atomic $\mathscr{L}$-formulae.*

(b) *For every $\mathscr{L}$-formula φ of the form $\varphi \equiv \exists\nu(\varphi_1 \wedge \ldots \wedge \varphi_n)$, where each φ_i is an atomic $\mathscr{L}$-formula, there is a quantifier-free $\mathscr{L}$-formula ψ with $\mathrm{free}(\psi) = \mathrm{free}(\varphi_1 \wedge \ldots \wedge \varphi_n) \setminus \{\nu\}$ such that $\mathbf{\Phi} \vdash \varphi \leftrightarrow \psi$.*

Then $\mathbf{\Phi}$ admits quantifier elimination.

Proof. The following steps show how to transform any $\mathscr{L}$-formula into an equivalent quantifier-free one using the above statements (a) and (b).

Step 1. Using Theorem 2.14, we can transform any $\mathscr{L}$-sentence into an $\mathscr{L}$-sentence in PNF.

Step 2. Using Tautology (Q.0), we can eliminate all universal quantifiers, i.e., every $\mathscr{L}$-formula in PNF is equivalent to an $\mathscr{L}$-formula of the form

$$(\neg)\exists\nu_1 \ldots (\neg)\exists\nu_n\varphi,$$

where $\nu_1, \ldots, \nu_n$ are variables and φ is quantifier-free.

Step 3. Given an $\mathscr{L}$-sentence of the form $(\neg)\exists\nu_1 \ldots (\neg)\exists\nu_n\varphi$ as above, one can transform the quantifier-free part φ into DNF by the Disjunctive Normal Form Theorem, i.e., all of the conjuncts are atomic or negated atomic formulae. Moreover, using (a) we can replace each negated atomic formula by a disjunction of conjunctions of atomic formulae and can thus transform the quantifier-free part into DNF in such a way that all conjuncts are atomic.

Step 4. From Step 3 we obtain an $\mathscr{L}$-formula of the form

$$(\neg)\exists\nu_1 \ldots (\neg)\exists\nu_n\Big(\big(\varphi_{1,1} \wedge \ldots \wedge \varphi_{1,k_1}\big) \vee \cdots \vee \big(\varphi_{m,1} \wedge \cdots \wedge \varphi_{m,k_m}\big)\Big),$$

where each $\varphi_{i,j}$ is an atomic or negated atomic $\mathscr{L}$-formula. Using Tautology (U.2) this is equivalent to

$$(\neg)\exists\nu_1 \ldots (\neg)\exists\nu_{n-1}(\neg)\big(\exists\nu_n(\varphi_{1,1}\wedge\ldots\wedge\varphi_{1,k_1})\vee\cdots\vee\exists\nu_n(\varphi_{m,1}\wedge\cdots\wedge\varphi_{m,k_m})\big).$$

Now using (a) and (b), each of the formulae $\exists\nu_n(\varphi_{i,1}\wedge\cdots\wedge\varphi_{i,k_i})$ is equivalent to a corresponding quantifier-free $\mathscr{L}$-formula ψ_i.

Step 5. In the case that there is a negation symbol $\neg$ in front of the existential quantifier $\exists\nu_n$, we use (a) to eliminate the negation, and then return to Step 3 in order to restore the DNF.

Steps 3–5 have to be repeated finitely many times until no more quantifiers are left. Thus, the above described algorithm yields a quantifier-free disjunction of conjunctions of almost atomic formulae, as desired. ⊣

Note that one could simplify Theorem 12.2 by omitting (a) and requiring instead in (b) that each φ_i is either atomic or the negation of an atomic formula.

Completeness of Presburger Arithmetic

We will now show that PrA is complete. Using the previous result, one might be tempted to first show that PrA admits quantifier elimination and then show that quantifier-free $\mathscr{L}_{\mathsf{PrA}}$-sentences can be either proven or disproven. While the second step will be verified in Lemma 12.7, the first one is not possible: An example is the $\mathscr{L}_{\mathsf{PrA}}$-formula

$$\exists y(x = \mathsf{2}y),$$

stating that $x \equiv_{\mathsf{2}} \mathsf{0}$ is not equivalent to a quantifier-free formula; another example for an $\mathscr{L}_{\mathsf{PA}}$-formula which is not equivalent to a quantifier-free formula is the formula $\exists z(x + z = y)$, stating that $x \leq y$ (see Exercise 12.0).

However, this problem can be overcome by extending the language $\mathscr{L}_{\mathsf{PrA}}$ to admit the binary relations $<$ and $\equiv_m$ for $m \in \mathbb{N}$: Let $\mathscr{L}_{\mathsf{PrA}^*}$ denote the language $\mathscr{L}_{\mathsf{PrA}} \cup \{<\} \cup \{\equiv_m |\ m \in \mathbb{N}\}$. By Theorem 6.0 we get that the

completeness of PrA with respect to $\mathscr{L}_{\mathsf{PrA}}$ is equivalent to the completeness of PrA with respect to the extended language $\mathscr{L}_{\mathsf{PrA}^*}$.

In the following paragraphs, we will show that PrA admits quantifier elimination with respect to $\mathscr{L}_{\mathsf{PrA}^*}$. The first step is to introduce a normal form for equations and congruences with respect to a fixed variable:

LEMMA 12.3. *Let ν be a variable. Then, for $n, m \in \mathbb{N}$, every atomic $\mathscr{L}_{\mathsf{PrA}^*}$-formula is logically equivalent to a formula of the form*

$$\underline{n}\nu + \tau = \tau', \quad \underline{n}\nu + \tau \equiv_m \tau', \quad \underline{n}\nu + \tau < \tau', \quad \tau' < \underline{n}\nu + \tau,$$

where ν does not occur in τ, τ'.

Proof. Since all cases are similar, we may assume that φ is an equation. We prove by induction on the term construction that for every $\mathscr{L}_{\mathsf{PrA}^*}$-term τ there exist $n \in \mathbb{N}$ and an $\mathscr{L}_{\mathsf{PrA}^*}$-term τ' such that $\mathsf{PrA} \vdash \tau = \underline{n}\nu + \tau'$.

- Suppose first that τ is atomic. If $\tau \equiv \mathsf{0}$, then obviously $\mathsf{PrA} \vdash \tau = \underline{\mathsf{0}}\nu + \mathsf{0}$. If $\tau \equiv \nu$, then $\mathsf{PrA} \vdash \tau = \underline{\mathbf{1}}\nu + \mathsf{0}$, and if $\tau \equiv w$ for some variable $w \not\equiv \nu$, then we can set $n \equiv \mathbf{0}$ and $\tau' \equiv w$.
- Assume now that $\tau \equiv \mathsf{s}\tau'$. By induction, we may assume that $\mathsf{PrA} \vdash \tau' = \underline{n}\nu + \tau''$ for some $n \in \mathbb{N}$ and some $\mathscr{L}_{\mathsf{PrA}^*}$-term τ'' such that ν does not occur in τ''. Then $\mathsf{PrA} \vdash \tau = \mathsf{s}\tau' = \mathsf{s}(\underline{n}\nu + \tau'') = \underline{n}\nu + \mathsf{s}\tau''$ by PA_3.
- Finally, let $\tau \equiv \tau_1 + \tau_2$, where $\mathsf{PrA} \vdash \tau_1 = \underline{n}\nu + \tau_1'$ and $\mathsf{PrA} \vdash \tau_2 = \underline{m}\nu + \tau_2'$, where $n, m \in \mathbb{N}$ and τ_1', τ_2' are terms such that ν does not occur in τ_1' and τ_2'. Then $\mathsf{PrA} \vdash \tau = \tau_1 + \tau_2 = (\underline{n}\nu + \tau_1') + (\underline{m}\nu + \tau_2') = \underline{n + m}\nu + \tau'$, where $\tau' \equiv \tau_1' + \tau_2'$.

It follows that every equation is equivalent to an equation of the form $\underline{n}\nu + \tau = \underline{m}\nu + \tau'$. Without loss of generality, we may assume that $n \geq m$, and hence $\mathsf{PrA} \vdash \underline{n} \geq \underline{m}$. Therefore, by LEMMA 12.0, we have

$$\mathsf{PrA} \vdash \underline{n}\nu + \tau = \underline{m}\nu + \tau' \;\leftrightarrow\; \underline{n - m}\nu + \tau = \tau',$$

which completes the proof. ⊣

We say that an atomic $\mathscr{L}_{\mathsf{PrA}^*}$-formula φ is in **ν-normal form**, if it is of the form

$$\underline{n}\nu + \tau = \tau', \quad \underline{n}\nu + \tau \equiv_m \tau', \quad \underline{n}\nu + \tau < \tau', \quad \tau' < \underline{n}\nu + \tau\,,$$

where $n, m \in \mathbb{N}$ and ν does not occur in τ, τ'. In that case, we call the number $n \in \mathbb{N}$ the **ν-coefficient** of φ.

LEMMA 12.4. *Let $\varphi_1, \ldots, \varphi_n$ be atomic $\mathscr{L}_{\mathsf{PrA}^*}$-formulae in ν-normal form. Then $\varphi_1 \wedge \ldots \wedge \varphi_n$ is logically equivalent to a conjunction of atomic $\mathscr{L}_{\mathsf{PrA}^*}$-formulae in ν-normal form, each of whose ν-coefficient is either $\mathbf{0}$ or $\mathbf{1}$.*

Proof. Without loss of generality, we may assume that each φ_i has a ν-coefficient $k_i \in \mathbb{N}$ with $k_i > \mathbf{0}$. Note that it is easy to check that $x\,R\,y \Leftrightarrow_{\mathsf{PrA}} \underline{k}x\,R\,\underline{k}y$, where R is either $=$, $<$, or $>$, and $k \in \mathbb{N}$ with $k \not\equiv \mathbf{0}$; a similar property also holds for congruences (see Exercise 12.1). In particular, by replacing k_i by $k \equiv \mathrm{lcm}^{\mathbb{N}}(k_1, \ldots, k_n)$ (see Exercise 9.3), we may assume that each formula φ_i has the same ν-coefficient k. Now if $\varphi_i \equiv \underline{k}\nu + \tau_i\,R_i\,\tau_i'$, then we can replace $\underline{k}\nu$ by w and obtain

$$\varphi_1 \wedge \ldots \wedge \varphi_m \;\Leftrightarrow_{\mathsf{PrA}}\; \psi_1 \wedge \ldots \wedge \psi_n \wedge w \equiv_k 0,$$

where φ_i is the formula $w + \tau_i \equiv_i \tau_i'$. ⊣

Lemma 12.5. *Let ν be a variable. If $\varphi_1, \ldots, \varphi_n$ are atomic $\mathscr{L}_{\mathsf{PrA}^*}$-formulae such that either φ_1 is an equation or each φ_i is a congruence, then there are atomic $\mathscr{L}_{\mathsf{PrA}^*}$-formulae $\psi_1, \ldots, \psi_n$ such that ψ_i is of the same type as φ_i, ν does not occur in $\psi_2, \ldots, \psi_n$, and $\varphi_1 \wedge \ldots \wedge \varphi_n \Leftrightarrow_{\mathsf{PrA}} \psi_1 \wedge \ldots \wedge \psi_n$.*

Proof. By induction, we may assume that $n = \mathbf{2}$. By Lemma 12.4, we may further suppose that each φ_i is in ν-normal form with ν-coefficient $\mathbf{1}$. There are two cases:

Case 1. $\varphi_1 \equiv \nu + \tau_1 = \tau_1'$ and $\varphi_2 \equiv \nu + \tau_2\,R\,\tau_2'$ for some terms $\tau_1, \tau_1', \tau_2, \tau_2'$ in which ν does not occur and R is either $=$, $<$, $>$ or $\equiv_m$ for some $m \in \mathbb{N}$. In this case, one can show that

$$\varphi_1 \wedge \varphi_2 \;\Leftrightarrow_{\mathsf{PrA}}\; \psi_1 \wedge \psi_2,$$

where $\psi_1 \equiv \varphi_1$ and $\psi_2 \equiv (\tau_1' + \tau_2)\,R\,(\tau_1 + \tau_2')$. Indeed, suppose that $\varphi_1 \wedge \varphi_2$ holds. Then we have

$$\tau_1' + \tau_2 = (\nu + \tau_1) + \tau_2 = \tau_1 + (\nu + \tau_2) \quad \text{and} \quad \big(\tau_1 + (\nu + \tau_2)\big)\,R\,\big(\tau_1 + \tau_2'\big).$$

Hence, we have $\psi_1 \wedge \psi_2$ as desired. The converse is similar.

Case 2. φ_1 is the formula $\nu + \tau_1 \equiv_{m_1} \tau_1'$ and φ_2 is $\nu + \tau_2 \equiv_{m_2} \tau_2'$. Then by Exercise 12.1 and by applying Lemma 12.4 to scale the ν-coefficients, we may suppose that $m_1 \equiv m_2$. The rest of the proof is the same as for the first case. ⊣

Theorem 12.6. *The theory* PrA *admits quantifier elimination with respect to the language $\mathscr{L}_{\mathsf{PrA}^*}$.*

Proof. We will check that PrA satisfies the assumptions of Theorem 12.2 with respect to the extended language $\mathscr{L}_{\mathsf{PrA}^*}$.

For the first condition, note that

$$\neg(\tau = \tau') \;\Leftrightarrow_{\mathsf{PrA}}\; \tau < \tau' \;\vee\; \tau' < \tau\,,$$

$$\neg(\tau < \tau') \;\Leftrightarrow_{\mathsf{PrA}}\; \tau = \tau' \;\vee\; \tau' < \tau\,,$$
$$\neg(\tau \equiv_m \tau') \;\Leftrightarrow_{\mathsf{PrA}}\; \bigvee_{k=\mathbf{1}}^{m-\mathbf{1}} (\tau + \underline{k} \equiv_m \tau') \quad \text{for every } m \in \mathbb{N},$$

where the last equivalence follows from Lemma 12.1.

We now turn to the second assumption. Let $\varphi_1, \ldots, \varphi_n$ be atomic $\mathscr{L}_{\mathsf{PrA}^*}$-formulae. We have to show that $\varphi \equiv \exists\nu(\varphi_1 \wedge \ldots \wedge \varphi_n)$ is logically equivalent to a quantifier-free $\mathscr{L}_{\mathsf{PrA}^*}$-formula ψ such that $\mathrm{free}(\psi) \equiv \mathrm{free}(\varphi) \setminus \{\nu\}$. Due to Tautology (T.2) and Lemma 12.5, we may suppose that each φ_i is in ν-normal form with ν-coefficient $\mathbf{1}$. We distinguish between the following four cases:

Case 1. There is an equation φ_i among $\varphi_1, \ldots, \varphi_n$. By finitely many applications of Lemma 12.5 in combination with Tautology (T.2), it is sufficient to check that $\exists\nu\varphi_i$ is logically equivalent to a quantifier-free $\mathscr{L}_{\mathsf{PrA}^*}$-formula, which is the case since $\exists\nu(\nu + \tau = \tau') \Leftrightarrow_{\mathsf{PrA}} \tau = \tau' \vee \tau < \tau'$.

Case 2. Each of the formulae $\varphi_1, \ldots, \varphi_n$ is a congruence. Then, by Tautology (T.2) and Lemma 12.5, it suffices to check that the quantifier in $\exists\nu(\nu + \tau \equiv_n \tau')$ can be eliminated. This is obviously possible, since by Lemma 12.1, there are $k, l \in \mathbb{N}$ such that $\tau \equiv_n \underline{k}$ and $\tau' \equiv \underline{l}$, and therefore, we can choose $\nu = \underline{l + n - k}$ in order to obtain a true formula.

Case 3. All formulae among $\varphi_1, \ldots, \varphi_n$ are inequalities. Since $<$ is a linear relation, we may order the lower and upper bounds in the following sense: For example, if $\nu + \tau_i < \tau_i'$ and $\nu + \tau_j < \tau_j'$ are two inequalities, then one can view them as upper bounds for ν, since if there is such a ν, then $\nu < \tau_i' - \tau_i$ and $\nu < \tau_j' - \tau_j$, where subtraction is defined as in Lemma 8.8. Now, by linearity of $<$, we have either $\tau_i' - \tau_i \leq \tau_j' - \tau_j$ or $\tau_j' - \tau_j < \tau_i' - \tau_i$. In other words, $\varphi_i \wedge \varphi_j$ is equivalent to

$$(\tau_i' + \tau_j \leq \tau_i + \tau_j' \wedge \nu + \tau_i < \tau_i') \vee (\tau_i + \tau_j' < \tau_i' + \tau_j \wedge \nu + \tau_j < \tau_j'),$$

where φ_i is the stronger bound in the first disjunct, and φ_j is the stronger bound in the second one. In a similar way, we can order the upper bounds. Using the distributive laws as well as Tautology (U.2), we may thus suppose that there is at most one lower and one upper bound. Moreover, note that

$$\exists\nu(\nu + \tau < \tau') \Leftrightarrow_{\mathsf{PrA}} \tau < \tau',$$
$$\exists\nu(\tau' < \nu + \tau) \Leftrightarrow_{\mathsf{PrA}} \mathsf{0} = \mathsf{0}\,.$$

On the other hand,

$$\exists\nu(\nu + \tau_1 < \tau_1' \wedge \tau_2' < \nu + \tau_2) \Leftrightarrow_{\mathsf{PrA}} \tau_1 + \tau_2' + \mathsf{1} < \tau_1' + \tau_2.$$

Therefore, the existential quantifier can be eliminated in both cases.

Case 4. There is at least one congruence and no equation among $\varphi_1, \ldots, \varphi_n$. As in the second case, without loss of generality we may assume that there is exactly one congruence and at most one inequality of each type. Without loss of generality, we only handle the case that there is exactly one lower and one upper bound, i.e., $\varphi \equiv \varphi_1 \wedge \varphi_2 \wedge \varphi_3$, where $\varphi_2 \equiv \tau_2' < \nu + \tau_2$ and $\varphi_3 \equiv \nu + \tau_3 < \tau_3'$. Then $\exists \nu \varphi$ is logically equivalent to the formula ψ with

$$\psi \equiv \bigvee_{k=\mathbf{1}}^{m} \Big(\tau_3 + \tau_2' + \underline{k} < \tau_3' + \tau_2 \ \wedge \ \tau_1 + \tau_2' + \underline{k} \equiv_m \tau_1' + \tau_2 \Big).$$

Note that if there is a ν such that φ holds, then it is of the form $\nu = \tau_2' - \tau_2 + x$ for some $x > \mathsf{0}$ such that $\nu < \tau_3' - \tau_3$, with the additional requirement that the congruence φ_1 be satisfied; one can then take x to be the smallest such solution, i.e., x is among $\underline{1}, \ldots, \underline{m}$. Clearly, $\nu \notin \mathrm{free}(\psi)$, and hence, ψ is as desired. ⊣

Lemma 12.7. *For every quantifier-free $\mathscr{L}_{\mathsf{PrA}^*}$-sentence φ, we have*

$$\textit{either} \quad \mathsf{PrA} \vdash \varphi \quad \textit{or} \quad \mathsf{PrA} \vdash \neg\varphi\,.$$

Proof. We will first check the claim for atomic $\mathscr{L}_{\mathsf{PrA}^*}$-sentences by proving that for every $\mathscr{L}_{\mathsf{PrA}^*}$-term τ which does not contain any variables, there is an $n \in \mathbb{N}$ such that $\mathsf{PrA} \vdash \tau = \underline{n}$. We proceed by induction on term construction:

- If $\tau \equiv \mathsf{0}$, then there is nothing to check.
- If $\tau \equiv \mathsf{s}\tau'$ and $\mathsf{PrA} \vdash \tau' = \underline{n}$, then $\mathsf{PrA} \vdash \tau = \mathsf{s}\tau' = \mathsf{s}\underline{n} = \underline{\mathsf{s}n}$ by N_0.
- If we have $\tau \equiv \tau_1 + \tau_2$, $\mathsf{PrA} \vdash \tau_1 = \underline{n_1}$, and $\mathsf{PrA} \vdash \tau_2 = \underline{n_2}$, then N_1 implies $\mathsf{PrA} \vdash \tau = \tau_1 + \tau_2 = \underline{n_1} + \underline{n_2} = \underline{n_1 + n_2}$.

Therefore, every atomic $\mathscr{L}_{\mathsf{PrA}^*}$-sentence is equivalent to a formula of the form $\underline{m} = \underline{n}, \underline{m} < \underline{n}$ or $\underline{m} \equiv_k \underline{n}$ for some $k \geq \mathbf{2}$, and is therefore $\mathbb{N}$-conform. Since $\mathbb{N}$-conformity is preserved under negation, conjunctions and disjunctions, the claim follows. ⊣

Corollary 12.8. *The theory* PrA *is complete.*

Proof. This follows immediately from Theorem 12.6 and Lemma 12.7. ⊣

Note that the proof of the completeness of PrA actually provides a *decision procedure* for $\mathscr{L}_{\mathsf{PrA}}$-sentences. In fact, there is an algorithm which computes, with respect to a given $\mathscr{L}_{\mathsf{PrA}}$-sentence φ, a quantifier-free $\mathscr{L}_{\mathsf{PrA}}$-sentence ψ such that $\varphi \Leftrightarrow_{\mathsf{PrA}} \psi$, which, by Lemma 12.7, can easily be decided in the sense that *either* $\mathsf{PrA} \vdash \psi$ *or* $\mathsf{PrA} \vdash \neg\psi$. In order to illustrate the algorithm, we provide an explicit example:

Example 12.9. We illustrate the quantifier elimination process using the example

$$\forall x \exists y \exists z (x < \underline{\mathbf{2}} z \wedge \underline{\mathbf{3}} z < x + y \wedge z \equiv_{\mathbf{2}} 0).$$

In a first step, we have to eliminate the quantifier $\exists z$. In order to achieve this, we first have to uniformise the z-coefficient using Lemma 12.4, obtaining the equivalent formula

$$\forall x \exists y \exists z \Big(\underline{\mathbf{3}} x < \underline{\mathbf{6}} z \wedge \underline{\mathbf{6}} z < \underline{\mathbf{2}} x + \underline{\mathbf{2}} y \wedge \underline{\mathbf{6}} z \equiv_{\mathbf{12}} 0 \Big).$$

Now, we can replace $\underline{\mathbf{6}} z$ by w, which yields

$$\forall x \exists y \exists w \Big(\underline{\mathbf{3}} x < w \wedge w < \underline{\mathbf{2}} x + \underline{\mathbf{2}} y \wedge w \equiv_{\mathbf{12}} 0 \wedge w \equiv_{\mathbf{12}} 0 \Big).$$

Clearly, the first congruence is a consequence of the second one and can thus be removed. Using the second case in the proof of Theorem 12.6, we can eliminate the variable w, and by further transformations we obtain

$$\begin{aligned} & \forall x \exists y \Big(\bigvee_{k=\mathbf{1}}^{\mathbf{12}} \big(\underline{\mathbf{3}} x + \underline{k} < \underline{\mathbf{2}} x + \underline{\mathbf{2}} y \wedge \underline{\mathbf{3}} x + \underline{k} \equiv_{\mathbf{6}} 0 \big) \Big) \\ \Leftrightarrow_{\mathsf{PrA}} \ & \forall x \Big(\bigvee_{k=\mathbf{1}}^{\mathbf{6}} \big(\underline{\mathbf{3}} x + \underline{k} \equiv_{\mathbf{6}} 0 \wedge \exists y (x + \underline{k} < \underline{\mathbf{2}} y) \big) \Big) \\ \Leftrightarrow_{\mathsf{PrA}} \ & \forall x \Big(\bigvee_{k=\mathbf{1}}^{\mathbf{6}} \underline{\mathbf{3}} x + \underline{k} \equiv_{\mathbf{6}} 0 \Big), \end{aligned}$$

which is provable by PrA due to Lemma 12.1. For the second equivalence, note that $\exists y (x + \underline{k} < \underline{\mathbf{2}} y)$ holds, which can be seen by taking $y = x + \underline{k}$. Following the algorithm, one would now have to replace the universal quantifier by a negated existential quantifier and negate the disjunction, and then restore the disjunctive normal form. Since this is very laborious, we will not pursue this further.

Let $\mathbb{N}^*$ be the $\mathscr{L}_{\mathsf{PrA}}$-structure $\big(\mathbb{N}, \mathbf{s}, +, \mathbf{0}\big)$, i.e., $\mathbb{N}^*$ is the same as $\mathbb{N}$, except that we do not have the binary function $\cdot^{\mathbb{N}}$ in $\mathbb{N}^*$. Since $\mathbb{N}$ is a model of PA, we find that $\mathbb{N}^*$ is a model of PrA. Now, since PrA is complete, we obtain that $\mathbf{Th}(\mathbb{N}^*)$ coincides with $\mathbf{Th}(\mathsf{PrA})$, i.e., for every $\mathscr{L}_{\mathsf{PrA}}$-sentence σ we have

$$\mathbb{N}^* \vDash \sigma \quad \Longleftrightarrow \quad \mathsf{PrA} \vdash \sigma \,.$$

The reader might now wonder why one does not take Presburger Arithmetic rather than Peano Arithmetic as the standard axiomatisation of arithmetic. Even though PrA is weaker than PA, it has the advantage that it is complete. However, the disadvantage of PrA is, that it is much too weak to serve as a proper axiomatisation of arithmetic. In fact, not even multiplication can be defined within PrA. This is a consequence of the following result, which states

that every set of natural numbers defined by some $\mathscr{L}_{\mathsf{PrA}}$-formula is *eventually periodic*.

LEMMA 12.10. *Let $\varphi(x)$ be an $\mathscr{L}_{\mathsf{PrA}}$-formula with one free variable x. Then*

$$\mathbb{N}^* \vDash \exists p > \mathsf{0}\, \exists n_0\, \forall n \geq n_0 \left(\varphi(n) \leftrightarrow \varphi(n+p)\right).$$

Proof. By THEOREM 12.6, it suffices to check the claim for quantifier-free $\mathscr{L}_{\mathsf{PrA}^*}$-formulae φ. We proceed by induction on the construction of the formula φ.

- If φ is an atomic formula, then, for some $m, n, p, k \in \mathbb{N}$, φ is logically equivalent to one of the following formulae:

$$\begin{aligned} \underline{m}v + \underline{n} &= \underline{l}\,, \\ \underline{m}v + \underline{n} &< \underline{l}\,, \\ \underline{m}v + \underline{n} &> \underline{l}\,, \\ \underline{m}v + \underline{n} &\equiv_k \underline{l}\,. \end{aligned}$$

  This follows immediately from LEMMATA 12.3 and 12.7. The first three cases are trivial, since one can choose $n_0 \equiv l$ and $p \equiv \mathbf{1}$ — in the first two cases we have $\neg\varphi(n_1)$ for all $n_1 \geq n_0$, and in the third case $\varphi(n)$ holds. Finally, suppose that φ is $\underline{m}v + \underline{n} \equiv_k \underline{l}$. Without loss of generality, we may assume that $n \equiv \mathbf{0}$. If l does not divide $\gcd(k, m)$, then φ is never satisfied and hence the claim is trivial. Otherwise, by EXERCISE 12.2, we may assume that $m \equiv \mathbf{1}$ and choose $p \equiv k$ and $n_0 \equiv \mathbf{0}$.
- If the claim holds for φ, then by TAUTOLOGY (H.0) it also holds for $\neg\varphi$ with the same witnesses $p, n_0 \in \mathbb{N}$ as for φ.
- Suppose that the claim holds for φ and ψ with witnesses n_0, p_0 and n_1, p_1, respectively, i.e.,

$$\mathbb{N}^* \vDash \forall n \geq n_0 \left(\varphi(n) \leftrightarrow \varphi(n+p_0)\right) \text{ and } \mathbb{N}^* \vDash \forall n \geq n_1 \left(\psi(n) \leftrightarrow \psi(n+p_1)\right).$$

  Now, let $n_2 :\equiv \max(n_0, n_1)$ and $p_2 :\equiv \operatorname{lcm}(p_0, p_1)$. Then, in $\mathbb{N}^*$ we have $\varphi(n) \leftrightarrow \varphi(n+p_2)$ and $\psi(n) \leftrightarrow \psi(n+p_2)$ for all $n \geq n_2$, and hence, the claim follows for $\varphi \vee \psi$ and $\varphi \wedge \psi$ by TAUTOLOGY (H.2) and (H.3), respectively.

⊣

THEOREM 12.11. *Multiplication is not definable in* PrA*, i.e., in* PrA *we cannot define a binary function* $\mathrm{mult}(\cdot, \cdot)$ *which satisfies the axioms* PA_4 *and* PA_5.

Proof. Assume toward a contradiction that there exists an $\mathscr{L}_{\mathsf{PrA}}$-formula $\varphi(x, y, z)$ such that $\mathsf{PrA} \vdash \forall x\, \forall y\, \exists! z\, \varphi(x, y, z)$ and

$$\mathrm{mult}(x, y) = z \;:\Longleftrightarrow\; \varphi(x, y, z)\,.$$

If there was such a formula φ, then $\psi(z) :\equiv \exists x \varphi(x, x, z)$ would be a formula defining z to be the square of some x. By Lemma 12.10, there are $p, n_0 \in \mathbb{N}$ with $p > 0$ such that

$$\mathbb{N}^* \vDash \forall n \geq n_0 \left(\psi(n) \leftrightarrow \psi(n+p)\right).$$

This, however, is impossible, since ψ is not eventually periodic: To see this, take, for example, $m := \max(n_0, p)$. Then $\mathbb{N}^* \vDash \psi(m^2)$, but $\mathbb{N}^* \nvDash \psi(m^2 + p)$, since $m \geq p$ implies

$$m^2 < (m^2 + p) < (m + \mathbf{1})^2,$$

and there are no squares between m^2 and $(m + \mathbf{1})^2$. ⊣

The previous theorem is the reason why Presburger Arithmetic is not considered as the standard axiomatisation of arithmetic. While multiplication cannot be expressed using the successor function and addition, exponentiation can be introduced using the successor function, addition and multiplication, as we have seen in Chapter 9. The main difference between PrA and PA lies in the fact that in PA allows *recursive definitions* using Gödel's β-function.

Non-standard models of PrA

In what follows, we recapitulate which axioms of PrA are actually necessary in order to prove its completeness. In fact, the proof only uses some particular instances of the Induction Schema; namely those which are concerned with equations and congruences. In order to axiomatise PrA, we surely need the axioms PA_0–PA_3. Moreover, in order to prove all required statements about equations, inequalities and congruences, we add the following axioms:

PrA_4: $\forall x \forall y (x + y = y + x)$

PrA_5: $\forall x \forall y \forall z (x + (y + z) = (x + y) + z)$

PrA_6: $\forall x \big(x = \mathsf{0} \vee \exists y (x = \mathsf{s}y)\big)$

PrA_7: $\forall x \forall y (x < y \vee x = y \vee y < x)$

PrA_8: $\forall x \big(\bigvee_{k=\mathbf{0}}^{n-\mathbf{1}} x \equiv_n \underline{k}\big)$ for every $n \in \mathbb{N}$

Note hat PrA_8 is actually an axiom scheme, just like the Induction Schema. However, it is considerably simpler than the Induction Schema in the sense that it is indexed by standard natural numbers instead of formulae. Each of these axioms is used in the proof of the completeness of PrA. For example, the commutative and associative laws of addition, i.e., PrA_4 and PrA_5, are used in the proof of Lemma 12.0. Moreover, by analysing all steps in the proof, it becomes clear that the axioms PA_0–PA_3 and PrA_4–PrA_8 suffice. By

completeness, we thus obtain that the theory given by these axioms coincides with PrA.

Now that we are in possession of a simplified system of axioms, we can describe non-standard models of PrA. Note that since PA extends PrA, every non-standard model of PA is also a non-standard model of PrA. However, there are also non-standard models of PrA which have a simpler structure than non-standard models of PA. The simplest non-standard model of PrA is surely the $\mathscr{L}_{\mathsf{PrA}}$-structure $\mathbf{M}$ with domain M consisting of all rational polynomials in the indeterminate X of degree at most $\mathbf{1}$, such that *either* the leading coefficient is positive and the constant coefficient is an integer, *or* the leading coefficient equals $\mathbf{0}$ and the constant coefficient is a natural number. More formally, M consists of all polynomials of the form

$$qX + a \ \in \ \mathbb{Q}[X]\,,$$

where $q \in \mathbb{Q}$, $q \geq \mathbf{0}$, and $a \in \mathbb{Z}$ or $a \in \mathbb{N}$, depending on whether $q > \mathbf{0}$ or $q = \mathbf{0}$. The interpretations of $\mathbf{0}$, $\mathbf{s}$, and $+$ are the obvious ones. Note that for the ordering $<$ we then obtain

$$\mathbf{M} \vDash pX + a < qX + b \quad \Longleftrightarrow \quad p <^{\mathbb{Q}} q \vee \big(p = q \wedge a <^{\mathbb{Z}} b\big),$$

which is essentially the lexicographic ordering. Since this is a linear order, it follows that $\mathbf{M} \vDash \mathsf{PrA}_7$. The other axioms, except for PrA_8, are obviously satisfied. In order to see that PrA_8 also holds in $\mathbf{M}$, note that

$$\mathbf{M} \vDash qX + a = \underline{n} \cdot \left(\frac{q}{n} X\right) + a \equiv_n a,$$

and hence, by taking $k \in \{\mathbf{0}, \ldots, n - \mathbf{1}\}$ to be the modulus of a, we obtain that $\mathbf{M} \vDash qX + a \equiv_n \underline{k}$.

Another model of PrA is obtained by admitting all rational polynomials

$$\sum_{k=\mathbf{0}}^{n} a_k = a_n X^n + \ldots + a_1 X + a_0 \ \in \ \mathbb{Q}[X]$$

with positive leading coefficient $a_n > \mathbf{0}$ and integer constant coefficient $a_0 \in \mathbb{Z}$, where $a_0 \in \mathbb{N}$ in the case when $n = \mathbf{0}$. Notice that both models are *not* models of PA: In PA, one can define exponentiation, and hence, there exists, e.g., the number $(\mathbf{1}X)^{\mathbf{1}X}$, which obviously does not have a polynomial representation.

Notes

The method of quantifier elimination was already used by Skolem in 1919 to prove the completeness of the first-order theory of a class of boolean algebrae. Presburger Arithmetic is named after the Polish mathematician Mojżesz Presburger, who proved its consistency,

completeness and decidability in 1929 (see [43]). However, his original proof is given for the integers rather than the natural numbers. The proof which is presented here follows the one presented in [8]. Skolem independently discovered Presburger's result in 1930 (see [51]) and further showed the same to be true for Skolem Arithmetic, i.e., the theory of natural numbers with multiplication but without addition.

Exercises

12.0 Show that there is no quantifier-free $\mathscr{L}_{\mathsf{PrA}}$-formula which is logically equivalent to the formula $\exists y(x = \mathbf{2}y)$, and conclude that Presburger Arithmetic does not admit quantifier elimination.

12.1 Prove that $\tau \equiv_n \tau' \Leftrightarrow_{\mathsf{PrA}} \underline{m}\tau \equiv_{mn} \underline{m}\tau'$ for all $\mathscr{L}_{\mathsf{PrA}}$-terms τ and τ' and every $m \in \mathbb{N}$.

12.2 Use Bézout's Lemma (see Exercise 8.2) in the standard model $\mathbb{N}$ to prove that in every formula of the form $\underline{m}v + \underline{n} \equiv_k \underline{l}$ for $m, n, k \in \mathbb{N}$ such that $\gcd(m, k) = \mathbf{1}$ is logically equivalent to a formula of the form $v \equiv_k \underline{l'}$ for some $l' \in \mathbb{N}$. Note that this essentially consists of proving that m has an inverse element modulo k.

12.3 Let $\mathscr{L}$ be the set of constant symbols $\{c_n : n \in \mathbb{N}\}$ and let T be the set of sentences $\{\neg(c_n = c_m) : n, m \in \mathbb{N}, n \not\equiv m\}$.

Show that T admits quantifier elimination.

12.4 Show that the theory $\mathbf{Th}(\mathbb{N}, <, \mathbf{s}, \mathbf{0})$ admits quantifier elimination and conclude that addition "+" is not definable in $\mathbf{Th}(\mathbb{N}, <, \mathbf{s}, \mathbf{0})$.

12.5 Let $\mathscr{L}$ consist of the binary relation symbol $\sim$ and the set of constant symbols $\{c_n : n \in \mathbb{N}\}$, and for every $n \in \mathbb{N}$ let M_n be an infinite subset of $\mathbb{N}$ such that $\mathbb{N}$ is the disjoint union of the M_n's. Let T be the theory of one equivalence relation with infinitely many infinite equivalence classes, i.e., consisting of the following axioms:

- $\forall x(x \sim x)$
- $\forall x, y(x \sim y \to y \sim x)$
- $\forall x, y, z(x \sim y \wedge y \sim z \to x \sim z)$
- $\forall x \exists x_1 \ldots \exists x_n \big(x \sim x_1 \wedge \ldots \wedge x \sim x_n \wedge \bigwedge_{i \neq j} \neg(x_i = x_j)\big)$ for each $n \in \mathbb{N}$
- $\forall x \exists x_1 \ldots \exists x_n (\bigwedge_{i=1}^{n} \neg(x \sim x_i) \wedge \bigwedge_{i \neq j} \neg(x_i \sim x_j))$ for each $n \in \mathbb{N}$

Show that T admits quantifier elimination.

12.6 Prove that the divisibility relation is not definable in PrA.

Part IV
The Axiom System ZFC

In this part, we first present the axioms of Set Theory, including the Axiom of Choice. Then we discuss the consistency of this axiomatic system and provide standard as well as non-standard models of Set Theory. In the last three chapters, we use Set Theory to prove the Löwenheim-Skolem Theorems, to construct models of PA, and to construct different models of the real numbers.

Chapter 13
The Axioms of Set Theory (ZFC)

In this chapter, we shall present and discuss the axioms of Zermelo-Fraenkel Set Theory including the Axiom of Choice, denoted ZFC. It will turn out that within this axiom system, we can develop all of first-order mathematics, and therefore, the axiom system ZFC serves as a foundation of mathematics. We will start with Zermelo's first axiomatisation of Set Theory and will show how basic mathematics can be developed within this system. Then we will introduce Zermelo's Axiom of Choice, Fraenkel's Axiom Schema of Replacement, and the Axiom of Foundation. Finally, we will discuss the notions of ordinal and cardinal numbers.

Before we begin presenting the axioms of Set Theory, let us say a few words about Set Theory in general: The signature of Set Theory $\mathscr{L}_{\mathsf{ST}}$ contains only one non-logical symbol, namely the binary **membership relation** denoted by $\in$, i.e., $\mathscr{L}_{\mathsf{ST}} = \{\in\}$. Furthermore, there exists just one type of objects, namely *sets*. However, to make life easier, instead of $\in(a, b)$ we write $a \in b$ (or also $b \ni a$ on rare occasions) and say that "a is an element of b", or that "a belongs to b". Furthermore, we write $a \notin b$ as an abbreviation of $\neg(a \in b)$. Later we will extend the signature of Set Theory $\mathscr{L}_{\mathsf{ST}}$ by defining some constants (like $\emptyset$ and ω), relations (like $\subseteq$), and operations (like the power set operation $\mathscr{P}$), but as we know from Chapter 6, all that can be expressed in Set Theory using defined constants, functions, and relations, can also be expressed by formulae containing the non-logical binary relation symbol $\in$ only.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_14

Zermelo's Axiom System (Z)

In 1905, Zermelo began to axiomatise Set Theory. In 1908, he published his first axiomatic system consisting of the following seven axioms:

1. Axiom der Bestimmtheit
   which corresponds to the Axiom of Extensionality
2. Axiom der Elementarmengen
   which includes the Axiom of Empty Set as well as the Axiom of Pairing
3. Axiom der Aussonderung
   which corresponds to the Axiom Schema of Separation
4. Axiom der Potenzmenge
   which corresponds to the Axiom of Power Set
5. Axiom der Vereinigung
   which corresponds to the Axiom of Union
6. Axiom der Auswahl
   which corresponds to the Axiom of Choice
7. Axiom des Unendlichen
   which corresponds to the Axiom of Infinity

The axioms 1–5 and axiom 7 (i.e., all axioms except the Axiom of Choice) form the so-called *Zermelo's axiom system*, denoted by Z, which will be discussed below.

Let us start with the axiom which states the existence of a set, namely the so-called *empty set*.

0. The Axiom of Empty Set

$$\exists x \forall z (z \notin x).$$

This axiom postulates the existence of a set without any elements, i.e., an empty set.

1. The Axiom of Extensionality

$$\forall x \forall y \big(\forall z (z \in x \leftrightarrow z \in y) \rightarrow x = y \big).$$

This axiom says that any sets x and y having the same elements are equal. Notice that the converse — which is: $x = y$ implies that x and y have the same elements — is just a consequence of the logical axiom L_{15}.

The Axiom of Extensionality also shows that the empty set, postulated by the Axiom of Empty Set, is unique: If x_0 and x_1 are empty sets, then we have

$\forall z(z \notin x_0 \wedge z \notin x_1)$, which implies $\forall z(z \in x_0 \leftrightarrow z \in x_1)$, and therefore, $x_0 = x_1$. Thus, with the Axiom of Empty Set and the Axiom of Extensionality we can prove $\exists! x\, \forall z(z \notin x)$, and therefore, we can denote the unique empty set by the constant symbol $\emptyset$.

Similarly, we define the binary relation symbol $\subseteq$, called **subset**, by stipulating

$$x \subseteq y :\Longleftrightarrow \forall z(z \in x \rightarrow z \in y).$$

Notice that for every x we have $\emptyset \subseteq x$. Furthermore, we define the binary relation symbol $\subsetneq$, called **proper subset**, by stipulating

$$x \subsetneq y :\Longleftrightarrow x \subseteq y \wedge x \neq y.$$

So far, we have at least one set, namely the empty set $\emptyset$, for which we have $\emptyset \subseteq \emptyset$.

2. The Axiom of Pairing

$$\forall x \forall y \exists u \forall z \big(z \in u \leftrightarrow (z = x \vee z = y)\big)$$

Notice that by the Axiom of Extensionality, the set u is uniquely defined by the sets x and y. Therefore, we can define the binary function symbol $\{\,\cdot\,,\,\cdot\,\}$ by stipulating

$$\{x, y\} = u :\Longleftrightarrow \forall z\big(z \in u \leftrightarrow (z = x \vee z = y)\big).$$

Notice that by the Axiom of Extensionality we have $\{x, x\} = \{x\}$, where $\{x\}$ denotes the set which contains the single element x. Thus, by the Axiom of Pairing, if x is a set, then also $\{x\}$ is a set. Now, starting with $\emptyset$, an iterated application of the Axiom of Pairing yields for example the sets $\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\{\{\emptyset\}\}\}, \ldots$, as well as $\{\emptyset, \{\emptyset\}\}, \{\{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ldots$.

Notice also that by the Axiom of Extensionality we have $\{x, y\} = \{y, x\}$. Therefore, it does not matter in which order the elements of a 2-element set are written down. However, with the Axiom of Pairing we can easily define **ordered pairs**, denoted by $\langle x, y\rangle$, as follows:

$$\langle x, y\rangle := \big\{\{x\}, \{x, y\}\big\}$$

It is not hard to show that $\langle x, y\rangle = \langle x', y'\rangle$ if and only if $x = x'$ *and* $y = y'$. Thus, we can define the binary function symbol $\langle\,\cdot\,,\,\cdot\,\rangle$ by stipulating

$$\langle x, y\rangle = u :\Longleftrightarrow \forall z\Big(z \in u \leftrightarrow \big(z = \{x\} \vee z = \{x, y\}\big)\Big).$$

Similarly, one could also define ordered triples, ordered quadruples, et cetera, but the notation becomes quite hard to read. However, once we have more axioms at hand, we can easily define arbitrarily large tuples.

3. The Axiom of Union

$$\forall x \exists u \forall z \big(z \in u \leftrightarrow \exists w \in x (z \in w)\big)$$

With this axiom we can define the unary function symbol $\bigcup$, called **union**, by stipulating

$$\bigcup x = u :\Longleftrightarrow \forall z \big(z \in u \leftrightarrow \exists w \in x (z \in w)\big).$$

Informally, for all sets x there exists the union of x which consists of all sets which belong to at least one element of x. For example, $x = \bigcup\{x\}$.

Similarly, we define the binary function symbol $\cup$ by stipulating

$$x \cup y = u :\Longleftrightarrow u = \bigcup\{x, y\}.$$

The set $x \cup y$ is called the **union** of x and y.

Now, with the Axiom of Union and the Axiom of Pairing, and by stipulating $x + 1 := x \cup \{x\}$, we can build, for example, the following sets:

$$\begin{aligned} 0 &:= \emptyset \\ 1 &:= 0 + 1 = 0 \cup \{0\} = \{0\} \\ 2 &:= 1 + 1 = 1 \cup \{1\} = \{0, 1\} \\ 3 &:= 2 + 1 = 2 \cup \{2\} = \{0, 1, 2\}, \end{aligned}$$

and so on. This construction leads to the following definition: A set x such that $\forall y (y \in x \to (y \cup \{y\}) \in x)$ is called **inductive**. More formally, we define the unary relation symbol ind by stipulating

$$\mathrm{ind}(x) :\Longleftrightarrow \forall y \Big(y \in x \to \big(y \cup \{y\}\big) \in x\Big).$$

Obviously, the empty set $\emptyset$ is inductive, i.e., $\mathrm{ind}(\emptyset)$; but of course, this definition only makes sense if also some other inductive sets exist. However, in order to make sure that non-empty inductive sets exist as well, we need the following axiom.

4. The Axiom of Infinity

$$\exists I \big(\emptyset \in I \wedge \mathrm{ind}(I)\big)$$

Informally, the Axiom of Infinity postulates the existence of a non-empty inductive set containing $\emptyset$. All the sets $0, 1, 2, \ldots$ constructed above — which we recognise as natural numbers — must belong to every inductive set. Thus, if there was a set which contains just the natural numbers, it would be the "smallest" inductive set containing the empty set. In order to construct this set, we need some more axioms.

5. The Axiom Schema of Separation

For each formula $\varphi(z,\vec{p})$ with $\mathrm{free}(\varphi) \subseteq \{z,\vec{p}\}$, the following formula is an axiom:

$$\forall x \forall \vec{p} \exists y \forall z \big(z \in y \leftrightarrow \big(z \in x \wedge \varphi(z,\vec{p})\big)\big),$$

where $\vec{p}$ is an abbreviation for $p_1,\ldots,p_n$, and correspondingly $\forall\vec{p}$ stands for $\forall p_1 \ldots \forall p_n$. One can think of the sets $p_1,\ldots,p_n$ as parameters of φ, which are usually some fixed sets. Informally, for each set x and every $\mathscr{L}_{\mathsf{ST}}$-formula $\varphi(z)$,

$$\big\{z \in x : \varphi(z)\big\}$$

is a set. Notice that the Axiom Schema of Separation just allows us to separate sets with a given property from a given set, but not to build collections of sets with a given property. For example, for a set x and $\varphi(z) \equiv z \notin z$, $\{z \in x : \varphi(z)\}$ is a set, but the collection $\{z : \varphi(z)\}$ is not a set.

As a first application of the Axiom Schema of Separation, we define the *intersection* of two sets x_0 and x_1: We use x_0 as a parameter and define $\varphi(z,x_0) \equiv z \in x_0$. Then, by the Axiom Schema of Separation, there exists a set $y = \{z \in x_1 : \varphi(z,x_0)\}$, i.e.,

$$z \in y \leftrightarrow (z \in x_1 \wedge z \in x_0)\,.$$

In other words, for any sets x_0 and x_1, the collection of all sets which belong to both x_0 and x_1 is a set. This set is called the **intersection** of x_0 and x_1 and is denoted by $x_0 \cap x_1$. More formally, we define the binary function symbol $\cap$ by stipulating

$$x_0 \cap x_1 = y :\Longleftrightarrow \forall z\big(z \in y \leftrightarrow z \in x_1 \wedge z \in x_0\big).$$

In general, for non-empty sets x we define the unary function symbol $\bigcap$ by stipulating

$$\bigcap x = y :\Longleftrightarrow y = \big\{u \in \bigcup x : \forall z \in x\,(u \in z)\big\},$$

which is the intersection of all sets which belong to x. In order to see that $\bigcap x$ is a set which is uniquely determined by x, let $\varphi(z,x) \equiv \forall y \in x\,(z \in y)$ and apply the Axiom Schema of Separation to $\bigcup x$. Notice that $x \cap y = \bigcap\{x,y\}$.

Another example is $\varphi(z,y) \equiv z \notin y$, where y is a parameter. In this case, $\{z \in x : z \notin y\}$ is a set, denoted by $x \setminus y$, which is called the **set-theoretic difference** of x and y. More formally, we define the binary function symbol $\setminus$ by stipulating

$$x \setminus y = u :\Longleftrightarrow \forall z\big(z \in u \leftrightarrow z \in x \wedge z \notin y\big).$$

The next axiom gives us for any set x the set of all subsets of x.

6. The Axiom of Power Set

$$\forall x \exists y \forall z (z \in y \leftrightarrow z \subseteq x)$$

Informally, the Axiom of Power Set states that for each set x there is a set $\mathscr{P}(x)$, called the **power set** of x, which consists of all subsets of x. More formally, we define the unary function symbol $\mathscr{P}$ by stipulating

$$\mathscr{P}(x) = y :\Longleftrightarrow \forall z (z \in y \leftrightarrow z \subseteq x).$$

The Set ω

As an application of the axioms which we have so far, we define the smallest non-empty inductive set containing $\emptyset$, denoted by ω, which will be the smallest set containing the natural numbers (see Chapter 16): By the Axiom of Infinity, there exists an non-empty inductive set I_0. Now, with the Axiom of Power Set and the Axiom Schema of Separation, we can define the set

$$\omega := \bigcap \big\{ X \in \mathscr{P}(I_0) : \emptyset \in X \wedge \mathrm{ind}(X) \big\}.$$

We have to show that the set ω is the smallest set which is inductive and contains $\emptyset$: By definition, ω is inductive and contains $\emptyset$. Now, let I be an inductive set with $\emptyset \in I$, and let $X_0 := \omega \cap I$. On the one hand, X_0 is inductive and $\emptyset \in X_0$. On the other hand, since $X_0 \subseteq \omega$, we have $X_0 \in \mathscr{P}(I_0)$, which implies that $\omega \subseteq X_0$. Therefore, ω is the unique inductive set containing $\emptyset$, which is contained in every inductive set containing $\emptyset$.

Later in Chapter 16, we shall see that ω is the domain of the standard model of Peano Arithmetic PA.

Functions, Relations, and Models

With the axioms which we have so far (i.e., with Zermelo's axiom system Z), we can define notions like functions and relations.

Cartesian Products and Functions

Let us first define Cartesian products: For arbitrary sets A and B we define the binary function symbol $\times$ by stipulating

$$A \times B := \big\{ \langle x, y \rangle : x \in A \wedge y \in B \big\},$$

where $\langle x, y\rangle = \{\{x\}, \{x, y\}\}$. The set $A \times B$ is called **Cartesian product** of A and B. Thus, the Cartesian product of two sets A and B is a subset of $\mathscr{P}(\mathscr{P}(A \cup B))$.

Now, we define **functions** $f : A \to B$ which map the elements of a set A to elements of a set B as certain subsets of $A \times B$. The set of all such functions is denoted by ${}^{A}B$, where we define

$$ {}^{A}B := \big\{ f \subseteq A \times B : \forall x \in A \, \exists! y \in B \big(\langle x, y\rangle \in f \big) \big\}. $$

For $f \in {}^{A}B$ (i.e., $f : A \to B$), we usually write $f(x) = y$ instead of $\langle x, y\rangle \in f$ and say that y is the **image** of x under f. Moreover, the set A is called the **domain** of f, denoted by $\operatorname{dom} f$. If $S \subseteq A$, then the **image** of S under f is denoted by $f[S] = \{f(x) : x \in S\}$, and $f|_S = \{\langle x, y\rangle \in f : x \in S\}$ is the restriction of f to S. Furthermore, for a function $f : A \to B$, $f[A]$ is called the **range** of f, denoted by $\operatorname{ran}(f)$.

Here are some special functions:

- A function $f : A \to B$ is **surjective**, or **onto**, if

  $$ \forall y \in B \, \exists x \in A \big(f(x) = y \big) . $$

  In order to emphasise the fact that the function f is surjective, one can write $f : A \twoheadrightarrow B$.

- A function $f : A \to B$ is **injective**, or **one-to-one**, if we have

  $$ \forall x_1 \in A \, \forall x_2 \in A \big(f(x_1) = f(x_2) \to x_1 = x_2 \big). $$

  In order to emphasise the fact that f is injective, one can write $f : A \hookrightarrow B$.

- A function $f : A \to B$ is **bijective** if it is injective and surjective. If $f : A \to B$ is bijective, then

  $$ \forall y \in B \, \exists! x \in A \big(\langle x, y\rangle \in f \big), $$

  which implies that

  $$ f^{-1} := \big\{ \langle y, x\rangle : \langle x, y\rangle \in f \big\} \in {}^{B}A $$

  is a function which is even bijective. Therefore, if a bijective function exists from A to B, then there is also one from B to A and we sometimes just say that there is a **bijection** between A and B. Notice that if $f : A \hookrightarrow B$ is injective, then f is a bijection between A and $f[A]$.

- If f is a function from A to B and g is a function from B to C, then the composition $g \circ f$ is a function from A to C, where

$$g \circ f := \Big\{ \langle x, z \rangle \in A \times C : \exists y \in B \big(\langle x, y \rangle \in f \wedge \langle y, z \rangle \in g \big) \Big\}.$$

- If f is a function with domain α, for some ordinal number α, then we call f a **sequence** of length α. If $f(\beta) =: x_\beta$ for $\beta < \alpha$, then we may write $f = \langle x_\beta : \beta < \alpha \rangle$.

Cartesian Products and Relations

Let us now turn back to Cartesian products: For a set A and a natural number $n \in \omega$, we define the n-fold Cartesian product, denoted A^n, by stipulating

$$\underbrace{A \times \ldots \times A}_{n\text{-times}} := \langle \cdots \langle \langle A \times A \rangle \times A \rangle \times \cdots \times A \rangle .$$

In order to simplify the notation, we identify the elements $\langle a_0, \ldots, a_{n-1} \rangle \in A^n$ with functions $f \in {}^nA$ by stipulating:

$$\begin{array}{ccc} A^n & \longleftrightarrow & {}^nA \\ \langle a_0, \ldots, a_{n-1} \rangle & \longmapsto & \big\{ \langle k, a_k \rangle : k \in n \big\} \\ \langle f(0), \ldots, f(n-1) \rangle & \longleftarrow\!\shortmid & f \end{array}$$

With this identification, we can define Cartesian products A^I for arbitrary sets I by identifying the elements of the Cartesian product with functions $f \in {}^IA$.

Notice that we are not yet able to define Cartesian products of arbitrary sets, but we are able to define relations as subsets of finite Cartesian products.

- For any set A and any $n \in \omega$, a set $R \subseteq A^n$ is called an ***n*-ary relation** on A.
- If $n = 2$, then $R \subseteq A \times A$ is also called a **binary relation**. For binary relations R we usually write xRy instead of $\langle x, y \rangle \in R$.
- A binary relation R on A is a **linear ordering** on A, if for any elements $x, y \in A$ we have xRy or $x = y$ or yRx, where these three cases are mutually exclusive.
- A linear ordering R on A is a **well-ordering** on A, if every non-empty subset $S \subseteq A$ has an R-minimal element, i.e., there exists a $x_0 \in S$ such that for each $y \in S$ we have $x_0 R y$. Notice that, since R is a linear ordering, the R-minimal element x_0 is unique. If there is a well-ordering R on A, then we say that A is *well-orderable*. The question whether each set is well-orderable has to be postponed until we have the Axiom of Choice.

Other important binary relations are the so-called equivalence relations: Let S be an arbitrary non-empty set. A binary relation $\sim$ on S is an **equivalence relation** if it is

- *reflexive* (i.e., for all $x \in S$: $x \sim x$),
- *symmetric* (i.e., for all $x, y \in S$: $x \sim y \leftrightarrow y \sim x$), and
- *transitive* (i.e., for all $x, y, z \in S$: $x \sim y \wedge y \sim z \ \rightarrow \ x \sim z$).

The **equivalence class** of an element $x \in S$, denoted by $[x]^{\sim}$, is the set $\{y \in S : x \sim y\}$. If it is clear from the context which relation $\sim$ is meant, we simply write $[x]$ for $[x]^{\sim}$. We would like to recall the fact that for any $x, y \in S$ we have *either* $[x]^{\sim} = [y]^{\sim}$ *or* $[x]^{\sim} \cap [y]^{\sim} = \emptyset$. A set $A \subseteq S$ is a set of **representatives** if A has exactly one element in common with each equivalence class $[x]^{\sim}$. We would like to mention that the existence of a set of representatives generally relies on the Axiom of Choice.

Zermelo-Fraenkel Set Theory with Choice (ZFC)

In 1922, Fraenkel and Skolem independently improved and extended Zermelo's original axiomatic system, and the final version was again presented by Zermelo in 1930. The two axioms which we have to add to Zermelo's system from 1908 are the Axiom Schema of Replacement and the Axiom of Foundation. In this section, we will present the remaining axioms of the so-called *Zermelo–Fraenkel Set Theory* with the Axiom of Choice, denoted by ZFC, which consists of Zermelo's axiom system Z together with the Axiom Schema of Replacement, the Axiom of Foundation, and the Axiom of Choice.

7. The Axiom Schema of Replacement

In the first form in which we present the Axiom Schema of Replacement, it states that for every first-order formula $\varphi(x, y, \bar{p})$ with $\text{free}(\varphi) = \{x, y, \bar{p}\}$, where $\bar{p}$ denotes a finite sequence of parameters, the following formula is an axiom:

$$\forall A \, \forall \bar{p} \big(\forall x \in A \, \exists! y \, \varphi(x, y, \bar{p}) \rightarrow \exists B \, \forall x \in A \, \exists y \in B \, \varphi(x, y, \bar{p}) \big)$$

In order to reformulate the Axiom Schema of Replacement, we introduce the notion of a *class function*: Let $\varphi(x, y, \bar{p})$ be a formula with $\text{free}(\varphi) = \{x, y, \bar{p}\}$ such that

$$\forall \bar{p} \forall x \, \exists! y \, \varphi(x, y, \bar{p}) \, .$$

Then, for each parameter set $\bar{p}$, the unary function symbol F, defined by stipulating

$$F(x) = y \ :\Longleftrightarrow \ \varphi(x, y, \bar{p}),$$

is called a **class function**. Now, the Axiom Schema of Replacement states that for every set A and for each class function F, $F[A] := \{F(x) : x \in A\}$

is contained in a set. With the Axiom Schema of Separation, we get slightly more.

FACT 13.0. *If F is a class function and A is a set, then $F[A]$ is a set.*

Proof. By the Axiom Schema of Replacement, there exists a set B such that $F[A] \subseteq B$, and by the Axiom Schema of Separation we obtain that $\{y \in B : \exists x \in A\,(F(x) = y)\}$ is a set which obviously corresponds to $F[A]$. ⊣

We can therefore replace the Axiom Schema of Replacement with the following, somewhat stronger statement:

$$\forall A\,\forall \bar{p}\Big(\forall x \in A\,\exists! y\,\varphi(x, y, \bar{p}) \to \exists B\,\forall y\,\big(y \in B \leftrightarrow \exists x \in A\,\varphi(x, y, \bar{p})\big)\Big)$$

In other words, images of sets under class functions are sets. In this context, we would like to mention that we can derive the Axiom Schema of Separation from this stronger form. To see this, let $\psi(z)$ be a first-order formula with $\text{free}(\psi) = \{z\}$, let x be a set, and define the class function F_ψ by stipulating

$$F_\psi(z) := \begin{cases} \{z\} & \text{if } \psi(z), \\ \emptyset & \text{otherwise.} \end{cases}$$

Then, by the stronger version of the Axiom Schema of Replacement, $F_\psi[x]$ is a set, and by the Axiom of Union, $\bigcup F_\psi[x] = \{z \in x : \psi(z)\}$ is also a set.

We would also like to mention that with this stronger version of the Axiom Schema of Replacement, the Axiom of Empty Set is also redundant (see EXERCISE 13.0).

With the Axiom Schema of Replacement, we can now define arbitrary Cartesian products: Let F be a class function and let I be an arbitrary set. Furthermore, for every $\iota \in I$ let $A_\iota := F(\iota)$ and let $A := \bigcup F[I]$. Then the set

$$\bigtimes_{\iota \in I} A_\iota := \Big\{ f \in {}^I A : \forall \iota \in I\big(f(\iota) \in A_\iota\big) \Big\}$$

is called the Cartesian product of the sets A_ι ($\iota \in I$). As a matter of fact, we would like to mention that with the axioms we have so far, we cannot prove that Cartesian products $\bigtimes_{\iota \in I} A_\iota$ of non-empty sets A_ι are non-empty.

8. The Axiom of Foundation

$$\forall x\,\big(x \neq \emptyset \to \exists y(y \in x \wedge (y \cap x = \emptyset))\big)$$

As a consequence of the Axiom of Foundation, we see that there is no infinite descending sequence $x_0 \ni x_1 \ni x_2 \ni \cdots$, since otherwise, the

set $\{x_0, x_1, x_2, \ldots\}$ would contradict the Axiom of Foundation. In particular, there is no set x such that $x \in x$ and there are also no cycles like $x_0 \in x_1 \in \cdots \in x_n \in x_0$. As a matter of fact, we would like to mention that if one assumes the Axiom of Choice, then the non-existence of such infinite descending sequences can be proved to be equivalent to the Axiom of Foundation.

The axiom system containing the axioms 0–8 is called **Zermelo–Fraenkel Set Theory** and is denoted by ZF.

9. The Axiom of Choice (AC)

$$\forall \mathscr{F}\, \exists f \left(f \text{ is a function from } \mathscr{F} \text{ to } \bigcup \mathscr{F} \wedge \left(\emptyset \notin \mathscr{F} \rightarrow \forall x \in \mathscr{F}\, (f(x) \in x)\right)\right),$$

or equivalently,

$$\forall \mathscr{F}\Big(\emptyset \notin \mathscr{F} \rightarrow \exists f\Big(f \in {}^{\mathscr{F}}\bigcup \mathscr{F} \wedge \forall x \in \mathscr{F}\big(f(x) \in x\big)\Big)\Big).$$

Informally, every family of non-empty sets has a choice function.

One can show that AC is equivalent to the statement that Cartesian products of non-empty sets are non-empty. More formally, let $\mathscr{F} = \{A_\iota : \iota \in I\}$ be a family of non-empty sets (i.e., for each $\iota \in I$, $A_\iota \neq \emptyset$). Then the Cartesian product

$$\bigtimes_{\iota \in I} A_\iota$$

is non-empty. In order to see this, let f be a choice function of $\mathscr{F}$. Then

$$\Big\{\big\langle \iota, f(A_\iota)\big\rangle : \iota \in I\Big\} \in \bigtimes_{\iota \in I} A_\iota\,,$$

and hence, $\bigtimes_{\iota \in I} A_\iota$ is non-empty.

ZF together with the Axiom of Choice AC is denoted by ZFC. Later on, we shall see that the axiom system ZFC is a foundation of first-order mathematics.

Well-Ordered Sets and Ordinal Numbers

In 1904, Zermelo [59] published his first proof of the so-called Well-Ordering Principle, which states that every set can be well-ordered (see THEOREM 13.3), and in 1908 he published a second proof (see [60]). In the proof presented below, we essentially follow Zermelo's first proof, but first we have to introduce the notion of ordinal numbers.

Ordinal Numbers

One of the most important concepts in Set Theory is the notion of *ordinal number*, which can be seen as a transfinite extension of the natural numbers. In order to define the concept of ordinal numbers, we must first give some definitions: Let $z \in x$. Then z is called an **$\in$-minimal element** of x, denoted by $\min_{\in}(z, x)$, if $\forall y(y \notin z \vee y \notin x)$, or equivalently, for any y in z we have $y \notin x$, or more formally,

$$\min_{\in}(z, x) \ :\Longleftrightarrow \ z \in x \ \wedge \ \forall y(y \in z \to y \notin x).$$

A set x is **ordered by** $\in$ if for any sets $y_1, y_2 \in x$ we have $y_1 \in y_2$ or $y_1 = y_2$ or $y_1 \ni y_2$, where the three cases do not have to be mutually exclusive. More formally,

$$\mathrm{ord}_{\in}(x) \ :\Longleftrightarrow \ \forall y_1, y_2 \in x \left(y_1 \in y_2 \vee y_1 = y_2 \vee y_1 \ni y_2\right).$$

Now, a set x is called **well-ordered by** $\in$ if it is ordered by $\in$ and if every non-empty subset of x has an $\in$-minimal element. More formally,

$$\mathrm{wo}_{\in}(x) \ :\Longleftrightarrow \ \mathrm{ord}_{\in}(x) \ \wedge \ \forall y \in \mathscr{P}(x) \left(y \neq \emptyset \to \exists z \in y \ \min_{\in}(z, y)\right).$$

Furthermore, a set x is called **transitive** if each element of x is a subset of x, i.e.,

$$\mathrm{trans}(x) \ :\Longleftrightarrow \ \forall y(y \in x \to y \subseteq x).$$

Notice that if x is transitive and $z \in y \in x$, then this implies $z \in x$. A set is called an **ordinal number**, or just an **ordinal**, if it is transitive and well-ordered by $\in$, i.e.,

$$\mathrm{ordinal}(x) :\Longleftrightarrow \ \mathrm{trans}(x) \ \wedge \ \mathrm{wo}_{\in}(x).$$

Ordinal numbers are usually denoted by Greek letters like α, β, γ, λ, et cetera, and the collection of all ordinal numbers is denoted by Ω. We will see later that Ω is not a set. However, we can consider "$\alpha \in \Omega$" as an abbreviation of $\mathrm{ordinal}(x)$, which is just a property of α, and thus, there is no harm in using the symbol Ω in this way, even though Ω is *not* an object of the set-theoretic universe.

Now, we are ready to prove the following result.

THEOREM 13.1.

(a) *If $\alpha \in \Omega$, then either $\alpha = \emptyset$ or $\emptyset \in \alpha$.*

(b) *If $\alpha \in \Omega$, then $\alpha \notin \alpha$.*

(c) *If $\alpha, \beta \in \Omega$, then $\alpha \in \beta$ or $\alpha = \beta$ or $\alpha \ni \beta$, where these three cases are mutually exclusive.*

(d) *If $\alpha \in \beta \in \Omega$, then $\alpha \in \Omega$.*

(e) *If $\alpha \in \Omega$, then also $\alpha \cup \{\alpha\} \in \Omega$.*

(f) *Ω is transitive and is well-ordered by $\in$. More precisely, Ω is transitive and ordered by $\in$, and every non-empty collection $C \subseteq \Omega$ has an $\in$-minimal element.*

Proof. (a) Since $\alpha \in \Omega$, α is well-ordered by $\in$. Thus, either $\alpha = \emptyset$, or, since $\alpha \subseteq \alpha$, α contains an $\in$-minimal element, say $x_0 \in \alpha$. If $x_0 \neq \emptyset$, then we find a $z \in x_0$, and by transitivity of α, we get $z \in \alpha$, which implies that x_0 is not an $\in$-minimal element of α. Hence, we must have $x_0 = \emptyset$, which shows that $\emptyset \in \alpha$.

(b) Assume towards a contradiction that $\alpha \in \alpha$. Then $\{\alpha\}$ is a non-empty subset of α and therefore contains an $\in$-minimal element. Now, since $\{\alpha\}$ just contains the element α, the $\in$-minimal element of $\{\alpha\}$ must be α, but on the other hand, $\alpha \in \alpha$ implies that α is not $\in$-minimal, a contradiction.

(c) First, notice that by (b), the three cases $\alpha \in \beta$, $\alpha = \beta$, $\alpha \ni \beta$ are mutually exclusive.

Let $\alpha, \beta \in \Omega$ be given. If $\alpha = \beta$, then we are done. So, let us assume that $\alpha \neq \beta$. Without loss of generality, we may assume that $\alpha \setminus \beta \neq \emptyset$.

We first show that $\alpha \cap \beta$ is the $\in$-minimal element of $\alpha \setminus \beta$: Let γ be an $\in$-minimal element of $\alpha \setminus \beta$. Since α is transitive and $\gamma \in \alpha$, $\forall u(u \in \gamma \to u \in \alpha)$, and since γ is an $\in$-minimal element of $\alpha \setminus \beta$, $\forall u(u \in \gamma \to u \in \beta)$, which implies $\gamma \subseteq \alpha \cap \beta$. On the other hand, if there is a $w \in (\alpha \cap \beta) \setminus \gamma$, then, since α is ordered by $\in$ and $\gamma \neq w$ ($\gamma \notin \beta \ni w$), we must have $\gamma \in w$, and since β is transitive and $w \in \beta$, this implies that $\gamma \in \beta$, which contradicts the fact that $\gamma \in (\alpha \setminus \beta)$. Hence, $\gamma = \alpha \cap \beta$ is the $\in$-minimal element of $\alpha \setminus \beta$. Now, if also $\beta \setminus \alpha \neq \emptyset$, then we would find that $\alpha \cap \beta$ is also the $\in$-minimal element of $\beta \setminus \alpha$, which is obviously a contradiction.

Thus, $\alpha \setminus \beta \neq \emptyset$ implies that $\beta \setminus \alpha = \emptyset$, or in other words, $\beta \subseteq \alpha$, which is the same as saying $\beta = \alpha \cap \beta$. Consequently, we see that β is the $\in$-minimal element of $\alpha \setminus \beta$, and, in particular, we have $\beta \in \alpha$.

(d) Let $\alpha \in \beta \in \Omega$. Since β is transitive, α is ordered by $\in$. So, it remains to show that α is transitive and well-ordered by $\in$.

well-ordered by $\in$: Because β is transitive, every subset of α is also a subset of β and consequently contains an $\in$-minimal element.

transitive: Let $\delta \in \gamma \in \alpha$. We have to show that $\delta \in \alpha$. Since β is transitive, $\delta \in \beta$, and since β is ordered by $\in$, we have either $\delta \in \alpha$ or $\delta = \alpha$ or $\alpha \in \delta$. If $\delta \in \alpha$, we are done, and if $\delta = \alpha$ or $\alpha \in \delta$, then the set $\{\alpha, \gamma, \delta\} \subseteq \beta$ does not have an $\in$-minimal element, which contradicts the fact that β is well-ordered by $\in$.

(e) We have to show that $\alpha \cup \{\alpha\}$ is transitive and well-ordered by $\in$.

transitive: If $\beta \in (\alpha \cup \{\alpha\})$, then either $\beta \in \alpha$ or $\beta = \alpha$, and in both cases we have $\beta \subseteq (\alpha \cup \{\alpha\})$.

well-ordered by $\in$: Since α is an ordinal, $\alpha \cup \{\alpha\}$ is ordered by $\in$. Let now $x \subseteq (\alpha \cup \{\alpha\})$ be a non-empty set. If $x = \{\alpha\}$, then α is obviously an

$\in$-minimal element of x. Otherwise, $x \cap \alpha \neq \emptyset$, and since $\alpha \in \Omega$, $x \cap \alpha$ has an $\in$-minimal element, say γ. Since α is transitive, we have $x \cap \gamma = \emptyset$ (otherwise, γ would not be $\in$-minimal in $x \cap \alpha$), which implies that γ is $\in$-minimal in x.

(f) Ω *is transitive and ordered by* $\in$: This is a consequence of part (d) and part (c), respectively.

Ω *is well-ordered by* $\in$: Let $C \subseteq \Omega$ be a non-empty collection of ordinals. If $C = \{\alpha\}$ for some $\alpha \in \Omega$, then α is the $\in$-minimal element of C. Otherwise, C contains an ordinal δ_0 such that $\delta_0 \cap C \neq \emptyset$ and let $x := \delta_0 \cap C$. Then x is a non-empty *set* of ordinals. Now, let $\alpha \in x$ and let γ be an $\in$-minimal element of $x \cap (\alpha \cup \{\alpha\})$. By definition, $\gamma \in (\alpha \cup \{\alpha\})$, and since $(\alpha \cup \{\alpha\}) \in \Omega$, $\gamma \subseteq (\alpha \cup \{\alpha\})$. Thus, every ordinal $\gamma' \in \gamma$ belongs to $\alpha \cup \{\alpha\}$, but by the definition of γ, γ' cannot belong to $x \cap (\alpha \cup \{\alpha\})$, which implies that γ is also $\in$-minimal in x, and consequently in C. ⊣

Notice that if Ω is a set, then by (f), Ω is an ordinal number, and therefore $\Omega \in \Omega$, which contradicts (b). Thus, the collection of all ordinals Ω is not a set, but a so-called *class*. In general, a collection of sets, satisfying for example a certain formula, which is not necessarily a set is called a **class**. For example, Ω is a class which is *not* a set (it consists of all transitive sets which are well-ordered by $\in$). Even though proper classes (i.e., classes which are not sets) do not belong to the set-theoretic universe, it is sometimes convenient to handle them like sets, e.g., taking intersections or extracting certain subsets or subclasses from them.

Since $\in$ constitutes a linear ordering by (c), we use the following notation:

$$\begin{aligned} \alpha < \beta &:\Longleftrightarrow \alpha \in \beta \\ \alpha \leq \beta &:\Longleftrightarrow \alpha < \beta \vee \alpha = \beta \end{aligned}$$

By Theorem 13.1.(e) we know that if $\alpha \in \Omega$, then also $(\alpha \cup \{\alpha\}) \in \Omega$. Now, for ordinals $\alpha \in \Omega$, let $\alpha + 1 := \alpha \cup \{\alpha\}$. Part (b) of the following result — which is just a consequence of Theorem 13.1 — motivates this notation.

Corollary 13.2.

(a) *If* $A \subseteq \Omega$ *is a set of ordinals, then* $\bigcup A$ *is an ordinal.*

(b) *If* $\alpha, \beta \in \Omega$ *and* $\alpha \in \beta$*, then* $\alpha + 1 \subseteq \beta$*. In other words,* $\alpha + 1$ *is the least ordinal which contains* α*.*

(c) *For every ordinal* $\alpha \in \Omega$ *we have either* $\alpha = \bigcup \alpha$ *or there exists a* $\beta \in \Omega$ *such that* $\alpha = \beta + 1$*.*

Proof. (a) For every $\beta \in \bigcup A$ there is an ordinal $\gamma \in A$ such that $\beta \in \gamma$. Thus, by Theorem 13.1.(d), β is an ordinal which implies that $\bigcup A \subseteq \Omega$ is a set of ordinals, and by Theorem 13.1.(f) we get that $\bigcup A$ is well-ordered by $\in$. Furthermore, if $\alpha \in \beta \in \gamma \in \bigcup A$, then, since γ is transitive, we have $\alpha \in \gamma$, which implies $\alpha \in \bigcup A$, i.e., $\bigcup A$ is transitive.

(b) Assume $\alpha \in \beta$, then $\{\alpha\} \subseteq \beta$, and since β is transitive, we also have $\alpha \subseteq \beta$; thus, $\alpha + 1 = \alpha \cup \{\alpha\} \subseteq \beta$.

(c) Since α is transitive, $\bigcup\alpha \subseteq \alpha$. Thus, if $\alpha \neq \bigcup\alpha$, then $\alpha \setminus \bigcup\alpha \neq \emptyset$. Let β be $\in$-minimal in $\alpha \setminus \bigcup\alpha$. Then $\beta \in \alpha$ and $\beta + 1 \in \Omega$, and by part (b) we have $\beta + 1 \subseteq \alpha$. On the one hand, $\alpha \in \beta + 1$ would imply that $\alpha \in \alpha$, a contradiction to THEOREM 13.1.(b). On the other hand, $\beta + 1 \in \alpha$ would imply that $\beta \in \bigcup\alpha$, which contradicts the choice of β. Thus, we must have $\beta + 1 = \alpha$. ⊣

This leads to the following definitions: An ordinal α is called a **successor ordinal** if there exists an ordinal β such that $\alpha = \beta + 1$; otherwise, it is called a **limit ordinal**. In particular, $\emptyset$ is a limit ordinal. Notice that $\alpha \in \Omega$ is a limit ordinal if and only if $\bigcup\alpha = \alpha$.

With these definitions one can show that ω, defined above as the least non-empty inductive set, is in fact the least non-empty limit ordinal. In particular, we have $\bigcup\omega = \omega$.

Now we are ready to prove the following

THEOREM 13.3. *The Well-Ordering Principle is equivalent to the Axiom of Choice.*

Proof. ($\Rightarrow$) Let $\mathscr{F}$ be any family of non-empty sets and let $<$ be any well-ordering on $\bigcup\mathscr{F}$. Define $f : \mathscr{F} \to \bigcup\mathscr{F}$ by stipulating $f(x)$ to be the $<$-minimal element of x.

($\Leftarrow$) Let M be a set. If $M = \emptyset$, then M is well-ordered and we are done. Therefore, assume that $M \neq \emptyset$ and let $\mathscr{P}^*(M) := \mathscr{P}(M) \setminus \{\emptyset\}$. Furthermore, let

$$f : \mathscr{P}^*(M) \to M$$

be an arbitrary but fixed choice function for the family $\mathscr{P}^*(M)$, which exists by the Axiom of Choice.

Now, an injective function

$$w_\alpha : \alpha \hookrightarrow M$$

from some ordinal $\alpha \in \Omega$ into M is called an ***f*-set** if for all $\gamma \in \alpha$ we have

$$w_\alpha(\gamma) = f\big(M \setminus \{w_\alpha(\delta) : \delta \in \gamma\}\big).$$

For example, $w_1 = \{\langle 0, f(M)\rangle\}$ is an f-set – in fact, w_1 is the unique f-set with domain $\{0\}$. In general, for every $\alpha \in \Omega$ there is at most one f-set w_α with domain α. In order to see this, assume that w_α and w'_α are two distinct f-sets with domain α. Because w_α and w'_α are distinct and $\alpha \in \Omega$, there exists an $\in$-minimal $\gamma \in \alpha$ such that $w_\alpha(\gamma) \neq w'_\alpha(\gamma)$, but since for all $\delta \in \gamma$ we have $w_\alpha(\delta) = w'_\alpha(\delta)$, which contradicts the fact that

$$w_\alpha(\gamma) = f\big(M \setminus \{w_\alpha(\delta) : \delta \in \gamma\}\big) = f\big(M \setminus \{w'_\alpha(\delta) : \delta \in \gamma\}\big) = w'_\alpha(\gamma).$$

Thus, if there exists an f-set w_α for some $\alpha \in \Omega$, then this f-set w_α is the unique f-set with $\mathrm{dom}(w_\alpha) = \alpha$. Moreover, if w_β and w_α are f-sets and $\beta \in \alpha$, then $w_\alpha|_\beta = w_\beta$ (i.e., the restriction of w_α to β is equal to w_β).

Because every f-set w_α induces a well-ordering on $\mathrm{ran}(w_\alpha) \subseteq M$, by the Axiom Schema of Separation, the collection of all f-sets is a set, say S. Now, we define the class function F by stipulating

$$F(w) := \begin{cases} \alpha & \text{if } w \text{ is an } f\text{-set with } \mathrm{dom}(w) = \alpha, \\ \emptyset & \text{otherwise.} \end{cases}$$

Notice that by the Axiom Schema of Replacement, $F[S]$ is a set of ordinals, and therefore, by Corollary 13.2.(a), $\bigcup F[S]$ is an ordianl, say α_0.

On S we define the ordering $\prec$ as follows: For two distinct f-sets w_α and w_β, let

$$w_\alpha \prec w_\beta \iff \alpha \in \beta.$$

Since Ω is well-ordered by $\in$, S is well-ordered by $\prec$ and for $w_{\alpha_0} := \bigcup S$ we have $\mathrm{dom}(w_{\alpha_0}) = \alpha_0$. Let now

$$M' := \big\{x \in M : \exists \gamma \in \alpha_0 \big(w_{\alpha_0}(\gamma) = x\big)\big\}.$$

Then $M' = M$ and $w_{\alpha_0} \in S$, since otherwise, w_{α_0} can be extended to the f-set

$$w_{\alpha_0} \cup \big\{\big\langle \alpha_0, f(M \setminus M')\big\rangle\big\},$$

which is a contradiction to the definition of S. Therefore, the injective function $w_{\alpha_0} : \alpha_0 \hookrightarrow M$ is surjective. In other words, w_{α_0} is a bijection between α_0 and M. Finally, define the binary relation $<$ on M by stipulating

$$x < y :\iff w_{\alpha_0}^{-1}(x) \in w_{\alpha_0}^{-1}(y)\,.$$

Then, since α_0 is well-ordered by $\in$, M is well-ordered by $<$. ⊣

As an immediate consequence of Theorem 13.3 we get the following

Corollary 13.4. *For each well-ordering $<$ of a set A there exists a unique ordinal α and a unique bijective function $w_\alpha : \alpha \to A$ such that for all $\beta, \gamma \in \alpha$,*

$$\beta \in \gamma \iff w_\alpha(\beta) < w_\alpha(\gamma)\,.$$

Proof. Let $<$ be a well-ordering of a set A. Define $f : \mathscr{P}^*(A) := \mathscr{P}(A) \setminus \{\emptyset\}$ by stipulating $f(x)$ to be the $<$-minimal element of x. Then the second part of the proof of Theorem 13.3 gives us a bijection $w_\alpha : \alpha \to A$ with desired property. ⊣

We conclude this chapter with the development of ordinal and cardinal arithmetic. Since this topic is not essential outside of set theory, some results are given without proof (for proofs we refer the reader to Halbeisen [22]).

Ordinal Arithmetic

The next result is the Transfinite Recursion Theorem, a very powerful tool which is used, for example, to define ordinal arithmetic (see below) or to build the cumulative hierarchy of sets (see Chapter 14).

Theorem 13.5 (Transfinite Recursion Theorem). *Let F be a class function which is defined for all sets. Then there is a unique class function G defined on Ω such that for each $\alpha \in \Omega$ we have*

$$G(\alpha) = F(G|_\alpha), \quad \text{where } G|_\alpha = \big\{\langle \beta, G(\beta)\rangle : \beta \in \alpha\big\}.$$

By transfinite recursion we are able to define addition, multiplication, and exponentiation of arbitrary ordinal numbers (see Exercise 13.1):

***Ordinal Addition*:** For arbitrary ordinals $\alpha \in \Omega$, we define

(a) $\alpha + 0 := \alpha$,
(b) $\alpha + 1 := \alpha \cup \{\alpha\}$,
(c) $\alpha + (\beta + 1) := (\alpha + \beta) + 1$ for all $\beta \in \Omega$,
(d) and if $\beta \in \Omega$ is non-empty and a limit ordinal, then $\alpha + \beta := \bigcup_{\delta\in\beta}(\alpha+\delta)$.

Notice that, for example, $1 + \omega = \omega \neq \omega + 1$, which shows that addition of ordinals is in general not commutative.

***Ordinal Multiplication*:** For arbitrary ordinals $\alpha \in \Omega$, we define

(a) $\alpha \cdot 0 := 0$,
(b) $\alpha \cdot (\beta + 1) := (\alpha \cdot \beta) + \alpha$ for all $\beta \in \Omega$,
(c) and if $\beta \in \Omega$ is a limit ordinal, then $\alpha \cdot \beta := \bigcup_{\delta\in\beta}(\alpha \cdot \delta)$.

Notice that, for example, $2 \cdot \omega = \omega \neq \omega + \omega = \omega \cdot 2$, which shows that multiplication of ordinals is in general not commutative.

***Ordinal Exponentiation*:** For arbitrary ordinals $\alpha \in \Omega$, we define

(a) $\alpha^0 := 1$,
(b) $\alpha^{\beta+1} := \alpha^\beta \cdot \alpha$ for all $\beta \in \Omega$,
(c) and if $\beta \in \Omega$ is non-empty and a limit ordinal, then $\alpha^\beta := \bigcup_{\delta\in\beta}(\alpha^\delta)$.

By definition, we obtain that addition, multiplication, and exponentiation of ordinals are binary operations on Ω. Moreover, one can prove that addition and multiplication of ordinals are also associative and that the left distributive law holds (but not the right distributive law). In order to prove a property for all ordinals, the following generalisation of the induction principle for natural numbers is a very powerful tool.

THEOREM 13.6 (TRANSFINITE INDUCTION PRINCIPLE). *Suppose that $\varphi(x)$ is an $\mathscr{L}_{\mathsf{ST}}$-formula and suppose that the following conditions hold:*

(a) $\varphi(0)$

(b) $\forall\alpha \in \Omega\big(\varphi(\alpha) \to \varphi(\alpha+1)\big)$

(c) $\big(\forall\beta < \alpha\ \varphi(\beta)\big) \to \varphi(\alpha)$ *if $\alpha \in \Omega$ is a limit ordinal.*

Then $\varphi(\alpha)$ holds for all ordinals $\alpha \in \Omega$.

EXAMPLE 13.7. We prove the left distributive law of ordinal arithmetic, where we assume that the associativity of addition has already been shown. Let $\alpha, \beta \in \Omega$ be fixed ordinals.

(a) By definition, we have $\alpha \cdot (\beta + 0) = \alpha \cdot \beta = (\alpha \cdot \beta) + (\alpha \cdot 0)$.

(b) Assume that $\alpha \cdot (\beta + \gamma)$ holds. Then we obtain

$$\begin{aligned}\alpha \cdot \big(\beta + (\gamma+1)\big) &= \alpha \cdot \big((\beta+\gamma)+1\big)\\ &= \alpha \cdot (\beta+\gamma) + \alpha\\ &= \big((\alpha\cdot\beta) + (\alpha\cdot\gamma)\big) + \alpha\\ &= (\alpha\cdot\beta) + \big((\alpha\cdot\gamma)+\alpha\big)\\ &= (\alpha\cdot\beta) + \big(\alpha\cdot(\gamma+1)\big).\end{aligned}$$

(c) Suppose that γ is a limit ordinal and that $\alpha \cdot (\beta + \delta) = (\alpha \cdot \beta) + (\alpha \cdot \delta)$ for all $\delta < \gamma$. Then we have

$$\begin{aligned}\alpha \cdot (\beta+\gamma) = \alpha \cdot \bigcup_{\delta<\gamma}(\beta+\delta) = \bigcup_{\delta<\gamma}\alpha\cdot(\beta+\delta) = \bigcup_{\delta<\gamma}\big((\alpha\cdot\beta)+(\alpha\cdot\delta)\big)\\ = (\alpha\cdot\beta) + \bigcup_{\delta<\gamma}\alpha\cdot\delta = (\alpha\cdot\beta)+(\alpha\cdot\gamma).\end{aligned}$$

Hence, by the TRANSFINITE INDUCTION PRINCIPLE, the left distributive law holds for all ordinals.

Let us consider the set ω again. The ordinals belonging to ω are called **natural numbers**. Since ω is the smallest non-empty limit ordinal, all natural numbers, except 0, are successor ordinals. Thus, for each $n \in \omega$ we have either $n = 0$ or there is an $m \in \omega$ such that $n = m + 1$. Since by definition,

$$k < n \iff k \in n$$

for each $n \in \omega$ we have $n = \{k \in \omega : k < n\}$, i.e., $n = \{0, 1, \ldots, n-1\}$. In particular, for every $n \in \omega$, n is a set containing exactly n elements.

With ordinal addition, multiplication, and exponentiation we can define sums, products, and powers of natural numbers within ZF. In fact, we can define these operations in Z already (see Exercise 13.5).

Cardinal Numbers and Cardinal Arithmetic

By Corollary 13.4 we know that for each well-ordering $<$ of a set A there exists a unique ordinal α and a unique bijective function $w_\alpha : \alpha \to A$ such that for all $\beta, \gamma \in \alpha$, we have $\beta \in \gamma \iff w_\alpha(\beta) < w_\alpha(\gamma)$. The unique ordinal α which corresponds to a well-ordering $<$ of A is called the **order type** of the well-ordering $<$.

In the presence of AC, we are now able to define cardinal numbers as ordinals: For any set A we define the cardinality of A, denoted by $|A|$, by stipulating

$$|A| := \min\left\{\alpha \in \Omega : \alpha \text{ is the order type of a well-ordering of } A\right\}.$$

By definition we have

$$|A| = \min\left\{\alpha \in \Omega : \text{there is a bijection between } \alpha \text{ and } A\right\}.$$

In order to see that this definition makes sense, notice that by AC, every set A is well-orderable, and that by the above remark, every well-ordering on A corresponds to exactly one ordinal. Therefore, for each set A, the set of all order types of well-orderings of A is a non-empty set of ordinals. Let $C \subseteq \Omega$ be this set of ordinals. Then, by Theorem 13.1.(f), C has an $\in$-minimal element $\min C$, which shows that $|A|$ is indeed an ordinal.

For example, we have $|n| = n$ for every $n \in \omega$, and $|\omega| = \omega$; but in general, for $\alpha \in \Omega$, we do not have $|\alpha| = \alpha$. For example, $|\omega + 1| \neq \omega + 1$, since $|\omega + 1| = \omega$ and $\omega \neq \omega + 1$. However, there are also other ordinals α besides $n \in \omega$ and ω itself for which we have $|\alpha| = \alpha$. This leads to the following definition:

An ordinal number $\kappa \in \Omega$ such that $|\kappa| = \kappa$ is called a **cardinal number**, or just a **cardinal**. Cardinal numbers are usually denoted by Greek letters like κ, λ, μ, et cetera, or by $\aleph$'s. For example, the cardinal number ω is denoted by $\aleph_0$, which is the cardinality of countably infinite sets.

A cardinal κ is **infinite** if $\kappa \notin \omega$, otherwise, it is **finite**. In other words, a cardinal is finite if and only if it is a natural number.

Since cardinal numbers are just a special kind of ordinal, they are well-ordered by $\in$. However, for cardinal numbers κ and λ we usually write $\kappa < \lambda$ instead of $\kappa \in \lambda$, i.e.,

$$\kappa < \lambda \quad \iff \quad \kappa \in \lambda.$$

Notice that for any cardinals κ and λ, if $\kappa \leq \lambda$ and $\lambda \leq \kappa$ then $\kappa = \lambda$. Thus, for any sets A and B, in order to show that $|A| = |B|$, it is enough to find injections from $A \hookrightarrow B$ and $B \hookrightarrow A$, since this implies $|A| \leq |B|$ and $|B| \leq |A|$, and hence, $|A| = |B|$.

The next result implies that there are arbitrarily large cardinal numbers.

THEOREM 13.8 (CANTOR'S THEOREM). *For every set A, $|A| < |\mathscr{P}(A)|$.*

Proof. Let A be an arbitrary set. Obviously, we have $|A| \leq |\mathscr{P}(A)|$. If we had $|A| = |\mathscr{P}(A)|$, then there would be a bijection between A and $\mathscr{P}(A)$. In particular, there would be a surjection $A \twoheadrightarrow \mathscr{P}(A)$. Therefore, in order to prove $|A| < |\mathscr{P}(A)|$, it is enough to show that there is no surjection $f : A \twoheadrightarrow \mathscr{P}(A)$.

If $A = \emptyset$, then $\mathscr{P}(A) = \{\emptyset\}$ and $f = \emptyset$; hence, f is not a surjection.

If $A \neq \emptyset$, consider the set

$$\Gamma := \big\{x \in A : x \notin f(x)\big\}.$$

On the one hand, since $\Gamma \subseteq A$, $\Gamma \in \mathscr{P}(A)$. On the other hand, for each $x \in A$ we have

$$x \in \Gamma \iff x \notin f(x),$$

and therefore, there is no $x \in A$ such that $f(x) = \Gamma$, which shows that f is not surjective. ⊣

A set A is called **countable** if $|A| \leq \omega$, and it is called **uncountable** if $|A| > \omega$. For example, one can show that the set $\mathbb{Q}$ of rational numbers is countable and that the set $\mathbb{R}$ of real numbers is uncountable (see EXERCISE 13.8).

Let κ be a cardinal. The smallest cardinal number which is greater than κ is denoted by κ^+, i.e.,

$$\kappa^+ = \min\big\{\alpha \in \Omega : \kappa < |\alpha|\big\}.$$

Notice that by THEOREM 13.8, $\kappa < 2^\kappa$ for every cardinal κ. In particular, for every cardinal κ, $\{\alpha \in \Omega : \kappa < |\alpha|\}$ is non-empty and therefore κ^+ exists.

A cardinal μ is called a **successor cardinal** if there exists a cardinal κ such that $\mu = \kappa^+$; otherwise, it is called a **limit cardinal**. In particular, every positive integer $n \in \omega$ is a successor cardinal and ω is the smallest non-zero limit cardinal. By induction on $\alpha \in \Omega$ we define $\aleph_{\alpha+1} := \aleph_\alpha^+$, where $\aleph_0 := \omega$, and $\aleph_\alpha := \bigcup_{\delta\in\alpha} \aleph_\delta$ for limit ordinals α; notice that $\bigcup_{\delta\in\alpha} \aleph_\delta$ is

a cardinal (see Solution to Exercise 13.2.(a)). In particular, $\aleph_\omega$ is the smallest uncountable limit cardinal and $\aleph_1 = \aleph_0^+$ is the smallest uncountable cardinal. The collection $\{\aleph_\alpha : \alpha \in \Omega\}$ is the class of all infinite cardinals, i.e., for every infinite cardinal κ there is an $\alpha \in \Omega$ such that $\kappa = \aleph_\alpha$. Notice that the collection of cardinals is — like the collection of ordinals — a proper *class* and not a *set*. Now, we define addition, multiplication, and exponentiation of cardinals as follows:

***Cardinal Addition*:** For cardinals κ and μ, let

$$\kappa + \mu := |(\kappa \times \{0\}) \cup (\mu \times \{1\})|.$$

***Cardinal Multiplication*:** For cardinals κ and μ, let

$$\kappa \cdot \mu := |\kappa \times \mu|.$$

***Cardinal Exponentiation*:** For cardinals κ and μ, let

$$\kappa^\mu := |{}^\mu\kappa|.$$

As a consequence of these definitions we get the following result (for a proof see Solution to Exercise 13.3).

Fact 13.9. *Addition and multiplication of cardinals are associative and commutative, and we have the distributive law for multiplication over addition, and for all cardinals κ, λ, μ, we have*

$$\kappa^{\lambda+\mu} = \kappa^\lambda \cdot \kappa^\mu, \quad \kappa^{\mu\cdot\lambda} = (\kappa^\lambda)^\mu, \quad (\kappa \cdot \lambda)^\mu = \kappa^\mu \cdot \lambda^\mu.$$

Notice that there is a bijection $f : \mathscr{P}(\kappa) \to {}^\kappa 2$ given by

$$f(X)(\lambda) := \begin{cases} 1, & \lambda \in X \\ 0, & \lambda \notin X \end{cases}$$

for $\lambda < \kappa$. Hence $2^\kappa = |\mathscr{P}(\kappa)|$, and therefore, Theorem 13.8 states that for every cardinal κ we have $\kappa < 2^\kappa$. Now, the Continuum Hypothesis (CH) states that $2^{\aleph_0} = \aleph_1$, and the Generalised Continuum Hypothesis (GCH) states that $2^{\aleph_\alpha} = \aleph_{\alpha+1}$ for all $\alpha \in \Omega$. It is worth mentioning that CH — and consequently also GCH — is not provable within ZFC (see Halbeisen [22, Ch. 15]).

One of the main features of the arithmetic of infinite cardinals is given by the following result.

Proposition 13.10. *For any ordinal numbers $\alpha, \beta \in \Omega$, we have*

$$\aleph_\alpha + \aleph_\beta = \aleph_\alpha \cdot \aleph_\beta = \aleph_{\alpha\cup\beta} = \max\{\aleph_\alpha, \aleph_\beta\}.$$

In particular, for every infinite cardinal κ and for every $n \in \omega$, $\kappa^n = \kappa$.

Proof. It is enough to show that for all $\alpha \in \Omega$ we have $\aleph_\alpha \cdot \aleph_\alpha = \aleph_\alpha$. For $\alpha = 0$ we have $\aleph_0 = \omega$, and by the bijection $f : \omega \times \omega \to \omega$ defined by

$$f(\langle n, m\rangle) := n + \frac{(n+m)(n+m+1)}{2},$$

we have $|\omega \times \omega| = \omega$, which shows that $\aleph_0 \cdot \aleph_0 = \aleph_0$. Assume towards a contradiction that there exists an $\alpha \in \Omega$ such that $\aleph_\alpha \cdot \aleph_\alpha > \aleph_\alpha$. Then there exists a least ordinal α_0 with this property, i.e.,

$$\alpha_0 = \bigcap \{\alpha \in \Omega : \aleph_\alpha \cdot \aleph_\alpha > \aleph_\alpha\}.$$

On $\aleph_{\alpha_0} \times \aleph_{\alpha_0}$ we define an ordering $<$ by stipulating

$$\langle \gamma_1, \delta_1\rangle < \langle \gamma_2, \delta_2\rangle \quad \Longleftrightarrow \quad \begin{cases} (\gamma_1 \cup \delta_1) \in (\gamma_2 \cup \delta_2), \text{ or} \\ (\gamma_1 \cup \delta_1) = (\gamma_2 \cup \delta_2) \wedge \gamma_1 \in \gamma_2, \text{ or} \\ (\gamma_1 \cup \delta_1) = (\gamma_2 \cup \delta_2) \wedge \gamma_1 = \gamma_2 \wedge \delta_1 \in \delta_2. \end{cases}$$

With respect to the ordering $<$, the first few elements of $\aleph_{\alpha_0} \times \aleph_{\alpha_0}$ are

$$\begin{aligned} &\langle 0,0\rangle < \langle 0,1\rangle < \langle 1,0\rangle < \langle 1,1\rangle \\ &\qquad < \langle 0,2\rangle < \langle 1,2\rangle < \langle 2,0\rangle < \langle 2,1\rangle < \langle 2,2\rangle < \langle 0,3\rangle < \cdots \end{aligned}$$

and in general, for $\alpha \in \beta \in \aleph_{\alpha_0}$ we have $\langle \alpha, \beta\rangle < \langle \beta, \alpha\rangle$. It is easily verified that $<$ is a linear ordering on $\aleph_{\alpha_0} \times \aleph_{\alpha_0}$, and we leave it as an exercise to the reader to show that $<$ is even a well-ordering.

Now, let $\eta \in \Omega$ be the order type of the well-ordering $<$ on $\aleph_{\alpha_0} \times \aleph_{\alpha_0}$ and let $\Gamma : \eta \to \aleph_{\alpha_0} \times \aleph_{\alpha_0}$ be the unique order preserving bijection between η and $\aleph_{\alpha_0} \times \aleph_{\alpha_0}$. In particular, for any $\alpha, \alpha' \in \eta$ we get

$$\alpha \in \alpha' \quad \textit{if and only if} \quad \Gamma(\alpha) < \Gamma(\alpha').$$

Because $\aleph_{\alpha_0} < |\aleph_{\alpha_0} \times \aleph_{\alpha_0}|$ we have $\aleph_{\alpha_0} < |\eta|$. Let now $\langle \gamma_0, \delta_0\rangle := \Gamma(\aleph_{\alpha_0})$. Then, since $\gamma_0, \delta_0 \in \aleph_{\alpha_0}$, for $\nu = \max\{\gamma_0, \delta_0\}$ we have

$$|\nu| < \aleph_{\alpha_0} \quad \textit{and} \quad \aleph_{\alpha_0} \leq |\nu \times \nu|.$$

Thus, for $\aleph_\beta = |\nu|$ we get $\aleph_\beta < \aleph_{\alpha_0}$ and $\aleph_\beta \cdot \aleph_\beta > \aleph_\beta$, which is a contradiction to the choice of α_0. ⊣

For a cardinal κ, let $\operatorname{fin}(\kappa)$ denote the set of all finite subsets of κ, and let $\operatorname{seq}(\kappa)$ denote the set of all finite sequences which we can build with elements of κ. For a proof of the following fact see the Solution to Exercise 13.4.

Fact 13.11. *For every infinite cardinal κ, we have $\kappa = |\operatorname{fin}(\kappa)| = |\operatorname{seq}(\kappa)|$.*

Notes

In 1905, Zermelo began to axiomatise Set Theory, and in 1908 he published his first axiomatic system consisting of the seven above-mentioned axioms. In 1930, he presented in [62] his second axiomatic system, which he called the ZF-system, where he incorporated ideas of Fraenkel [10], Skolem [50], and von Neumann [38, 39, 40]. In fact, he added the Axiom Schema of Replacement (which was already used implicitly by Cantor in 1899) and the Axiom of Foundation to his former system, cancelled the Axiom of Infinity and did not explicitly mention the Axiom of Choice. More details can be found, for example, in the notes of Halbeisen [22, Ch. 3].

Exercises

13.0 Show that the Axiom of Empty Set follows from the Axiom Schema of Separation.

13.1 (a) Define by transfinite recursion addition of ordinals.

Hint: For each $\alpha \in \Omega$ define a class function F_α by stipulating $F_\alpha(x) := \emptyset$ if x is *not* a function; if x *is* a function, then let

$$F_\alpha(x) = \begin{cases} \alpha & \text{if } x = \emptyset, \\ x(\beta) \cup \{x(\beta)\} & \text{if } \operatorname{dom}(x) = \beta + 1 \text{ and } \beta \in \Omega, \\ \bigcup_{\delta \in \beta} x(\delta) & \text{if } \operatorname{dom}(x) = \beta \text{ and } \beta \in \Omega \setminus \{\emptyset\} \text{ is a limit ordinal}, \\ \emptyset & \text{otherwise.} \end{cases}$$

(b) Define by transfinite recursion multiplication of ordinals.

(c) Define by transfinite recursion exponentiation of ordinals.

13.2 (a) Show that for limit ordinals $\alpha \in \Omega$, $\bigcup_{\delta \in \alpha} \aleph_\delta$ is a cardinal.

(b) Show that any infinite cardinal κ is of the form $\aleph_\beta$ for some ordinal β.

13.3 Prove Fact 13.9.

13.4 (a) If κ is an infinite cardinal, then $\kappa = |\operatorname{seq}(\kappa)|$.

Hint: Notice that $|\operatorname{seq}(\kappa)| = \left|\bigcup_{n \in \omega} \kappa^n\right| = \aleph_0 \cdot \kappa$.

(b) If κ is an infinite cardinal, then $\kappa = |\operatorname{fin}(\kappa)|$.

13.5 Show that we can construct a model of PA within the axiom system Z. In particular, addition and multiplication of ordinals in ω can be defined without the Axiom Schema of Replacement (i.e., without the help of the Transfinite Recursion Theorem).

13.6 Let the unary relation symbol trans^* be defined as follows:

$$\operatorname{trans}^*(x) \ :\Longleftrightarrow\ \operatorname{trans}(x) \wedge \forall y \in x \big(\operatorname{trans}(y)\big)$$

Show that for all sets x we have:

$$\operatorname{trans}^*(x) \to \forall y_1, y_2 \in x \big(y_1 \subseteq y_2 \to (y_1 = y_2 \vee y_1 \in y_2)\big)$$

13.7 Prove that the following statements are equivalent for all sets x:

(a) ordinal(x)

(b) trans$^*(x)$

(c) $\text{ord}_{\in}(x) \wedge \text{trans}(x)$

13.8 Show that $\mathbb{R}$ is uncountable. In particular, show that $|\mathbb{R}| = |(0,1)| = |\mathscr{P}(\omega)|$, where $(0,1)$ is the set $\{r \in \mathbb{R} : 0 < r < 1\}$.

Remark: A construction of $\mathbb{R}$ is given in Chapter 17. In this exercise, you may use any well-known properties of $\mathbb{R}$.

Chapter 14
Models of Set Theory

Zermelo writes in [61, p. 262] that he was not able to show that the seven axioms for Set Theory given in that article are consistent. Even though it is essential whether a theory is consistent or not, we know that whenever a theory is strong enough to prove the axioms of PA, then there is no way to prove its consistency within this theory (see Chapter 11). Therefore, since we can prove within ZF that PA is consistent, we cannot prove the consistency of ZF within ZF. On the the other hand, we know that every consistent theory has a model. In particular, if ZF is consistent, then it has a model.

In what follows, we first show what models of ZF look like, then we briefly discuss non-standard models of ZF, and finally we give a construction of Gödel's model of ZFC, which shows that AC is relatively consistent with ZF.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_15

The Cumulative Hierarchy of Sets

Let us assume that ZF is consistent. Then, by Gödel's Completeness Theorem 5.5 we know that there must be a model $\mathbb{M} = (\mathbf{M}, \in^{\mathbb{M}})$ of ZF. Surprisingly, the domain $\mathbf{M}$ of the model $\mathbb{M}$ always has the structure of a cumulative hierarchy of sets. For this, we first construct within ZF a certain class $\mathbf{V}$ which has the structure of a cumulative hierarchy, and then we show that every set in $\mathbf{M}$ corresponds to an element in $\mathbf{V}$, which is — in abuse of notation — denoted $\mathbf{M} = \mathbf{V}^{\mathbb{M}}$. To construct $\mathbf{V}$, we first define the sets

$$\begin{aligned} \mathrm{V}_0 &:= \emptyset\,, \\ \mathrm{V}_\alpha &:= \bigcup_{\beta\in\alpha} \mathrm{V}_\beta \quad \text{if } \alpha \text{ is a limit ordinal}\,, \\ \mathrm{V}_{\alpha+1} &:= \mathscr{P}(\mathrm{V}_\alpha)\,, \end{aligned}$$

and then we define the class $\mathbf{V}$ by stipulating

$$\mathbf{V} := \bigcup_{\alpha\in\Omega} \mathrm{V}_\alpha.$$

The class $\mathbf{V}$ we defined is called the **cumulative hierarchy** of sets. To carry out the construction of $\mathbf{V}$ in the framework of ZF, we define the class function F by stipulating $F(x) := \emptyset$ if x is *not* a function; and if x *is* a function, then let

$$F(x) = \begin{cases} \emptyset & \text{if } x = \emptyset, \\ \mathscr{P}\big(x(\beta)\big) & \text{if } \mathrm{dom}(x) = \beta + 1 \text{ and } \beta \in \Omega, \\ \bigcup_{\delta\in\beta} x(\delta) & \text{if } \mathrm{dom}(x) = \beta \text{ and } \beta \in \Omega \setminus \{\emptyset\} \text{ is a limit ordinal}, \\ \emptyset & \text{otherwise.} \end{cases}$$

By the Transfinite Recursion Theorem 13.5, there is a unique class function G defined on Ω such that for each $\alpha \in \Omega$ we have $G(\alpha) = F(G|_\alpha)$. In particular, for each $\alpha \in \Omega$ we have $G(\alpha) = \mathrm{V}_\alpha$.

Notice that by the Axiom Schema of Replacement, for each $\alpha \in \Omega$, V_α is a *set*. Moreover, we can easily prove the following

Fact 14.0. *For any $\alpha, \beta \in \Omega$ we have:*

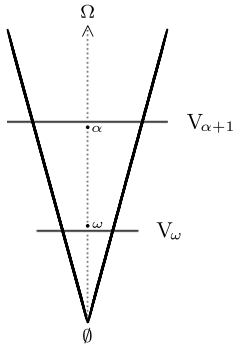
(a) V_α *is transitive.*

(b) *The class* $\mathbf{V}$ *is transitive.*

(c) *If* $\alpha \in \beta$, *then* $\mathrm{V}_\alpha \subsetneq \mathrm{V}_\beta$.

(d) $\alpha \subseteq \mathrm{V}_\alpha$ *and* $\alpha \in \mathrm{V}_{\alpha+1}$.

The cumulative hierarchy of sets is visualised by the following figure:



Before we can prove that the interpretation $\mathbf{V}^{\mathbb{M}}$ of $\mathbf{V}$ in the model $\mathbb{M}$ contains *all* sets of $\mathbf{M}$, we have to introduce the notion of *transitive closure*: Let S be an arbitrary set. By induction on $n \in \omega$, we define

$$S_0 = S, \qquad S_{n+1} = \bigcup S_n,$$

and finally

$$\mathrm{TC}(S) = \bigcup_{n\in\omega} S_n,$$

where $\bigcup_{n\in\omega} S_n := \bigcup\{S_n : n \in \omega\}$. For example, $x_1 \in S_1$ if and only if there is an $x_0 \in S_0$, such that $x_0 \ni x_1$, and in general, $x_{n+1} \in S_{n+1}$ if and only if

$$\exists x_0 \in S_0 \cdots \exists x_n \in S_n (x_0 \ni x_1 \ni \cdots \ni x_{n+1})\,.$$

Notice that by the Axiom of Foundation, every descending sequence of the form $x_0 \ni x_1 \ni \cdots$ must be finite. More precisely, every descending sequence $x_0 \ni x_1 \ni \cdots$ is of the form $x_0 \ni x_1 \ni \cdots \ni x_n$ for some $n \in \omega$.

By construction, $\mathrm{TC}(S)$ is transitive, i.e., $x \in \mathrm{TC}(S)$ implies $x \subseteq \mathrm{TC}(S)$, and we further have $S \subseteq \mathrm{TC}(S)$. Moreover, since every transitive set T must satisfy $\bigcup T \subseteq T$, it follows that the set $\mathrm{TC}(S)$ is the smallest transitive set which contains S. Thus,

$$\mathrm{TC}(S) = \bigcap\{T : T \supseteq S \text{ and } T \text{ is transitive}\}.$$

Consequently, the set $\mathrm{TC}(S)$ is called the **transitive closure** of S.

Now we are ready to show that $\mathbf{M} = \mathbf{V}^{\mathbb{M}}$.

THEOREM 14.1. *For every set x in $\mathbf{M}$ there is an ordinal α such that $x \in \mathrm{V}_\alpha^{\mathbb{M}}$. In particular, the domain of any model $\mathbb{M}$ of* ZF *has the structure of a cumulative hierarchy of sets.*

Proof. Let $\mathbb{M}$ be an arbitrary model of ZF with domain $\mathbf{M}$, where we write just "$\in$" instead of "$\in^{\mathbb{M}}$". Assume towards a contradiction that there exists a set x in $\mathbf{M}$ which does not belong to $\mathbf{V}^{\mathbb{M}}$. Let $\bar{x} := \mathrm{TC}(\{x\})$ (i.e., $\bar{x}$ is the transitive closure of $\{x\}$), and let $w := \{z \in \bar{x} : z \notin \mathbf{V}^{\mathbb{M}}\}$, i.e., $w = \bar{x} \setminus \{z' \in \bar{x} : \exists \alpha \in \Omega(z' \in \mathrm{V}_\alpha^{\mathbb{M}})\}$. Notice that $w \in \mathbf{M}$. Since $x \in w$ we have $w \neq \emptyset$, and by the Axiom of Foundation there is a $z_0 \in w$ such that $(z_0 \cap w) = \emptyset$. Since $z_0 \in w$ we have $z_0 \notin \mathbf{V}^{\mathbb{M}}$, which implies that $z_0 \neq \emptyset$, but for all $u \in z_0$ there is a least ordinal α_u such that $u \in \mathrm{V}_{\alpha_u}^{\mathbb{M}}$. By the Axiom Schema of Replacement, $\{\alpha_u : u \in z_0\}$ is a set, and moreover, $\alpha = \bigcup\{\alpha_u : u \in z_0\} \in \Omega$. This implies that $z_0 \subseteq \mathrm{V}_\alpha^{\mathbb{M}}$ and consequently we get $z_0 \in \mathrm{V}_{\alpha+1}^{\mathbb{M}}$, which contradicts the fact that $z_0 \notin \mathbf{V}^{\mathbb{M}}$. Thus, we have verified that in the model $\mathbb{M}$ we have $\mathbf{M} = \mathbf{V}^{\mathbb{M}}$. ⊣

Since the class $\mathbf{V}$ is defined as a union of iterated powers of $\emptyset$, i.e., $\mathbf{V} = \bigcup_{\alpha \in \Omega} \mathscr{P}^\alpha(\emptyset)$, $\mathbf{V}^{\mathbb{M}}$ depends on the interpretation $\mathscr{P}^{\mathbb{M}}$ in $\mathbb{M}$ of the unary power-set function symbol $\mathscr{P}$ in the extended language of set theory. It is therefore natural to ask whether we can interpret $\mathscr{P}$ in $\mathbb{M}$ also in a different way, say $\hat{\mathscr{P}}^{\mathbb{M}}$, such that the corresponding cumulative hierarchy $\hat{\mathbf{V}}^{\mathbb{M}}$ is a sub-class of $\mathbf{V}^{\mathbb{M}}$, but $\hat{\mathbb{M}} = (\hat{\mathbf{V}}^{\mathbb{M}}, \in)$ is still a model of ZF, or even of ZFC. By Theorem 14.1, we obtain that if $\hat{\mathbb{M}} \vDash \mathsf{ZF}$, then $\mathbf{V}^{\hat{\mathbb{M}}} = \hat{\mathbf{V}}^{\mathbb{M}}$. This is because $\mathscr{P}^{\hat{\mathbb{M}}}$, which is the interpretation of $\mathscr{P}$ in $\hat{\mathbb{M}}$, is the same function as $\hat{\mathscr{P}}^{\mathbb{M}}$. In particular, if $\hat{\mathbf{V}}^{\mathbb{M}}$ is a proper sub-class of $\mathbf{V}^{\mathbb{M}}$, then $\hat{\mathbb{M}} = (\mathbf{V}^{\hat{\mathbb{M}}}, \in)$ is a proper sub-model of $\mathbb{M} = (\mathbf{V}^{\mathbb{M}}, \in)$.

Now, since every model $\mathbb{M}$ of ZF is of the form $\mathbb{M} = (\mathbf{V}^{\mathbb{M}}, \in)$, where $\mathbf{V}$ is the formally defined cumulative hierarchy of sets, in abuse of terminology we will identify the model $\mathbb{M}$ with the class $\mathbf{V}$. In particular, we will consider $\mathbf{V}$ as the set-theoretic universe which contains all sets. Even though $\mathbf{V}$ is formally defined within ZF, this does not mean that ZF has a unique model, as we will see in the next section.

Non-Standard Models of ZF

On the one hand, in Chapter 7 we have constructed the standard model of PA with domain $\mathbb{N}$, on which we later defined the linear ordering $<$. On the other hand, in Chapter 13 we have constructed from ZF the set ω with the linear ordering given by the membership relation $\in$. Now, if, in a model $\mathbf{V}$ of ZF, the structure $(\omega, \in)$ is isomorphic to the structure $(\mathbb{N}, <)$, then we call $\mathbf{V}$ a **standard model** of ZF; otherwise, $\mathbf{V}$ is called a **non-standard model** of ZF. Recall that in ZF, a set x is called *finite* if and only if there exists a bijection between x and some element of ω. This leads to an alternative definition of standard models of ZF: A model $\mathbf{V}$ of ZF is a standard model

if and only if each set x which is finite in $\mathbf{V}$ is also finite with respect to the metamathematical notion of FINITENESS.

Before we show the existence of non-standard models of ZF, we would like to mention that we do not have a criteria to decide whether a model $\mathbf{V}$ of ZF is standard. In particular, when we assumed earlier that $\mathbf{M} \vDash \mathsf{ZF}$, it might have been that $\mathbf{M}$ was a non-standard model, and therefore we find that the non-standard models of ZF also have the structure of the cumulative hierarchy.

Now, we will show that if ZF is consistent, then there exists a non-standard model of ZF. For this purpose, we first extend the signature $\mathscr{L}_{\mathsf{ST}} = \{\in\}$ by adding countably many new constant symbols $\mathsf{c}_0, \mathsf{c}_1, \mathsf{c}_2, \ldots$, i.e., the new signature is $\{\in, \mathsf{c}_0, \mathsf{c}_1, \mathsf{c}_2, \ldots\}$. Then, we extend the axioms ZF by adding the formulae

$$\underbrace{\mathsf{c}_1 \in \mathsf{c}_0}_{\varphi_0}, \quad \underbrace{\mathsf{c}_2 \in \mathsf{c}_1}_{\varphi_1}, \quad \underbrace{\mathsf{c}_3 \in \mathsf{c}_2}_{\varphi_2}, \quad \ldots,$$

and let Ψ be the collection of these formulae. Now, if ZF has a model $\mathbf{V}$ and Φ is any finite subset of Ψ, then, by interpreting the finitely many c_i's $\mathsf{c}_0, \ldots, \mathsf{c}_n$ appearing in Ψ in a suitable way, e.g., by stipulating

$$\mathsf{c}_i := n - i,$$

$\mathbf{V}$ is also a model of $\mathsf{ZF} \cup \Phi$, which implies that $\mathsf{ZF} \cup \Phi$ is consistent. Thus, by the COMPACTNESS THEOREM, $\mathsf{ZF} \cup \Psi$ is also consistent and therefore has a model, say $\mathbf{V}^*$. Since $\mathbf{V}^* \vDash \mathsf{ZF} \cup \Psi$, we get that the Axiom of Foundation holds in $\mathbf{V}^*$. In particular, there must be a set $z \in \mathrm{TC}\left(\mathsf{c}_0^{\mathbf{V}^*}\right)$ such that

$$\mathbf{V}^* \vDash z \cap \mathrm{TC}\left(\mathsf{c}_0^{\mathbf{V}^*}\right) = \emptyset,$$

which implies that z must be different from all the sets $\mathsf{c}_n^{\mathbf{V}^*}$. On the other hand, by the Axiom of Foundation, the length of a decreasing sequence of the form

$$\mathsf{c}_0^{\mathbf{V}^*} \ni \mathsf{c}_1^{\mathbf{V}^*} \ni \mathsf{c}_2^{\mathbf{V}^*} \ni \cdots \ni z$$

must be finite in the sense of $\mathbf{V}^*$. In other words, the length of such a decreasing sequence must be an element of ω in the model $\mathbf{V}^*$, denoted by $\omega^{\mathbf{V}^*}$, which shows that $\omega^{\mathbf{V}^*}$ contains sets which are not finite with respect to the metamathematical notion of FINITENESS. In particular, the structures $(\omega^{\mathbf{V}^*}, \in)$ and $(\mathbb{N}, <)$ are not isomorphic, and hence, $\mathbf{V}^*$ is a non-standard model of ZF.

As a matter of fact, we would like to mention that from the consistency of ZF we obtain the existence of non-standard models of ZF, but without using the metamathematical notion of FINITENESS, we do not obtain the existence of standard models of ZF.

Gödel's Incompleteness Theorems for Set Theory

In this section, we indicate how Gödel's Incompleteness Theorems can be transferred to ZF and ZFC, respectively. In Chapter 16, we shall see that within ZF we can construct a model of PA with domain ω. In particular, we obtain

$$\mathrm{Con}(\mathsf{ZF}) \implies \mathrm{Con}(\mathsf{PA})\,.$$

Therefore, with respect to ω, we can define within ZF the non-logical symbols of $\mathscr{L}_{\mathsf{PA}}$ and can extend the language $\mathscr{L}_{\mathsf{ST}}$ to the language $\mathscr{L} := \mathscr{L}_{\mathsf{ST}} \cup \mathscr{L}_{\mathsf{PA}} \cup \{\omega\}$. Furthermore, we can extend the theory ZF to the theory $\mathsf{T} := \mathsf{ZF} \cup \mathsf{PA}$. Now, since $\mathscr{L} \supseteq \mathscr{L}_{\mathsf{PA}}$ is a gödelisable language and T is a gödelisable $\mathscr{L}$-theory, we can apply Theorem 10.11 and obtain that if ZF is consistent, then it is incomplete — the same applies to ZFC.
Moreover, within ZF we can define the $\mathscr{L}_{\mathsf{ST}}$-sentence

$$\mathrm{con}_{\mathsf{ZF}} :\Longleftrightarrow \neg\,\mathrm{prv}_{\mathsf{ZF}}(\ulcorner \emptyset = \{\emptyset\} \urcorner)$$

and show that if ZF is consistent, then $\mathsf{ZF} \nvdash \mathrm{con}_{\mathsf{ZF}}$ — where the same applies to ZFC.

Summing up, we obtain the following

Theorem 14.2 (Gödel's Incompleteness Theorems for Set Theory).
(a) *If* ZF *is consistent, then it is incomplete.*
(b) *If* ZF *is consistent, then* $\mathsf{ZF} \nvdash \mathrm{con}_{\mathsf{ZF}}$.

Correspondingly, the same holds for ZFC.

As a matter of fact, we would like to mention that for T as above, on the one hand we have

$$\mathrm{Con}(\mathsf{T}) \implies \mathrm{Con}(\mathsf{PA})\,,$$

but on the other hand, for $\mathrm{con}_{\mathsf{PA}} :\Longleftrightarrow \neg\,\mathrm{prv}_{\mathsf{PA}}(\ulcorner \mathsf{0} = \mathsf{1} \urcorner)$ we have

$$\mathrm{Con}(\mathsf{T}) \implies \mathsf{T} \nvdash \mathrm{con}_{\mathsf{PA}}\,.$$

Absoluteness

In order to investigate sub-models of $\mathbf{V}$ of fragments of ZFC, where we identify the class $\mathbf{V}$ with the model $(\mathbf{V}, \in)$, the notion of absoluteness will be crucial. If $\mathbf{M}$ is the domain of a sub-model $\mathbb{M}$ of $\mathbf{V}$, then $\mathbf{M}$ is a collection of sets of $\mathbf{V}$. However, $\mathbf{M}$ itself is not necessarily a set in $\mathbf{V}$, but as a collection of sets of $\mathbf{V}$, $\mathbf{M}$ is a class in $\mathbf{V}$. For a class $\mathbf{M}$ in $\mathbf{V}$, we can **relativise** a formula

$\varphi(\nu_1, \ldots, \nu_n)$ to the model $\mathbb{M} = (\mathbf{M}, \in)$ such that for all $x_1, \ldots, x_n \in \mathbf{M}$, the relativised formula $\varphi^{\mathbb{M}}(\nu_1, \ldots, \nu_n)$ has the following property:

$$\varphi^{\mathbb{M}}(x_1, \ldots, x_n) \iff \mathbb{M} \vDash \varphi(x_1, \ldots, x_n)$$

To do this, by THEOREM 1.7, we may assume that φ only contains $\neg$ and $\wedge$ as logical operators and $\exists$ as quantifier. We define $\varphi^{\mathbb{M}}$ recursively as follows:

$$\begin{aligned}
(x_i \in x_j)^{\mathbb{M}} &: \iff \mathbf{V} \vDash x_i \in x_j \\
(x_i = x_j)^{\mathbb{M}} &: \iff \mathbf{V} \vDash x_i = x_j \\
(\neg\varphi)^{\mathbb{M}} &: \iff \mathbf{V} \vDash \neg\varphi^{\mathbb{M}} \\
(\varphi \wedge \psi)^{\mathbb{M}} &: \iff \mathbf{V} \vDash \varphi^{\mathbb{M}} \wedge \psi^{\mathbb{M}} \\
(\exists\nu\, \varphi)^{\mathbb{M}} &: \iff \exists x \in \mathbf{M} :\ \mathbf{V} \vDash \varphi^{\mathbb{M}}(x)
\end{aligned}$$

If $\mathbf{M}_1$ and $\mathbf{M}_2$ are two classes in $\mathbf{V}$ such that $\mathbf{M}_1 \subseteq \mathbf{M}_2$, then an $\mathscr{L}_{\mathsf{ST}}$-formula $\varphi(\nu_1, \ldots, \nu_n)$ is called **absolute** between the models $\mathbb{M}_1 = (\mathbf{M}_1, \in)$ and $\mathbb{M}_2 = (\mathbf{M}_2, \in)$, if for all $x_1, \ldots, x_n \in \mathbf{M}_1$,

$$\varphi^{\mathbb{M}_1}(x_1, \ldots, x_n) \iff \varphi^{\mathbb{M}_2}(x_1, \ldots, x_n)\,.$$

In the case when $\mathbb{M}_2 = (\mathbf{V}, \in)$, then we say that φ is **absolute** for $\mathbb{M}_1$.

As for $\mathscr{L}_{\mathsf{PA}}$-formulae, we say that an $\mathscr{L}_{\mathsf{ST}}$-formula φ is a **Δ-formula** if it is built up from atomic formulae using $\neg$, $\wedge$, $\vee$, and bounded quantification, i.e., $\forall\nu \in \tau$ and $\exists\nu \in \tau$ for some term τ.

EXAMPLE 14.3. The formula $x \subseteq y$ can be expressed by a Δ-formula, since

$$x \subseteq y \iff \forall z \in x(z \in y).$$

Moreover, the formula stating that x is an ordinal is equivalent to a Δ-formula, since $\mathrm{trans}(x)$ is a Δ-formula and by EXERCISES 13.6 and 13.7.(b) we have

$$\mathrm{ordinal}(x) \iff \mathrm{trans}(x) \wedge \forall y \in x\big(\mathrm{trans}(y)\big)\,.$$

A proof of the following result can be found in Kunen [31]).

FACT 14.4. *Let $\mathbf{M}$ be a class and let $\mathbb{M} = (\mathbf{M}, \in)$. Then every $\mathscr{L}_{\mathsf{ST}}$-formula which is logically equivalent to a Δ-formula is absolute for $\mathbb{M}$.*

The following result is useful because it provides simple criteria for the axioms of ZF being valid in transitive classes (a proof can be found in Kunen [31]).

LEMMA 14.5. *Let* $\mathbf{M}$ *be a transitive class (i.e.,* $\mathbf{M}$ *is a class and for all* x, y *we have* $x \in y \in \mathbf{M} \to x \in \mathbf{M}$*), and let* $\mathbb{M} = (\mathbf{M}, \in)$*. Then we have:*

(a) $\mathbb{M} \vDash$ *Axiom of Extensionality*

(b) $\mathbb{M} \vDash$ *Axiom of Foundation*

(c) $\mathbb{M} \vDash$ *Axiom of Pairing* $\iff \forall x, y \in \mathbf{M}\big(\{x, y\} \in \mathbf{M}\big)$

(d) $\mathbb{M} \vDash$ *Axiom of Union* $\iff \forall x \in \mathbf{M}\big(\bigcup x \in \mathbf{M}\big)$

(e) $\mathbb{M} \vDash$ *Axiom of Infinity* $\iff \omega \in \mathbf{M}$

(f) $\mathbb{M} \vDash$ *Axiom of Power Set* $\iff \forall x \in \mathbf{M}\big(\mathscr{P}(x) \cap \mathbf{M} \in \mathbf{M}\big)$

Gödel's Constructible Model L

In this section, we present Gödel's constructible universe $\mathbf{L}$, which essentially consists of all sets which can be "described" and is therefore the smallest model of Set Theory. More precisely, in each step of the construction of $\mathbf{L}$ within some ground model $\mathbf{V} \vDash \mathsf{ZF}$, we only add sets which are definable from already constructed sets M by taking sets of the form

$$x = \big\{y \in M : (M, \in) \vDash \varphi(y, p_1, \ldots, p_n)\big\},$$

for some formulae φ and parameters $p_1, \ldots, p_n \in M$. The problem that we encounter at this point is that we do not know whether the satisfaction relation $(M, \in) \vDash \varphi(y, p_1, \ldots, p_n)$ can be defined by an $\mathscr{L}_{\mathsf{ST}}$-formula, which is crucial in order to apply the Axiom Schema of Separation to obtain the existence of the set x. To achieve this, we first gödelise $\mathscr{L}_{\mathsf{ST}}$-formulae within ZF in a similar way as we gödelised $\mathscr{L}_{\mathsf{PA}}$-formulae within PA in Chapter 9. However, in Set Theory the gödelisation is much simpler. We first gödelise atomic $\mathscr{L}_{\mathsf{ST}}$-formulae as follows:

$$\ulcorner v_i \in v_j \urcorner := \langle 0, i, j\rangle$$
$$\ulcorner v_i = v_j \urcorner := \langle 1, i, j\rangle$$

Suppose that φ and ψ have already been gödelised. Then we define:

$$\ulcorner \neg\varphi \urcorner := \langle 2, \ulcorner \varphi \urcorner\rangle$$
$$\ulcorner \varphi \wedge \psi \urcorner := \langle 3, \ulcorner \varphi \urcorner, \ulcorner \psi \urcorner\rangle$$
$$\ulcorner \varphi \vee \psi \urcorner := \langle 4, \ulcorner \varphi \urcorner, \ulcorner \psi \urcorner\rangle$$
$$\ulcorner \varphi \to \psi \urcorner := \langle 5, \ulcorner \varphi \urcorner, \ulcorner \psi \urcorner\rangle$$
$$\ulcorner \exists v_i \varphi \urcorner := \langle 6, i, \ulcorner \varphi \urcorner\rangle$$
$$\ulcorner \forall v_j \varphi \urcorner := \langle 7, j, \ulcorner \varphi \urcorner\rangle$$

We can then define the set of all codes of formalised $\mathscr{L}_{\mathsf{ST}}$-formulae by stipulating:

$$\begin{aligned} f \in \mathrm{Fml} :\Longleftrightarrow\ & \exists n \in \omega \text{ and } c : n+1 \to \mathrm{V}_\omega \text{ such that } f = c(n) \text{ and} \\ & \forall m \leq n\, \exists i, j \in \omega\, \exists k, l < m \big(c(m) = \langle 0, i, j\rangle \vee c(m) = \langle 1, i, j\rangle \\ & \quad \vee c(m) = \langle 2, c(k)\rangle \vee c(m) = \langle 3, c(k), c(l)\rangle \\ & \quad \vee c(m) = \langle 4, c(k), c(l)\rangle \vee c(m) = \langle 5, c(k), c(l)\rangle \\ & \quad \vee c(m) = \langle 6, i, c(k)\rangle \vee c(m) = \langle 7, i, c(k)\rangle\big)\,. \end{aligned}$$

Notice that Fml is a set which belongs to **V**. Furthermore, by induction on the construction of φ, one can show that for every $\mathscr{L}_{\mathsf{ST}}$-formula φ we have $\ulcorner\varphi\urcorner \in \mathrm{Fml}$.

Let M be a set or a class. As usual, we shall identify the $\mathscr{L}_{\mathsf{ST}}$-structure $(M, \in)$ with M. Furthermore, let $\vec{x} = \langle x_0, \dots, x_n\rangle \in M^{n+1}$, i.e., $\vec{x}$ is a function with $\operatorname{dom}\vec{x} = n+1$ and $\vec{x}(i) = x_i$ for all $0 \leq i \leq n$. Now, we formalise the satisfaction relation for $\mathscr{L}_{\mathsf{ST}}$-formulae with free variables among $\{v_0, \dots, v_n\}$ as follows:

$$\begin{aligned} M^{\ulcorner}\vDash^{\urcorner}\langle 0, i, j\rangle[\vec{x}] &:\Longleftrightarrow x_i \in x_j \\ M^{\ulcorner}\vDash^{\urcorner}\langle 1, i, j\rangle[\vec{x}] &:\Longleftrightarrow x_i = x_j \\ M^{\ulcorner}\vDash^{\urcorner}\langle 2, f\rangle[\vec{x}] &:\Longleftrightarrow \neg M^{\ulcorner}\vDash^{\urcorner} f[\vec{x}] \\ M^{\ulcorner}\vDash^{\urcorner}\langle 3, f, g\rangle[\vec{x}] &:\Longleftrightarrow M^{\ulcorner}\vDash^{\urcorner} f[\vec{x}] \wedge M^{\ulcorner}\vDash^{\urcorner} g[\vec{x}] \\ M^{\ulcorner}\vDash^{\urcorner}\langle 6, i, f\rangle[\vec{x}] &:\Longleftrightarrow \exists a \in M(M^{\ulcorner}\vDash^{\urcorner} f[\vec{x}\tfrac{i}{a}]), \end{aligned}$$

where for $0 \leq k \leq n$,

$$(\vec{x}\tfrac{i}{a})_k = \begin{cases} x_k & k \neq i, \\ a & k = i. \end{cases}$$

FACT 14.6. *Let M be a class in* **V** *and let $\varphi(\nu_0, \dots, \nu_n)$ be an $\mathscr{L}_{\mathsf{ST}}$-formula. Then for any $x_0, \dots, x_n \in M$ and $\vec{x} = \langle x_0, \dots, x_n\rangle$ we have*

$$M \vDash \varphi(x_0, \dots, x_n) \iff M^{\ulcorner}\vDash^{\urcorner}{}^{\ulcorner}\varphi^{\urcorner}[\vec{x}]\,.$$

Let now M be a set in **V**. We say that a set x is **definable** over M, if there exist an $\mathscr{L}_{\mathsf{ST}}$-formula $\varphi(\nu_0, \dots, \nu_n)$ and $p_1, \dots, p_n \in M$ such that

$$x = \big\{y \in M : M \vDash \varphi(y, p_1, \dots, p_n)\big\} = \big\{y \in M : \varphi^M(y, p_1, \dots, p_n)\big\}\,.$$

Furthermore, we define

$$\mathrm{Def}(M) := \big\{x \in \mathscr{P}(M) : x \textit{ is definable over } M\big\}\,.$$

Notice that since $M \in \mathbf{V}$, $\mathrm{Def}(M)$ belongs to $\mathbf{V}$ as well. Moreover, by definition we have $\mathrm{Def}(M) \subseteq \mathscr{P}(M)$.

In order to achieve that the set $\mathrm{Def}(M)$ itself is definable within $\mathbf{V}$, we have to use the formalised satisfaction relation. The reason is that we cannot quantify over $\mathscr{L}_{\mathsf{ST}}$-formulae, but need to quantify over elements of Fml instead. By the above definitions we obtain

$$x \in \mathrm{Def}(M) :\Longleftrightarrow \\ \exists f \in \mathrm{Fml}\, \exists n \in \omega\, \exists \vec{p} \in M^n\, \forall y \in M \big(y \in x \leftrightarrow M \ulcorner \vDash \urcorner f[\langle y, \vec{p} \rangle]\big)\,,$$

which shows that we can indeed define the set $\mathrm{Def}(M)$ within $\mathbf{V}$. Thus, by transfinite recursion we can define within $\mathbf{V}$ the **constructible hierarchy** as follows:

$$\begin{aligned} \mathrm{L}_0 &= \emptyset \\ \mathrm{L}_\alpha &= \bigcup_{\beta \in \alpha} \mathrm{L}_\beta \quad \text{if } \alpha \text{ is a limit ordinal} \\ \mathrm{L}_{\alpha+1} &= \mathrm{Def}(\mathrm{L}_\alpha) \end{aligned}$$

The **constructible universe** is then defined as

$$\mathbf{L} = \bigcup_{\alpha \in \Omega} \mathrm{L}_\alpha.$$

By transfinite recursion one can show that the class $\mathbf{L}$ is definable within $\mathbf{V}$. The goal is now to show that $\mathbf{L}$ is a model of ZFC. In order to simplify the notation, we will not distinguish between the sets L_α and the corresponding $\mathscr{L}_{\mathsf{ST}}$-structures $(\mathrm{L}_\alpha, \in)$; the same applies to the class $\mathbf{L}$ and the $\mathscr{L}_{\mathsf{ST}}$-structure $(\mathbf{L}, \in)$.

The following result shows that the structure of $\mathbf{L}$ is the same as the structure of $\mathbf{V}$. In particular, we obtain that $\mathbf{L}$ contains the same ordinals as $\mathbf{V}$.

Proposition 14.7.

(a) *If* $\alpha \in \beta \in \Omega$, *then* $\mathrm{L}_\alpha \subsetneq \mathrm{L}_\beta$.
(b) *For every* $\beta \in \Omega$, L_β *is transitive.*
(c) *For every* $\beta \in \Omega$, $\mathrm{L}_\beta \subseteq \mathrm{V}_\beta$.
(d) *For every* $\beta \in \Omega$, $\mathrm{L}_\beta \cap \Omega = \beta$.

Proof. The proof is by induction on $\beta \in \Omega$. For $\beta = 0$ we have $\mathrm{L}_\beta = \emptyset$ and therefore, the claim is trivial. The limit case follows immediately from the definition of L_β at the limit stage. Hence we may assume that $\beta = \gamma + 1$ is a successor ordinal. For (a) it suffices to check that $\mathrm{L}_\gamma \subseteq \mathrm{L}_\beta$ and $\mathrm{L}_\gamma \in \mathrm{L}_\beta$. For the first claim, let $x \in \mathrm{L}_\gamma$. By transitivity of L_γ we have $x \subseteq \mathrm{L}_\gamma$ and hence

$$x = \{y \in \mathrm{L}_\gamma : (\mathrm{L}_\gamma, \in) \vDash y \in x\} \in \mathrm{Def}(\mathrm{L}_\gamma) = \mathrm{L}_\beta.$$

For the second claim, note that

$$\mathrm{L}_\gamma = \{x \in \mathrm{L}_\gamma : (\mathrm{L}_\gamma, \in) \vDash x = x\} \in \mathrm{L}_\beta.$$

For (b) let $x \in \mathrm{L}_\beta$. Then (a) and the transitivity of L_γ imply $x \subseteq \mathrm{L}_\gamma \subseteq \mathrm{L}_\beta$. By our induction hypothesis we have $\mathrm{L}_\gamma \subseteq \mathrm{V}_\gamma$ and thus $\mathscr{P}(\mathrm{L}_\gamma) \subseteq \mathscr{P}(\mathrm{V}_\gamma) = \mathrm{V}_\beta$. Since $\mathrm{V}_\beta = \mathrm{Def}(\mathrm{L}_\gamma) \subseteq \mathscr{P}(\mathrm{L}_\gamma)$, (c) holds. For (d), observe first that $\mathrm{L}_\beta \cap \Omega \subseteq \mathrm{V}_\beta \cap \Omega = \beta$. For the reverse inclusion, by induction we have $\mathrm{L}_\gamma \cap \Omega = \gamma$ and therefore, by absoluteness of the formula ordinal(x)', we finally have

$$\gamma = \big\{\delta \in \mathrm{L}_\gamma : \mathrm{ordinal}(\delta)\big\} = \big\{\delta \in \mathrm{L}_\gamma : (\mathrm{L}_\gamma, \in) \vDash \mathrm{ordinal}(\delta)\big\} \in \mathrm{Def}(\mathrm{L}_\gamma) = \mathrm{L}_\beta\,. \qquad \dashv$$

Theorem 14.8 (Lévy's Reflection Theorem). *Let φ be an $\mathscr{L}_{\mathsf{ST}}$-formula and $\alpha \in \Omega$. Then there is a $\beta \in \Omega$ with $\beta \geq \alpha$ such that φ is absolute between L_β and* **L**.

Proof. By Theorem 1.7, we may assume that φ only contains $\neg$ and $\wedge$ as logical operators and $\exists$ as quantifier. Suppose that $\varphi_0, \ldots, \varphi_m$ is a list of all subformulae of φ with the property that all proper subformulae of φ_i occur among $\varphi_0, \ldots, \varphi_{i-1}$, i.e., φ is the formula φ_m. Furthermore, assume that all free variables of $\varphi_0, \ldots, \varphi_m$ are among $\{v_0, \ldots, v_n\}$. For the sake of simplicity, for $\vec{x} = \langle x_0, \ldots, x_n\rangle \in \mathbf{L}^{n+1}$ we define

$$\varphi(\vec{x}) :\equiv \varphi(x_0, \ldots, x_n)\,.$$

For every $i \leq m$, we define a class function $F_i : \mathbf{L}^{n+1} \to \Omega$ by stipulating

$$F_i(\vec{x}) = \begin{cases} \min_{\in}\left\{\beta \in \Omega : \exists b \in \mathrm{L}_\beta\, \psi^{\mathbf{L}}(\frac{\nu}{b}, \vec{x})\right\} & \text{if } \varphi_i(\vec{x}) \equiv \exists \nu\, \psi(\nu, \vec{x}) \\ & \quad \text{and } \exists a \in \mathbf{L}\, \psi^{\mathbf{L}}(\frac{\nu}{a}, \vec{x}), \\ 0 & \text{otherwise.} \end{cases}$$

Notice that the class function F_i guarantees that if φ_i is of the form $\exists \nu\, \psi(\nu)$ and there is a witness in **L** for $\psi(\nu)$, then there is already a witness for $\psi(\nu)$ in L_β.

Now, we recursively define a sequence of ordinals $\langle \beta_k : k \in \omega\rangle$ as follows: Let $\beta_0 := \alpha$. Suppose that β_k is given. Then let

$$\beta_{k+1} := \bigcup \big\{F_i(\vec{x}) : i \leq m \wedge \{x_0, \ldots, x_n\} \subseteq \mathrm{L}_{\beta_k}\big\}\,.$$

Finally, set $\beta := \bigcup_{k\in\omega} \beta_k$. We show by induction that for every $i \leq m$, the formula φ_i is absolute between L_β and **L**. Suppose that $\{x_0, \ldots, x_n\} \subseteq \mathrm{L}_\beta$.

Case 1. φ_i is atomic. Then φ_i is obviously absolute between L_β and **L**.

Case 2. φ_i is $\neg\varphi_j$ for some $j < i$ and φ_j is absolute between L_β and $\mathbf{L}$. By assumption, we have $\varphi_j^{\mathrm{L}_\beta}(\vec{x}) \iff \varphi_j^{\mathbf{L}}(\vec{x})$ and hence

$$\varphi_i^{\mathrm{L}_\beta}(\vec{x}) \iff \neg\varphi_j^{\mathrm{L}_\beta}(\vec{x}) \iff \neg\varphi_j^{\mathbf{L}}(\vec{x}) \iff \varphi_i^{\mathbf{L}}(\vec{x})\,.$$

Case 3. φ_i is $\varphi_j \wedge \varphi_k$ for $j, k < i$ and φ_j, φ_k are absolute between L_β and $\mathbf{L}$. Then we have

$$\varphi_i^{\mathrm{L}_\beta}(\vec{x}) \iff \varphi_j^{\mathrm{L}_\beta}(\vec{x}) \wedge \varphi_k^{\mathrm{L}_\beta}(\vec{x}) \iff \varphi_j^{\mathbf{L}}(\vec{x}) \wedge \varphi_k^{\mathbf{L}}(\vec{x}) \iff \varphi_i^{\mathbf{L}}(\vec{x})\,.$$

Case 4. φ_i is $\exists\nu\varphi_j$ for some $j < i$ such that φ_j is absolute between L_β and $\mathbf{L}$. Then on the one hand, since $\mathbf{L}$ and $\mathrm{L}_\beta \subseteq \mathbf{L}$, we have

$$\varphi_i^{\mathrm{L}_\beta}(\vec{x}) \Longrightarrow \exists\nu \in \mathrm{L}_\beta\, \varphi_j^{\mathrm{L}_\beta}(\vec{x}) \Longrightarrow \exists\nu \in \mathbf{L}\, \varphi_j^{\mathbf{L}}(\vec{x}) \Longrightarrow \varphi_i^{\mathbf{L}}(\vec{x})\,,$$

and on the other hand, by construction of β and since $\{x_0, \ldots, x_n\} \subseteq \mathrm{L}_\beta$, we have

$$\varphi_i^{\mathbf{L}}(\vec{x}) \Longrightarrow \exists\nu \in \mathbf{L}\, \varphi_j^{\mathbf{L}}(\vec{x}) \Longrightarrow \exists\nu \in \mathrm{L}_\beta\, \varphi_j^{\mathrm{L}_\beta}(\vec{x}) \Longrightarrow \varphi_i^{\mathrm{L}_\beta}(\vec{x})\,.$$

Hence, φ_i is absolute between L_β and $\mathbf{L}$. ⊣

$\mathbf{L} \vDash \mathsf{ZF}$

Now we are ready to show the following

Theorem 14.9. *The constructible universe is a model of* ZF*, i.e.,*

$$\mathbf{L} \vDash \mathsf{ZF}\,.$$

Proof. First notice that since $\emptyset \in \mathbf{L}$, the Axiom of Empty Set holds in $\mathbf{L}$, and since $\mathbf{L}$ is a transitive class, by Lemma 14.5 also the Axiom of Extensionality and the Axiom of Foundation hold in $\mathbf{L}$.

By applying Lemma 14.5, we now show that the following five axioms of ZF hold in $\mathbf{L}$:

Axiom of Pairing. Let $a, b \in \mathbf{L}$ and let $\alpha \in \Omega$ be such that $a, b \in \mathrm{L}_\alpha$. Then

$$\{a, b\} = \big\{x \in \mathrm{L}_\alpha : x = a \vee x = b\big\} \in \mathrm{L}_{\alpha+1} \subseteq \mathbf{L}.$$

Axiom of Union. Let $a \in \mathbf{L}$ and let $\alpha \in \Omega$ such that $a \in \mathrm{L}_\alpha$. Since $\mathbf{L}_\alpha$ is transitive, we have $\bigcup a \subseteq \mathrm{L}_\alpha$, and thus,

$$\begin{aligned}\bigcup a &= \big\{x \in \mathrm{L}_\alpha : \exists y(x \in y \wedge y \in a)\big\}\\ &= \big\{x \in \mathrm{L}_\alpha : \mathrm{L}_\alpha \vDash \exists y \in a(x \in y)\big\} \in \mathrm{L}_{\alpha+1}.\end{aligned}$$

Axiom of Infinity. By Proposition 14.7 we have $\omega \in \mathrm{L}_{\omega+1} \subseteq \mathbf{L}$.

Axiom of Power Set. Let $a \in \mathbf{L}$. By the Axiom of Power Set in $\mathbf{V}$ we obtain that $\mathscr{P}(a) \cap \mathbf{L}$ is a set, and thus, there is an $\alpha \in \Omega$ such that $\mathscr{P}(a) \cap \mathbf{L} \subseteq \mathrm{L}_\alpha$. Therefore, we have

$$\mathscr{P}(a) \cap \mathbf{L} = \big\{x \in \mathrm{L}_\alpha : x \subseteq a\big\} = \big\{x \in \mathrm{L}_\alpha : \mathrm{L}_\alpha \vDash x \subseteq a\big\} \in \mathrm{L}_{\alpha+1}\,,$$

since the subset relation is absolute.

It remains to show that the two axiom schema of ZF hold in $\mathbf{L}$ as well:

Axiom Schema of Separation. Let $\varphi(\nu_0, \ldots, \nu_n)$ be an $\mathscr{L}_{\mathsf{ST}}$-formula such that $\mathrm{free}(\varphi) \subseteq \{\nu_0, \ldots, \nu_n\}$, let $\{x, p_1, \ldots, p_n\} \subseteq \mathbf{L}$ and let $\vec{p} = \langle p_1, \ldots, p_n\rangle$. It suffices to prove that

$$\big\{y \in x : \mathbf{L} \vDash \varphi(y, \vec{p})\big\} \in \mathbf{L}.$$

Let $\alpha \in \Omega$ be such that $\{x, p_1, \ldots, p_n\} \subseteq \mathrm{L}_\alpha$. By Lévy's Reflection Theorem, there is a $\beta \in \Omega$ with $\beta \geq \alpha$ such that φ is absolute between L_β and $\mathbf{L}$. Then by transitivity of L_β we have

$$\begin{aligned}\big\{y \in x : \mathbf{L} \vDash \varphi(y, \vec{p})\big\} &= \big\{y \in x : \mathrm{L}_\beta \vDash \varphi(y, \vec{p})\big\} \\ &= \big\{y \in \mathrm{L}_\beta : \mathrm{L}_\beta \vDash \psi(y, \vec{p}, x)\big\} \in \mathrm{L}_{\beta+1},\end{aligned}$$

where $\psi(v_0, \ldots, v_{n+1})$ is the formula $v_0 \in v_{n+1} \wedge \varphi(v_0, \ldots, v_n)$.

Axiom Schema of Replacement. Let φ be an $\mathscr{L}_{\mathsf{ST}}$-formula with $n+2$ free variables and let $\{p_1, \ldots, p_n, A\} \in \mathbf{L}$. Suppose that φ defines a class function in $\mathbf{L}$, i.e.,

$$\forall x \in \mathbf{L}\, \exists! y \in \mathbf{L}\, \varphi(x, y, \vec{p}).$$

Consider the function F on A given by

$$F(x) = \min\nolimits_{\in} \big\{\beta \in \Omega : \exists y \in \mathrm{L}_\beta\, \varphi^{\mathbf{L}}(x, y, \vec{p})\big\}\,.$$

Since $A \in \mathbf{L}$, by the Axiom Schema of Replacement applied in $\mathbf{V}$ we have $X := F[A] \in \mathbf{V}$. Now, X is a set of ordinals, and therefore, $\alpha := \bigcup X \in \Omega$. Consider

$$B := \big\{y \in \mathbf{L} : \exists x \in A\, \varphi^{\mathbf{L}}(x, y, \vec{p})\big\}\,.$$

Since φ is functional and using the Axiom Schema of Replacement in $\mathbf{V}$, we obtain that B is a set satisfying $B \subseteq \mathrm{L}_\alpha$. Now, by Lévy's Reflection Theorem there is an ordinal $\beta \geq \alpha$ such that $A \in \mathrm{L}_\beta$ and φ is absolute between L_β and $\mathbf{L}$. Hence,

$$\begin{aligned}B &= \big\{y \in \mathbf{L} : \exists x \in A\, \varphi^{\mathbf{L}}(x, y, \vec{p})\big\} \\ &= \big\{y \in \mathrm{L}_\beta : \exists x \in A\, \varphi^{\mathrm{L}_\beta}(x, y, \vec{p})\big\} \in \mathrm{L}_{\beta+1}\,.\end{aligned}$$

Therefore, we have shown that $\mathbf{L}$ satisfies all axioms of ZF, i.e., $\mathbf{L} \vDash \mathsf{ZF}$. $\dashv$

$\mathbf{L} \vDash \mathsf{ZFC}$

In this section, we will show that it is possible to define a class-sized well-ordering of Gödel's model $\mathbf{L}$, from which it follows that $\mathbf{L}$ is in fact a model of the Axiom of Choice. Since our ground model $\mathbf{V}$, in which we carried out the construction of $\mathbf{L}$, was just a model of ZF, it may come as a surprise that $\mathbf{L} \vDash \mathsf{ZFC}$. In particular, we obtain that in *every* model $\mathbf{V}$ of ZF there exists a sub-model of ZFC, no matter whether or not AC holds in $\mathbf{V}$.

The idea of the proof is to show that each level L_α of the constructible hierarchy can be well-ordered, which can be used to construct a well-ordering of $\mathbf{L}$.

Suppose that for some $\alpha \in \Omega$, a well-ordering $\prec_\alpha$ of L_α is given such that for any $\beta \in \Omega$ with $\alpha < \beta$, $\prec_\beta$ is an *end-extension* of $\prec_\alpha$, i.e., $\prec_\beta$ satisfies the following two properties:

- For any $x, y \in \mathrm{L}_\alpha$, if $x \prec_\alpha y$ then $x \prec_\beta y$.
- If $x \in \mathrm{L}_\alpha$ and $y \in \mathrm{L}_\beta \setminus \mathrm{L}_\alpha$, then $x \prec_\beta y$.

Assuming the existence of such a well-ordering $\prec_\alpha$ for every $\alpha \in \Omega$, we obtain a class-sized well-ordering of $\mathbf{L}$ by stipulating

$$x \prec_{\mathbf{L}} y :\Longleftrightarrow \exists \alpha \in \Omega (x \prec_\alpha y).$$

We will define $\prec_\alpha$ by transfinite recursion. Note that the only non-trivial case will be the successor case. Recall that we have defined $\mathrm{L}_{\alpha+1}$ to be $\mathrm{Def}(\mathrm{L}_\alpha)$, and each element of $\mathrm{Def}(\mathrm{L}_\alpha)$ is of the form $x = D(\alpha, f, \vec{p})$, where $f \in \mathrm{Fml}$, $\vec{p} \in \mathrm{seq}(\mathrm{L}_\alpha)$ and

$$D(\alpha, f, \vec{p}) := \left\{ y \in \mathrm{L}_\alpha : \mathrm{L}_\alpha \ulcorner \vDash \urcorner f[\langle y, \vec{p} \rangle] \right\}.$$

Therefore, the task of defining a well-ordering on $\mathrm{L}_{\alpha+1}$ essentially reduces to ordering triples $(\alpha, f, \vec{p})$, whereby one has to take into account that different triples can generate the same set. Thus, we will also define recursively a well-ordering $\tilde{\prec}_\alpha$ on triples of the form $(\beta, f, \vec{p})$ for $\beta < \alpha$, $f \in \mathrm{Fml}$, and $\vec{p} \in \mathrm{seq}(\mathrm{L}_\alpha)$. Now, since the sequence of parameters $\vec{p}$ is in L_α, one needs to refer to $\prec_\alpha$ in order to define $\tilde{\prec}_{\alpha+1}$. Hence, we define both well-orderings by a simultaneous recursion.

Moreover, observe that ordering such triples further requires ordering Fml. By construction, we have $\mathrm{Fml} \subseteq \mathrm{V}_\omega$, and thus, Fml is countable. Hence, there is a well-ordering $\prec_{\mathrm{Fml}}$ of Fml. We proceed as follows:

- Let $\prec_0$ be the empty ordering, i.e., $\prec_0 := \emptyset$.
- Suppose that for some $\alpha \in \Omega$, $\prec_\alpha$ has already been defined. We first tackle $\tilde{\prec}_{\alpha+1}$. Let $\beta, \gamma < \alpha$, $f, g \in \mathrm{Fml}$, and let $\vec{p}, \vec{q} \in \mathrm{seq}(\mathrm{L}_\alpha)$ of length $l_{\vec{p}}$ and $l_{\vec{q}}$, respectively. Then we define:

$$\begin{aligned}\langle \beta, f, \vec{p}\rangle \tilde{\prec}_{\alpha+1} \langle \gamma, g, \vec{q}\rangle :\Longleftrightarrow\ & \beta < \gamma \vee (\beta = \gamma \wedge f \prec_{\mathrm{Fml}} g)\\ & \vee \big(\beta = \gamma \wedge f = g \wedge l_{\vec{p}} < l_{\vec{q}}\big)\\ & \vee \Big(\beta = \gamma \wedge f = g \wedge l_{\vec{p}} = l_{\vec{q}} \wedge\\ & \quad \exists n \in \omega \big(n = \min\big\{m < l_{\vec{p}}, l_{\vec{q}} : \vec{p}(m) \neq \vec{q}(m)\big\}\\ & \quad \wedge \vec{p}(n) \prec_\alpha \vec{q}(n)\big)\Big)\end{aligned}$$

Now, we are ready to define $\prec_{\alpha+1}$. For $x, y \in \mathrm{L}_{\alpha+1}$, we set

$$x \prec_{\alpha+1} y :\Longleftrightarrow \langle \beta, f, p\rangle \tilde{\prec}_{\alpha+1} \langle \gamma, g, q\rangle,$$

where $\langle \beta, f, \vec{p}\rangle$ and $\langle \gamma, g, \vec{q}\rangle$ are $\tilde{\prec}_{\alpha+1}$-minimal triples such that $x = D(\beta, f, \vec{p})$ and $y = D(\gamma, g, \vec{q})$.

- If α is a limit ordinal, then we set

$$\prec_\alpha := \bigcup_{\beta<\alpha} \prec_\beta .$$

By construction, $\prec_\beta$ is an end-extension of $\prec_\alpha$ for all $\alpha, \beta \in \Omega$ with $\alpha < \beta$. Therefore, the ordering $\prec_{\mathbf{L}}$ as defined above is a well-ordering of the entire constructible universe **L**. Note that the existence of a well-ordering of the whole model, a so-called **global well-ordering**, yields a strengthening of the Axiom of Choice, namely a class-sized choice function which chooses an element from every non-empty set.

As a consequence of the existence of a global well-ordering of **L**, we obtain the following

THEOREM 14.10. *The constructible universe is a model of the Axiom of Choice, i.e.,*

$$\mathbf{L} \vDash \mathsf{ZFC}\,.$$

Proof. For every family $\mathscr{F} \in \mathrm{L}$ such that $\emptyset \notin \mathscr{F}$, there is a choice function

$$f : \mathscr{F} \to \bigcup \mathscr{F},$$

where $f(x)$ is the $\prec_{\mathbf{L}}$-minimal element of $x \in \mathscr{F}$. ⊣

In particular, as a consequence of THEOREM 14.10 we obtain that the consistency of ZF implies the consistency of ZFC, i.e.,

$$\mathrm{Con}(\mathsf{ZF}) \Longrightarrow \mathrm{Con}(\mathsf{ZFC}).$$

Notes

The constructible universe $\mathbf{L}$ was introduced by Gödel in his 1938 paper [17], in which he proved both that if ZFC is consistent, then $\mathbf{L}$ is a model of the Axiom of Choice and the Continuum Hypothesis. The construction of $\mathbf{L}$ presented here is mainly taken from Koepke [30] (see also Kunen [31]). According to Bernays, Gödel originally used the old German script $\mathfrak{L}$ to denote the constructible universe, where $\mathfrak{L}$ is a capital C and not — as one could think — a capital L. In 1963, Cohen developed in [4] and [5] the method of forcing (see, for example, Halbeisen [22] for an introduction) to prove that both the Axiom of Choice and the Continuum Hypothesis are in fact independent of the axioms of ZF.

Exercises

14.0 Prove Fact 14.0.

14.1 Let $\mathbf{V} \vDash \mathsf{ZFC}$. For cardinals $\kappa \in \mathbf{V}$, we define the class

$$H_\kappa := \big\{x \in \mathbf{V} : |\operatorname{TC}(\{x\})| < \kappa\big\}.$$

(a) Prove $H_\kappa \subseteq \mathrm{V}_\kappa$ and conclude that H_κ is a set.

Hint: Set $\operatorname{rk}(x) = \alpha$ for the minimal $\alpha \in \Omega$ such that $x \in \mathrm{V}_\alpha$ and prove that $\{\operatorname{rk}(y) : y \in \operatorname{TC}(x)\} = \operatorname{rk}(x)$. We call $\operatorname{rk}(x)$ the *rank* of x.

(b) Find cardinals κ_1 and κ_2 such that $H_{\kappa_1} = \mathrm{V}_{\kappa_2}$ and $H_{\kappa_2} \neq \mathrm{V}_{\kappa_2}$.

14.2 Examine which of the axioms of ZFC hold in the structure $(H_\kappa, \in)$ for $\kappa \in \{\aleph_0, \aleph_1\}$.

Remark: $H_{\aleph_0}$ is called the *set of hereditarily finite sets* and $H_{\aleph_1}$ is called the *set of hereditarily countable sets*.

14.3 Show that Zermelo's axiom system Z does not imply the Axiom Schema of Replacement.

Hint: Show that $\mathrm{V}_{\omega+\omega}$ is a model of Z but the Axiom Schema of Replacement fails in $\mathrm{V}_{\omega+\omega}$.

14.4 Show that the axiom system of ZF is not equivalent to a FINITE set of axioms.

Hint: Notice that Lévy's Reflection Theorem also holds if we replace L_α by V_α. Use this fact and the Axiom of Foundation to show that the assumption that ZF is equivalent to a FINITE axiom system leads to a contradiction.

14.5 A cardinal κ is called inaccessible if it has the following properties:

(0) κ is uncountable.
(1) For all cardinals $\lambda < \kappa$ we have $2^\lambda < \kappa$.
(2) For all sets $A \subseteq \kappa$ with $|A| < \kappa$ we have $\bigcup A \in \kappa$.

Recall that by Corollary 13.2.(a), $\bigcup A$ is an ordinal, and notice that $\aleph_0$ has properties (1) and (2).

(a) Show that if κ is inaccessible, then $\mathrm{V}_\kappa \vDash \mathsf{ZFC}$.

Hint: Show first that for all ordinals $\alpha \in \kappa$, $|\mathrm{V}_\alpha| < \kappa$.

(b) Show that if ZFC is consistent, then the existence of an inaccessible cardinal cannot be proved within ZFC.

Chapter 15
Models and Ultraproducts

The goal of this chapter is to show that every consistent $\mathscr{L}$-theory has a model, no matter whether the signature $\mathscr{L}$ is countable or uncountable. In addition, we will show that if a consistent $\mathscr{L}$-theory T has an infinite model, then, on the one hand, T has arbitrarily large models, and on the other hand, T has a model of size at most $\max\{\aleph_0, |\mathscr{L}|\}$.

In order to prove these results, we shall work within a model of ZFC, in particular, we shall make use of the Axiom of Choice. Therefore, in contrast to the proofs of the corresponding results in Part II, the following proofs are not constructive in general. As a matter of fact, we would like to mention that even though the proofs are carried out in a model of ZFC, in general, they cannot be carried out in ZFC. In fact, we do not work with ZFC as a formal system, but we just take a model of ZFC and use it as a framework in which we carry out the proofs.

Filters and Ultrafilters

Let S be an arbitrary non-empty set and let $\mathscr{P}(S)$ be the *power-set* of S, i.e., the set of all subsets of S. A set $\mathscr{F} \subseteq \mathscr{P}(S)$ is called a **filter** over S, if $\mathscr{F}$ has the following properties:

- $S \in \mathscr{F}$ and $\emptyset \notin \mathscr{F}$
- $(x \in \mathscr{F} \wedge y \in \mathscr{F}) \to (x \cap y) \in \mathscr{F}$
- $(x \in \mathscr{F} \vee y \in \mathscr{F}) \to (x \cup y) \in \mathscr{F}$

In particular, if $x \in \mathscr{F}$ and $x \subseteq y$, then $y \in \mathscr{F}$. Thus, a filter over S is a set of subsets of S which does not contain the empty set and which is closed under finite intersections and supersets. For example, the set $\{S\}$ is a filter over S.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_16

A more interesting example of a filter over S is the set

$$\mathscr{F} := \big\{x \subseteq S : S \setminus x \text{ is finite}\big\},$$

which is the so-called *Fréchet-filter*. Now, a set $\mathscr{U} \subseteq \mathscr{P}(S)$ is called an **ultrafilter** over S, if $\mathscr{U}$ is a filter over S and for each $x \in \mathscr{P}(S)$, either $x \in \mathscr{U}$ or $(S \setminus x) \in \mathscr{U}$. In other words, a filter $\mathscr{U}$ is an ultrafilter if $\mathscr{U}$ is not properly contained in any filter. For example, for each $a \in S$, the set

$$\mathscr{U}_a := \big\{x \subseteq S : a \in x\big\}$$

is an ultrafilter over S, called *trivial ultrafilter*. In particular, every ultrafilter over a finite set is trivial. It is natural to ask whether there exist also non-trivial ultrafiters, for example, ultrafilters which contain the Fréchet-filter. Or in general, we can ask whether every filter can be extended to an ultrafiter. This is what the *Ultrafilter Theorem* states:

> Ultrafilter Theorem: If $\mathscr{F}$ is a filter over a set S, then $\mathscr{F}$ can be extended to an ultrafilter.

Surprisingly, we cannot prove the Ultrafilter Theorem without assuming some form of the Axiom of Choice. However, proving the Ultrafilter Theorem within ZFC is not so hard (see Solution to Exercise 15.1).

Ultraproducts and Ultrapowers

Let $\mathscr{L}$ be an arbitrary but fixed signature, let I be a non-empty set, and for each $\iota \in I$, let $\mathbf{M}_\iota$ be an $\mathscr{L}$-structure with domain A_ι. Furthermore, let $A := \bigtimes_{\iota \in I} A_\iota$ be the Cartesian product of the sets A_ι. Below, we shall identify the elements of A with functions $f : I \to \bigcup_{\iota \in I} A_\iota$, where for each $\iota \in I$, $f(\iota) \in A_\iota$. Finally, let $\mathscr{U} \subseteq \mathscr{P}(I)$ be an ultrafilter over I. With respect to $\mathscr{U}$, we define a binary relation $\sim$ on A by stipulating

$$f \sim g \;:\Longleftrightarrow\; \big\{\iota \in I : f(\iota) = g(\iota)\big\} \in \mathscr{U}\,.$$

Fact 15.0. *The relation $\sim$ is an equivalence relation.*

Proof. We have to show that $\sim$ is reflexive, symmetric, and transitive.

- For all $f \in A$, we obviously have $f \sim f$.
- For all $f, g \in A$, we obviously have $f \sim g \leftrightarrow g \sim f$.
- Let $f, g, h \in A$ and assume that $f \sim g$ and $g \sim h$. Furthermore, let $x := \big\{\iota \in I : f(\iota) = g(\iota)\big\}$ and $y := \big\{\iota \in I : g(\iota) = h(\iota)\big\}$. Then $x, y \in \mathscr{U}$, and since $\mathscr{U}$ is a filter, $x \cap y$ as well as every superset of $x \cap y$ belongs to $\mathscr{U}$. Thus, we have

$$x \cap y \subseteq \big\{\iota \in I : f(\iota) = h(\iota)\big\} \in \mathscr{U},$$

which shows that $f \sim h$.

⊣

Now, for each $f \in A$, let

$$[f] := \{g \in A : g \sim f\}$$

and let

$$A^* := \big\{[f] : f \in A\big\}.$$

We now construct the $\mathscr{L}$-structure $\mathbf{M}^*$ with domain A^* as follows:

- For every constant symbol $c \in \mathscr{L}$, let $f_c \in A$ be defined by stipulating

$$f_c(\iota) := c^{\mathbf{M}_\iota} \quad \text{for all } \iota \in I,$$

and let

$$c^{\mathbf{M}^*} := [f_c].$$

- For every n-ary function symbol $F \in \mathscr{L}$, let $F^{\mathbf{M}^*} : (A^*)^n \to A^*$ be such that

$$F^{\mathbf{M}^*}\big([f_0], \dots, [f_{n-1}]\big) = [f] \iff \Big\{\iota \in I : F^{\mathbf{M}_\iota}\big(f_0(\iota), \dots, f_{n-1}(\iota)\big) = f(\iota)\Big\} \in \mathscr{U}.$$

- For every n-ary relation symbol $R \in \mathscr{L}$, let $R^{\mathbf{M}^*} \subseteq (A^*)^n$ be such that

$$\big\langle [f_0], \dots, [f_{n-1}] \big\rangle \in R^{\mathbf{M}^*} \iff \Big\{\iota \in I : \big\langle f_0(\iota), \dots, f_{n-1}(\iota)\big\rangle \in R^{\mathbf{M}_\iota}\Big\} \in \mathscr{U}.$$

FACT 15.1. *The constants $c^{\mathbf{M}^*}$, the functions $F^{\mathbf{M}^*}$, and the relations $R^{\mathbf{M}^*}$ are well-defined.*

Proof. We just show that the functions $F^{\mathbf{M}^*} : (A^*)^n \to A$ are well-defined and leave the proofs for $c^{\mathbf{M}^*}$ and $R^{\mathbf{M}^*}$ as an exercise (see EXERCISE 15.2). Let $F \in \mathscr{L}$ be an n-ary function symbol and let $\langle f_0, \dots, f_{n-1}\rangle$ and $\langle g_0, \dots, g_{n-1}\rangle$ be elements in A^n such that for each $0 \leq i < n$ we have

$$f_i \sim g_i \quad \text{or equivalently} \quad [f_i] = [g_i]\,.$$

For $0 \leq i < n$ let

$$x_i := \big\{\iota \in I : f_i(\iota) = g_i(\iota)\big\}\,.$$

Then $x_i \in \mathscr{U}$ for each $0 \leq i < n$, and since $\mathscr{U}$ is a filter, we get that also $x_0 \cap \cdots \cap x_{n-1} \in \mathscr{U}$. Furthermore, we define $f, g \in A$ by stipulating

$$f(\iota) := F^{\mathbf{M}_\iota}\big(f_0(\iota), \ldots, f_{n-1}(\iota)\big) \quad \text{and} \quad g(\iota) := F^{\mathbf{M}_\iota}\big(g_0(\iota), \ldots, g_{n-1}(\iota)\big)\,.$$

Then we have

$$x_0 \cap \cdots \cap x_{n-1} \subseteq \big\{\iota \in I : f_0(\iota) = g_0(\iota) \wedge \cdots \wedge f_{n-1}(\iota) = g_{n-1}(\iota)\big\} \in \mathscr{U}\,,$$

and consequently, we obtain

$$\Big\{\iota \in I : F^{\mathbf{M}_\iota}\big(f_0(\iota), \ldots, f_{n-1}(\iota)\big) = F^{\mathbf{M}_\iota}\big(g_0(\iota), \ldots, g_{n-1}(\iota)\big)\Big\} \in \mathscr{U}\,.$$

Hence, $\big\{\iota \in I : f(\iota) = g(\iota)\big\} \in \mathscr{U}$, which shows that $[f] = [g]$ and implies that

$$F^{\mathbf{M}^*}\big([f_0], \ldots, [f_{n-1}]\big) \;=\; F^{\mathbf{M}^*}\big([g_0], \ldots, [g_{n-1}]\big)\,.$$

Therefore, the value of the function $F^{\mathbf{M}^*}$ does not depend on the particular representatives that we choose from the equivalence classes $[f_i]$. ⊣

The $\mathscr{L}$-structure $\mathbf{M}^*$ with domain A^* is called the **ultraproduct** of the $\mathscr{L}$-structures $\mathbf{M}_\iota$ ($\iota \in I$) with respect to the ultrafilter $\mathscr{U}$ over I. If for all $\iota \in I$ we have $\mathbf{M}_\iota = \mathbf{M}$ for some $\mathscr{L}$-structure $\mathbf{M}$, then $\mathbf{M}^*$ is called the **ultrapower** of $\mathbf{M}$ with respect to $\mathscr{U}$.

In the next section, we show that if each $\mathscr{L}$-structure $\mathbf{M}_\iota$ is a model of some $\mathscr{L}$-theory T, then also the ultraproduct $\mathbf{M}^*$ is a model of T.

Łoś's Theorem

As above, let $\mathscr{L}$ be an arbitrary signature, let I be a non-empty set, and for each $\iota \in I$, let $\mathbf{M}_\iota$ be an $\mathscr{L}$-structure with domain A_ι. Finally, let $\mathscr{U}$ be an ultrafilter over I and let $\mathbf{M}^*$ be the ultraproduct of the $\mathscr{L}$-structures $\mathbf{M}_\iota$ ($\iota \in I$) with respect to $\mathscr{U}$. The following result allows us to decide whether a given $\mathscr{L}$-sentence is valid in $\mathbf{M}^*$.

Theorem 15.2 (Łoś's Theorem). *For each $\mathscr{L}$-sentence σ, we have*

$$\mathbf{M}^* \vDash \sigma \quad \Longleftrightarrow \quad \big\{\iota \in I : \mathbf{M}_\iota \vDash \sigma\big\} \in \mathscr{U}\,.$$

Proof. By Theorem 1.7, for every $\mathscr{L}$-sentence σ there is an equivalent $\mathscr{L}$-sentence σ' which contains only $\neg$ and $\wedge$ as logical operators and $\exists$ as quantifier. Therefore, it is enough to prove Łoś's Theorem for the $\mathscr{L}$-sentence σ'.

The proof is by induction on the number of the symbols $\neg$, $\wedge$, and $\exists$ which appear in the $\mathscr{L}$-sentence σ'.

By construction of $\mathbf{M}^*$, Łoś's Theorem holds for atomic $\mathscr{L}$-sentences σ', i.e., for sentences σ' which are build with the rules (F0) and (F1).

Assume that $\sigma' \equiv \neg\sigma_0$ and that Łoś's Theorem holds for σ_0. Then we have:

$$\begin{aligned}
\mathbf{M}^* \vDash \neg\sigma_0 \quad &\Longleftrightarrow \quad \mathbf{M}^* \nvDash \sigma_0 \\
&\Longleftrightarrow \quad \big\{\iota \in I : \mathbf{M}_\iota \vDash \sigma_0\big\} \notin \mathscr{U} \\
&\Longleftrightarrow \quad I \setminus \big\{\iota \in I : \mathbf{M}_\iota \vDash \sigma_0\big\} \in \mathscr{U} \\
&\Longleftrightarrow \quad \big\{\iota \in I : \mathbf{M}_\iota \nvDash \sigma_0\big\} \in \mathscr{U} \\
&\Longleftrightarrow \quad \big\{\iota \in I : \mathbf{M}_\iota \vDash \neg\sigma_0\big\} \in \mathscr{U}
\end{aligned}$$

Now, assume that $\sigma' \equiv \sigma_1 \wedge \sigma_2$ and that Łoś's Theorem holds for σ_1 and σ_2. Then we have:

$$\begin{aligned}
\mathbf{M}^* \vDash \sigma_1 \wedge \sigma_2 \quad &\Longleftrightarrow \quad \mathbf{M}^* \vDash \sigma_1 \quad \text{and} \quad \mathbf{M}^* \vDash \sigma_2 \\
&\Longleftrightarrow \quad \underbrace{\big\{\iota \in I : \mathbf{M}_\iota \vDash \sigma_1\big\}}_{=:x_1} \in \mathscr{U} \ \text{and} \ \underbrace{\big\{\iota \in I : \mathbf{M}_\iota \vDash \sigma_2\big\}}_{=:x_2} \in \mathscr{U} \\
&\Longleftrightarrow \quad x_1 \cap x_2 \in \mathscr{U} \\
&\Longleftrightarrow \quad \big\{\iota \in I : \mathbf{M}_\iota \vDash \sigma_1 \wedge \sigma_2\big\} \in \mathscr{U}
\end{aligned}$$

Finally, assume that $\sigma' \equiv \exists\nu\sigma_0$ (for some variable ν) and that for some $[g] \in A^*$ we have

$$\mathbf{M}^*\tfrac{[g]}{\nu} \vDash \sigma_0(\nu) \ \Longleftrightarrow \big\{\iota \in I : \mathbf{M}_\iota\tfrac{g(\iota)}{\nu} \vDash \sigma_0(\nu)\big\} \in \mathscr{U}.$$

Then we have:

$$\begin{aligned}
\mathbf{M}^* \vDash \exists\nu\sigma_0 \quad &\Longleftrightarrow \quad \text{it exists } [g_0] \text{ in } A^* : \ \mathbf{M}^*\tfrac{[g_0]}{\nu} \vDash \sigma_0(\nu) \\
&\Longleftrightarrow \quad \text{it exists } [g_0] \text{ in } A^* : \underbrace{\big\{\iota \in I : \mathbf{M}_\iota\tfrac{g_0(\iota)}{\nu} \vDash \sigma_0(\nu)\big\}}_{=:x} \in \mathscr{U}
\end{aligned}$$

Because $x \subseteq \big\{\iota \in I : \mathbf{M}_\iota \vDash \exists\nu\sigma_0\big\}$, it follows that $\big\{\iota \in I : \mathbf{M}_\iota \vDash \exists\nu\sigma_0\big\} \in \mathscr{U}$, which shows that

$$\mathbf{M}^* \vDash \exists\nu\sigma_0 \quad \Longrightarrow \quad \big\{\iota \in I : \mathbf{M}_\iota \vDash \exists\nu\sigma_0\big\} \in \mathscr{U}.$$

In order to show the converse implication, we have to make use of the Axiom of Choice. If, for $\iota \in I$, $\mathbf{M}_\iota \vDash \exists\nu\sigma_0$, then let $a_\iota \in A_\iota$ be such that $\mathbf{M}_\iota\frac{a_\iota}{\nu} \vDash \sigma_0(\nu)$, otherwise, let a_ι be an arbitrary element of A_ι. Now, for the function

$$\begin{array}{rcl} g_0 : I & \to & \bigcup_{\iota \in I} A_\iota \\ \iota & \mapsto & a_\iota \end{array}$$

we have $\{\iota \in I : \mathbf{M}_\iota \vDash \exists \nu \sigma_0\} = \{\iota \in I : \mathbf{M}_\iota \frac{g_0(\iota)}{\nu} \vDash \sigma_0(\nu)\}$. In particular, if $\{\iota \in I : \mathbf{M}_\iota \vDash \exists \nu \sigma_0\} \in \mathscr{U}$, then also

$$\{\iota \in I : \mathbf{M}_\iota \frac{g_0(\iota)}{\nu} \vDash \sigma_0(\nu)\} \in \mathscr{U},$$

which shows that

$$\{\iota \in I : \mathbf{M}_\iota \vDash \exists \nu \sigma_0\} \in \mathscr{U} \quad \Longrightarrow \quad \mathbf{M}^* \vDash \exists \nu \sigma_0 .$$

Thus, we obtain

$$\mathbf{M}^* \vDash \exists \nu \sigma_0 \Longleftrightarrow \{\iota \in I : \mathbf{M}_\iota \vDash \exists \nu \sigma_0\} \in \mathscr{U} ,$$

which completes the proof. ⊣

The Completeness Theorem for Uncountable Signatures

In Chapter 5, we have proven Gödel's Completeness Theorem 5.5 (i.e., the Completeness Theorem for countable signatures). The proof given there was based on potentially infinite lists, and the metamathematical assumptions we made were very mild. In fact, our proof for Gödel's Completeness Theorem 5.5 can be carried out effectively in a kind of *algorithmic* way. In contrast to the proof for countable signatures, the proof of the Completeness Theorem for uncountable signatures — which will follow from the semantic form of the Compactness Theorem 2.17 — is much more formal. In particular, it makes use of Łoś's Theorem 15.2, which is based on the existence of ultrafilters and choice functions, and is carried out in a model of ZFC — but not in ZFC itself.

Theorem 15.3 (Semantic Form of the Compactness Theorem). *Let* T *be an* $\mathscr{L}$*-theory such that for every finite subset* $\Phi \subseteq \mathsf{T}$ *there is an* $\mathscr{L}$*-structure* $\mathbf{M}_\Phi$ *such that* $\mathbf{M}_\Phi \vDash \Phi$. *Then* T *has a model.*

Proof. Let I be the set of all finite subsets of T, i.e.,

$$I := \{\Phi \subseteq \mathsf{T} : \Phi \text{ is finite}\}.$$

For each $\Phi \in I$, let $\mathbf{M}_\Phi$ be an $\mathscr{L}$-structure with domain A_Φ such that $\mathbf{M}_\Phi \vDash \Phi$. Furthermore, for every $\Phi \in I$ let

$$\Delta(\Phi) := \{\Phi' \in I : \Phi \subseteq \Phi'\}.$$

In other words, $\Delta(\Phi)$ is the set of all finite supersets $\Phi' \supseteq \Phi$. In particular, for every $\Phi \in I$ we have $\Phi \in \Delta(\Phi)$ and $\Delta(\Phi) \subseteq I$. Now, for all $\Phi_1, \Phi_2 \in I$ we have $\Delta(\Phi_1) \cap \Delta(\Phi_2) = \Delta(\Phi_1 \cup \Phi_2)$, where $\Phi_1 \cup \Phi_2 \in I$. Therefore, the set

$$\mathscr{F} := \big\{\Psi \subseteq I : \exists \Phi \in I\big(\Delta(\Phi) \subseteq \Psi\big)\big\}$$

is a filter over I, which, by the Ultrafilter Theorem, can be extended to an ultrafilter $\mathscr{U}$.

Let $\mathbf{M}^*$ with domain A^* be the ultraproduct of the $\mathscr{L}$-structures $\mathbf{M}_\Phi$ ($\Phi \in I$) with respect to the ultrafilter $\mathscr{U}$ over I, and let $\sigma_0 \in \mathsf{T}$ be an arbitrary $\mathscr{L}$-sentence. Then $\{\sigma_0\} \in I$ and $\mathbf{M}_{\{\sigma_0\}} \vDash \sigma_0$. Moreover, for every $\Phi \in \Delta\big(\{\sigma_0\}\big)$ we have $\mathbf{M}_\Phi \vDash \sigma_0$. Therefore, we have

$$\Delta\big(\{\sigma_0\}\big) = \big\{\Phi \in I : \sigma_0 \in \Phi\big\} \subseteq \big\{\Phi \in I : \mathbf{M}_\Phi \vDash \sigma_0\big\}.$$

Now, since $\Delta\big(\{\sigma_0\}\big) \in \mathscr{F} \subseteq \mathscr{U}$, by Łoś's Theorem 15.2 we obtain

$$\mathbf{M}^* \vDash \sigma_0,$$

and since $\sigma_0 \in \mathsf{T}$ was arbitrary, this shows that $\mathbf{M}^* \vDash \mathsf{T}$. Hence, T has a model. ⊣

As a consequence of Theorem 15.3 and Gödel's Completeness Theorem 5.5, we obtain the Completeness Theorem for arbitrarily large signatures.

Theorem 15.4 (Completeness Theorem). *If $\mathscr{L}$ is an arbitrary signature and T is a consistent set of $\mathscr{L}$-sentences, then T has a model.*

Proof. Firstly, if T is consistent, then, by the Compactness Theorem 2.17, every finite subset $\Phi \subseteq \mathsf{T}$ is consistent. Secondly, as in the proof of Gödel's Completeness Theorem 5.5, for every finite subset of $\Phi \subseteq \mathsf{T}$ we can construct an $\mathscr{L}'$-structure $\mathbf{M}'_\Phi$ with domain A_Φ, such that $\mathbf{M}'_\Phi \vDash \Phi$, where $\mathscr{L}'$ is the finite subset of $\mathscr{L}$ consisting of all non-logical symbols which appear in sentences of Φ. Now, we extend each $\mathscr{L}'$-structure $\mathbf{M}'_\Phi$ to an $\mathscr{L}$-structure $\mathbf{M}_\Phi$ with the same domain A_Φ such that $\mathbf{M}_\Phi \vDash \Phi$ (see Exercise 3.2). Hence, for every finite subset of $\Phi \subseteq \mathsf{T}$ there is an $\mathscr{L}$-structure $\mathbf{M}_\Phi$ such that $\mathbf{M}_\Phi \vDash \Phi$, and therefore, we can apply Theorem 15.3 in order to construct a model $\mathbf{M}^* \vDash \mathsf{T}$. ⊣

As an immediate consequence of the Completeness Theorem 15.4 and the Soundness Theorem 3.7, we obtain the following

Corollary 15.5. *For any signature $\mathscr{L}$, a set T of $\mathscr{L}$-sentences has a model if and only if T is consistent.*

The Upward Löwenheim-Skolem Theorem

Next, we will show that every $\mathscr{L}$-theory which has an infinite model has arbitrarily large models.

THEOREM 15.6 (UPWARD LÖWENHEIM-SKOLEM THEOREM). *Let T be an $\mathscr{L}$-theory which has an infinite model, and let κ be an arbitrarily large cardinal. Then there exists a model $\mathbf{M}^* \models \mathsf{T}$ with domain A^* such that $|A^*| \geq \kappa$ (i.e., the cardinality of A^* is at least κ).*

Proof. For each $\gamma \in \kappa$, we define a constant symbol c_γ which does not belong to $\mathscr{L}$. Let $\mathscr{L}^* := \mathscr{L} \cup \{c_\gamma : \gamma \in \kappa\}$. Furthermore, let T^* be the $\mathscr{L}^*$-theory consisting of the sentences in T together with the sentences $c_\gamma \neq c_{\gamma'}$ (for any distinct $\gamma, \gamma' \in \kappa$). As in the proof of THEOREM 15.3, let I be the set of all finite subsets of T^*. Now, let $\mathbf{M} \models \mathsf{T}$ be a model with infinite domain A. For any $\Phi \in I$, we can extend the $\mathscr{L}$-structure $\mathbf{M}$ to an $\mathscr{L}^*$-structure $\mathbf{M}_\Phi$ such that

$$\mathbf{M}_\Phi \models \mathsf{T} + \Phi.$$

In order to see this, notice that the domain A of $\mathbf{M}$ is infinite and that there are just finitely many constant symbols c_γ which appear in Φ. Therefore, we can apply THEOREM 15.3 in order to construct an $\mathscr{L}^*$-structure $\mathbf{M}^*$ with domain A^* such that $\mathbf{M}^* \models \mathsf{T}^*$. Finally, by definition of T^*, the elements $c_\gamma^{\mathbf{M}^*}$ in A^* (for $\gamma \in \kappa$) are pairwise distinct, which shows that $|A^*| \geq \kappa$. ⊣

As an immediate consequence of the UPWARD LÖWENHEIM–SKOLEM THEOREM 15.6, we get the following

COROLLARY 15.7. *If an $\mathscr{L}$-theory T has a countably infinite model, then T also has an uncountable model. In particular, PA has an uncountable model.*

As a matter of fact, we would like to mention that the proof of the UPWARD LÖWENHEIM–SKOLEM THEOREM 15.6 can be carried out neither in the formal language of ZFC (since we use an infinite set of constant symbols), nor in the language of metamathematics (since we use THEOREM 15.3, which is based on ŁOŚ'S THEOREM 15.2 and therefore on ultrafilters).

The Downward Löwenheim-Skolem Theorem

The last result of this chapter provides an upper bound for the minimum size of a model of a given theory.

THEOREM 15.8 (DOWNWARD LÖWENHEIM-SKOLEM THEOREM). *If a consistent $\mathscr{L}$-theory T has an infinite model, then T has a model of size at most* $\max\{\aleph_0, |\mathscr{L}|\}$.

Proof. If the signature $\mathscr{L}$ is countable, then, by Gödel's Completeness Theorem 5.5, T has a model, which is — by construction — a countable model. Now, assume that $|\mathscr{L}|$ (i.e., the cardinality of $\mathscr{L}$) is uncountable. First notice that with the signature $\mathscr{L}$ we can build at most $|\mathscr{L}|$ terms. In order to see this, recall that a term is just a special finite string of logical and non-logical symbols, and by Fact 13.11, the cardinality of the set of such strings is $\max\{\aleph_0, |\mathscr{L}|\}$. Now, in order to build a model $\mathbf{M} \models \mathsf{T}$ of cardinality at most $\max\{\aleph_0, |\mathscr{L}|\}$, we can essentially follow the proof of Gödel's Completeness Theorem 5.5. However, instead of potentially infinite lists we have to work with actual infinite sequences of length at most $|\mathscr{L}|$. At the end of the construction, the domain of $\mathbf{M}$ will be a sequence of length at most $|\mathscr{L}|$ of sequences of length at most $|\mathscr{L}|$. ⊣

As an immediate consequence of the Downward Löwenheim–Skolem Theorem 15.8, we get the following

Corollary 15.9. *If T is a consistent $\mathscr{L}$-theory and the signature $\mathscr{L}$ is countable, then T has a countable model.*

As a matter of fact, we would like to mention that the proof of the Downward Löwenheim–Skolem Theorem 15.8 cannot be carried out in the formal language of ZFC either. Otherwise, since the signature of ZFC just contains the single symbol $\in$ and is therefore countable, we would be able to construct a countable model of ZFC within a model of ZFC. In particular, we would be able to prove within ZFC that ZFC is consistent, which obviously contradicts the Second Incompleteness Theorem.

Notes

Most of the material of this chapter is taken from Bell and Slomson [3, Ch. 5], where one can find some more historical background. Łoś's Theorem, also called the Fundamental Theorem of Ultraproducts, is due to the Polish mathematician Łoś (see, e.g., [33]). A first version of the Löwenheim-Skolem Theorems was proved by Löwenheim in 1915 (see [34]). Some years later, Skolem generalised Löwenheim's result in [49].

Exercises

15.0 Let S be a non-empty set.

(a) Show that a set $F \subseteq \mathscr{P}(S)$ can be extended to a filter $\mathscr{F}$ over S, if and only if no finite intersection of elements of F is empty, i.e., for every finite subset $\{x_0, \ldots, x_n\} \subseteq F$ of elements of F we have $\bigcap_{i=0}^{n} x_i \neq \emptyset$.

(b) Show that a set $\mathscr{U} \subseteq \mathscr{P}(S)$ is an ultrafilter if and only if every intersection of finitely many elements of $\mathscr{U}$ is non-empty and for all $x \subseteq S$ we have either $x \in \mathscr{U}$ or $S \setminus x \in \mathscr{U}$.

15.1 Find a proof of the Ultrafilter Theorem within ZFC.

Hint: First take a well-ordering of $\mathscr{P}(S)$, and then extend the filter $\mathscr{F}$ over S to an ultrafilter by transfinite induction.

15.2 Complete the proof of Fact 15.1, i.e., show that the constants $c^{\mathbf{M}^*}$ and the relations $R^{\mathbf{M}^*}$, defined in the construction of the $\mathscr{L}$-structure $\mathbf{M}^*$, are well-defined.

15.3 Prove Łoś's Theorem 15.2 for $\mathscr{L}$-sentences σ' which contain only $\neg$ and $\vee$ as logical operators and $\forall$ as quantifier.

15.4 Prove that there exists an ultraproduct of finite sets which is infinite.

Chapter 16
Models of Peano Arithmetic

The Standard Model of Peano Arithmetic in ZF

In this section, we will show that ZF is sufficiently strong to prove that PA is consistent. In fact, within a model $\mathbf{V}$ of ZF we can construct a model $\mathbb{N}_\omega$ of PA with domain ω. The model $\mathbb{N}_\omega$ which we obtain in $\mathbf{V}$ is the *standard model of* PA *with respect to* $\mathbf{V}$. In the case when the model $\mathbf{V}$ is a standard model of ZF, the model $\mathbb{N}_\omega$ is isomorphic to the standard model $\mathbb{N}$ of PA which we constructed in Chapter 7. However, if the model $\mathbf{V}$ is a non-standard model of ZF, then $\mathbb{N}_\omega$ is a non-standard model of PA (i.e., $\mathbb{N}_\omega$ is a model of PA which is not isomorphic to $\mathbb{N}$), and there is no way to obtain the standard model of PA within $\mathbf{V}$. In general, people living in $\mathbf{V}$, no matter whether $\mathbf{V}$ is a standard or a non-standard model of ZF, believe that $\mathbb{N}_\omega$ is the standard model $\mathbb{N}$.

Now, let $\mathbf{V}$ be a model of ZF. Within $\mathbf{V}$, we construct an $\mathscr{L}_{\mathsf{PA}}$-structure $\mathbb{N}_\omega$ with domain ω, and show that $\mathbb{N}_\omega$ is a model of PA. Recall that $\mathscr{L}_{\mathsf{PA}} = \{\mathsf{0}, \mathsf{s}, +, \cdot\}$. The $\mathscr{L}_{\mathsf{PA}}$-structure is defined by the following assignments which are based on ordinal arithmetic (see Exercise 13.5):

$$
\begin{aligned}
\mathsf{0}^{\mathbb{N}_\omega} &:= \emptyset \\
\mathsf{s}^{\mathbb{N}_\omega} &: \omega \to \omega \\
&\quad\; n \mapsto n+1 \\
+^{\mathbb{N}_\omega} &: \omega \times \omega \to \omega \\
&\quad\; \langle n, m\rangle \mapsto n+m \\
\cdot^{\mathbb{N}_\omega} &: \omega \times \omega \to \omega \\
&\quad\; \langle n, m\rangle \mapsto n \cdot m
\end{aligned}
$$

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_17

Before we show that the $\mathscr{L}_{\mathsf{PA}}$-structure $\mathbb{N}_\omega$ is a model of Peano Arithmetic, we first recall the axioms of PA:

PA_0: $\neg\exists x(\mathsf{s}x = \mathsf{0})$

PA_1: $\forall x\forall y(\mathsf{s}x = \mathsf{s}y \to x = y)$

PA_2: $\forall x(x + \mathsf{0} = x)$

PA_3: $\forall x\forall y(x + \mathsf{s}y = \mathsf{s}(x + y))$

PA_4: $\forall x(x \cdot \mathsf{0} = \mathsf{0})$

PA_5: $\forall x\forall y(x \cdot \mathsf{s}y = (x \cdot y) + x)$

If φ is any $\mathscr{L}_{\mathsf{PA}}$-formula with $x \in \text{free}(\varphi)$, then:

PA_6: $\big(\varphi(\mathsf{0}) \wedge \forall x(\varphi(x) \to \varphi(\mathsf{s}(x)))\big) \to \forall x\varphi(x)$

Let us now show that $\mathbb{N}_\omega \vDash \mathsf{PA}$:

- PA_0: Since $n + 1 = n \cup \{n\}$ and $n \in \{n\}$ (i.e., $n \cup \{n\} \neq \emptyset$), there is no $n \in \omega$ such that $n + 1 = \emptyset$.
- PA_1: If $n, m \in \omega$ and $n \neq m$, then, by Theorem 13.1.(c), we have either $n \in m$ or $m \in n$, and in both cases we get $n + 1 \neq m + 1$.
- PA_2 and PA_3: Follow immediately from (a) and (b) of ordinal addition.
- PA_4 and PA_5: Follow immediately from (a) and (b) of ordinal multiplication.
- PA_6: Let φ be an $\mathscr{L}_{\mathsf{PA}}$-formula with $x \in \text{free}(\varphi)$ and assume that

$$\varphi(\emptyset) \wedge \forall n \in \omega\big(\varphi(n) \to \varphi(n + 1)\big)\,.$$

  Furthermore, let $E := \{n \in \omega : \neg\varphi(n)\}$. Obviously, E is a subset of ω. If $E = \emptyset$, then $\forall n \in \omega\big(\varphi(n)\big)$ and we are done. Otherwise, if $E \neq \emptyset$, let m be the $\in$-minimal element of E. Now, m can *neither* be $\emptyset$, since we assumed $\varphi(\emptyset)$, *nor* a successor ordinal (i.e., of the form $n + 1$), since we assumed $\varphi(n) \to \varphi(n + 1)$ which is equivalent to $\neg\varphi(n + 1) \to \neg\varphi(n)$. Thus, there is no $\in$-minimal element of E, which is only possible when $E = \emptyset$.

Thus, $\mathbb{N}_\omega$ is a model of PA with domain ω.

In Chapter 7, we saw that there are non-standard models of PA. However, the existence of these models was obtained by the Compactness Theorem 2.17, and the proof cannot be carried out in ZFC. In the next section, we will now give a construction of non-standard models of PA which can be carried out in ZFC. Since the construction uses ultrapowers, it cannot be carried out without the aid of the Axiom of Choice.

A Non-Standard Model of Peano Arithmetic in ZFC

The non-standard model of PA which we now construct is the ultrapower of the standard model $\mathbb{N}_\omega$ with respect to some arbitrary but fixed non-trivial ultrafilter $\mathscr{U}$ over ω. First, let ${}^\omega\omega$ be the set of all functions $f : \omega \to \omega$. With respect to $\mathscr{U}$, we define the binary relation $\sim$ on ${}^\omega\omega$ by stipulating

$$f \sim g \;:\Longleftrightarrow\; \big\{n \in \omega : f(n) = g(n)\big\} \in \mathscr{U}\,.$$

Then the relation $\sim$ is an equivalence relation (see Chapter 15). For each $f \in {}^\omega\omega$, let

$$[f] := \{g \in {}^\omega\omega : g \sim f\},$$

and let

$$\omega^* := \big\{[f] : f \in {}^\omega\omega\big\}.$$

We now construct the $\mathscr{L}_{\mathsf{PA}}$-structure $\mathbb{N}^*_\omega$ with domain ω^* as follows:

- For the constant symbol $\mathsf{0} \in \mathscr{L}_{\mathsf{PA}}$, let $f_{\mathsf{0}} \in {}^\omega\omega$ be defined by stipulating

  $$f_{\mathsf{0}}(n) := 0 \quad \text{for all } n \in \omega,$$

  and let

  $$\mathsf{0}^{\mathbb{N}^*_\omega} := [f_{\mathsf{0}}].$$

- For the unary function symbol $\mathtt{s}$ in $\mathscr{L}_{\mathsf{PA}}$, we define $\mathtt{s}(f)$ by stipulating

  $$\mathtt{s}(f)(n) := f(n) + 1 \quad \text{for } n \in \omega,$$

  and let

  $$\mathtt{s}^{\mathbb{N}^*_\omega}\big([f]\big) := [\mathtt{s}(f)].$$

- For the binary function symbols $+$ and $\cdot$ in $\mathscr{L}_{\mathsf{PA}}$, we define $f + g$ and $f \cdot g$ (for $f, g \in {}^\omega\omega$) by stipulating for all $n \in \omega$

  $$(f + g)(n) := f(n) +^{\mathbb{N}_\omega} g(n)\,,$$
  $$(f \cdot g)(n) := f(n) \cdot^{\mathbb{N}_\omega} g(n)\,,$$

  and let

  $$[f] +^{\mathbb{N}^*_\omega} [g] := [f + g] \quad \text{and} \quad [f] \cdot^{\mathbb{N}^*_\omega} [g] := [f \cdot g].$$

By Łoś's Theorem 15.2, the $\mathscr{L}_{\mathsf{PA}}$-structure $\mathbb{N}^*_\omega$ is a model of PA. In order to see that $\mathbb{N}^*_\omega$ is a non-standard model of PA, first notice that $\mathbb{N}_\omega$ can be embedded — in a unique way — into $\mathbb{N}^*_\omega$ by the embedding

$$\begin{aligned} \omega &\to \omega^* \\ k &\mapsto [f_k]\,, \end{aligned}$$

where $f_k(n) := k$ for all $n \in \omega$. This shows that $\mathbb{N}_\omega$ is a substructure of $\mathbb{N}^*_\omega$.

In order to show that the models $\mathbb{N}_\omega$ and $\mathbb{N}^*_\omega$ are not isomorphic, let $g \in {}^\omega\omega$ be such that for all $n \in \omega$,

$$g(n) := n\,.$$

Then for all $k \in \omega$, $[f_k] < [g]$, which shows that the models $\mathbb{N}_\omega$ and $\mathbb{N}^*_\omega$ are not isomorphic, even though they are elementarily equivalent (see Exercise 16.1).

Exercises

16.0 Describe the model $\mathbb{N}^*_\omega$ that one obtains by choosing a principal ultrafilter rather than a non-principal one.

16.1 Show that the $\mathscr{L}_{\mathsf{PA}}$-structures $\mathbb{N}_\omega$ and $\mathbb{N}^*_\omega$ are elementarily equivalent.
Hint: Use Łoś's Theorem 15.2.

16.2 Show that the domain ω^* of $\mathbb{N}^*_\omega$ is uncountable. In particular, show that $\mathbb{N}^*_\omega$ is an uncountable model of PA.

16.3 Show that for any $[g], [g'] \in \omega^*$ with $[g] < [g']$, the cardinality of the set

$$\big\{[f] \in \omega^* : [g] \leq [f] \leq [g']\big\}$$

is either finite or uncountable.

Chapter 17
Models of the Real Numbers

In this chapter, we will first construct a model of the real numbers using Cauchy sequences of rational numbers. We also present a second model of the real numbers according to A'Campo [1]. This construction has the advantage that it only relies on the integers and not on the rational numbers, and that the definition of the multiplication is much simpler and natural than the classical definition based on equivalence classes of Cauchy sequences. Afterwards, we will show that both constructions yield isomorphic models. The constructions of the real numbers will be quite general, such that — depending on whether we start with the standard or a non-standard model of the natural numbers — we obtain the standard or a non-standard model of the real numbers.

We shall conclude this chapter by giving a brief introduction to the so-called Non-Standard Analysis, which is essentially just Analysis in a non-standard model of the reals.

Let us first introduce the axioms R of the real numbers. The language of R is $\mathscr{L}_{\mathsf{R}} = \{\mathtt{0}, \mathtt{1}, +, \cdot, <\}$, where $\mathtt{0}$ and $\mathtt{1}$ are constant symbols, $+$ and $\cdot$" are binary function symbols and $<$ is a binary relation symbol.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3_18

The first group of axioms are simply the field axioms:

R_0: $\forall x \forall y \forall z \big(x + (y + z) = (x + y) + z\big)$
R_1: $\forall x (x + \mathsf{0} = x)$
R_2: $\forall x \exists y (x + y = 0)$
R_3: $\forall x \forall y (x + y = y + x)$
R_4: $\forall x \forall y \forall z \big(x \cdot (y \cdot z) = (x \cdot y) \cdot z\big)$
R_5: $\forall x (x \cdot \mathsf{1} = x)$
R_6: $\forall x \big(x \neq \mathsf{0} \to \exists y (x \cdot y = \mathsf{1})\big)$
R_7: $\forall x \forall y (x \cdot y = y \cdot x)$
R_8: $\forall x \forall y \forall z \big(x \cdot (y + z) = (x \cdot y) + (x \cdot z)\big)$
R_9: $\mathsf{0} \neq \mathsf{1}$

The second group of axioms are the so-called order axioms:

R_{10}: $\forall x \neg (x < x)$
R_{11}: $\forall x \forall y \forall z (x < y \wedge y < z \to x < z)$
R_{12}: $\forall x \forall y (x < y \vee x = y \vee y < x)$
R_{13}: $\forall x \forall y \forall z (x < y \to x + z < y + z)$
R_{14}: $\forall x \forall y \forall z (x < y \wedge \mathsf{0} < z \to x \cdot z < y \cdot z)$

The last two axioms form the Completeness Axiom which is — in contrast to the other axioms — a so-called *second-order axiom* (i.e., a statement, not about the real numbers, but about *sets* of real numbers). The set $\mathcal{N}$ denotes either the standard or a non-standard model of the natural numbers.

R_{15}: Every Cauchy sequence of reals converges.
R_{16}: If $x > \mathsf{0}$ and $y > 0$ there exists $n \in \mathcal{N}$ such that $nx > y$.

Axiom R_{16} is also called the Archimedian Axiom.

Classical Construction of the Real Numbers

Let $\mathcal{N}$ be either ω or ω^*, where ω^* is the ultrapower of ω with respect to some non-trivial ultrafilter $\mathscr{U} \subseteq \mathscr{P}(\omega)$. In other words, $\mathcal{N}$ is either the domain of the model $\mathbb{N}_\omega$ or of $\mathbb{N}^*_\omega$. Recall that the former model is the standard model of PA within some model of ZF, whereas the latter model is a non-standard model of PA, constructed in a model of ZFC, which is elementarily equivalent to the corresponding model $\mathbb{N}_\omega$. Furthermore, let $\mathbf{N}$ be the structure $(\mathcal{N}, 0, +, \cdot)$, i.e., $\mathbf{N}$ is either $\mathbb{N}_\omega$ or $\mathbb{N}^*_\omega$.

From $\mathbf{N}$, we first construct a model $\mathbb{Z}_\mathcal{N}$ of the integers, then we construct a model $\mathbb{Q}_\mathcal{N}$ of the rationals, and finally we construct a model $\mathbb{R}^{\mathscr{C}}_\mathcal{N}$ of the reals using Cauchy sequences.

A Model of the Integers

On $\mathcal{N}$, we first define the binary function $\dot{-}$ by stipulating

$$x \dot{-} y = z \;:\Longleftrightarrow\; \exists u\big(y + u = x \;\wedge\; z = u\big) \vee \big(\neg\exists u(y + u = x) \;\wedge\; z = 0\big)\,.$$

Now, we define the set of integers $\mathbb{Z}_{\mathcal{N}}$ as a subset of $\mathcal{N} \times \mathcal{N}$ by stipulating

$$\mathbb{Z}_{\mathcal{N}} := \big\{\langle x, 0\rangle : x \in \mathcal{N}\big\} \cup \big\{\langle 0, y\rangle : y \in \mathcal{N}\big\}.$$

We identify the elements $x \in \mathcal{N}$ with integers of the form $\langle 0, y\rangle$.
On $\mathbb{Z}_{\mathcal{N}}$, we define the two binary functions $+$ and $\cdot$ as well as the unary function $-$ as follows:

$$\begin{aligned}
\langle x_0, y_0\rangle + \langle x_1, y_1\rangle = z &:\Longleftrightarrow z = \big\langle (x_0{+}x_1) \dot{-} (y_0{+}y_1),\, (y_0{+}y_1) \dot{-} (x_0{+}x_1)\big\rangle \\
\langle x_0, y_0\rangle \cdot \langle x_1, y_1\rangle = z &:\Longleftrightarrow z = \big\langle (x_0 \cdot y_1) + (y_0 \cdot x_1),\, (x_0 \cdot x_1) + (y_0 \cdot y_1)\big\rangle \\
-\langle x, y\rangle = z &:\Longleftrightarrow z = \langle y, x\rangle
\end{aligned}$$

In order to simplify the notation, we usually write $\langle x_0, y_0\rangle - \langle x_1, y_1\rangle$ instead of $\langle x_0, y_0\rangle + \big(-\langle x_1, y_1\rangle\big)$. Notice that $\langle x, y\rangle - \langle x, y\rangle = \langle 0, 0\rangle$.

We leave it as an exercise to the reader to check that the structure

$$\mathbb{Z}_{\mathcal{N}} := \big(\mathbb{Z}_{\mathcal{N}}, \langle 0, 0\rangle, \langle 0, 1\rangle, +, \cdot\big)$$

is a model of the ring of integers satisfying the axioms R_0–R_9 except R_6, where $\langle 0, 0\rangle$ and $\langle 0, 1\rangle$ are the neutral elements with respect to the binary operations $+$ and $\cdot$, respectively. Notice that in the case when $\mathcal{N}$ is equal to ω^*, then $\mathbb{Z}_{\mathcal{N}}$ is a non-standard model of the integers.

On $\mathbb{Z}_{\mathcal{N}}$, we define the binary relation $<$ and the unary function symbol $|\cdot|$ as follows:

$$\begin{aligned}
\langle x_0, y_0\rangle < \langle x_1, y_1\rangle \;&:\Longleftrightarrow\; y_0 + x_1 < y_1 + x_0 \\
\big|\langle x, y\rangle\big| = z \;&:\Longleftrightarrow\; \begin{cases} z = \langle x, y\rangle & \text{if } \langle x, y\rangle > \langle 0, 0\rangle \\ z = \langle y, x\rangle & \text{otherwise} \end{cases}
\end{aligned}$$

A Model of the Rational Numbers

Let $\mathcal{N}^+ := \mathcal{N} \setminus \{0\}$. On pairs $\langle x_0, y_0\rangle, \langle x_1, y_1\rangle \in \mathbb{Z}_{\mathcal{N}} \times \mathcal{N}^+$ we define an equivalence relation $\sim$ by stipulating

$$\langle x_0, y_0\rangle \sim \langle x_1, y_1\rangle :\Longleftrightarrow x_0 \cdot y_1 = x_1 \cdot y_0.$$

Now, we denote the equivalence classes by

$$\tfrac{x}{y} := \big[\langle x, y\rangle\big] = \big\{\langle x', y'\rangle \in \mathbb{Z}_{\mathcal{N}} \times \mathcal{N}^+ : \langle x', y'\rangle \sim \langle x, y\rangle\big\}$$

and call $\frac{x}{y}$ a **rational number**. Let $\mathbb{Q}_{\mathcal{N}}$ denote the set of all rational numbers, i.e.,

$$\mathbb{Q}_{\mathcal{N}} := \big\{\tfrac{x}{y} : x \in \mathbb{Z}_{\mathcal{N}}, y \in \mathcal{N}^+\big\}.$$

We can now introduce the two binary functions $+$ and $\cdot$ by

$$\tfrac{x_0}{y_0} + \tfrac{x_1}{y_1} := \tfrac{x_0 y_1 + y_0 x_1}{y_0 y_1}\,,$$
$$\tfrac{x_0}{y_0} \cdot \tfrac{x_1}{y_1} := \tfrac{x_0 \cdot x_1}{y_0 \cdot y_1}\,.$$

We leave it as an exercise for the reader to check that these functions are well-defined and that the structure $\mathbb{Q}_{\mathcal{N}} = (\mathbb{Q}_{\mathcal{N}}, \frac{0}{1}, \frac{1}{1}, +, \cdot)$ satisfies the field axioms R_0–R_9. As in the case of the integers, if $\mathcal{N}$ is ω^*, then $\mathbb{Q}_{\mathcal{N}}$ is a non-standard model of the rational numbers.

On $\mathbb{Q}_{\mathcal{N}}$, we define the binary relation $<$ and the unary function symbol $|\cdot|$ as follows:

$$\tfrac{x_0}{y_0} < \tfrac{x_1}{y_1} \quad :\Longleftrightarrow \quad x_0 \cdot y_1 < x_1 \cdot y_0$$
$$|\tfrac{x}{y}| = z \quad :\Longleftrightarrow \quad \begin{cases} z = \frac{x}{y} & \text{if } x \geq 0 \\ z = \frac{-x}{y} & \text{otherwise} \end{cases}$$

Again, it is easy to check that the order $<$ and the absolute value function are well-defined and satisfy the usual properties.

A Model of the Real Numbers using Cauchy Sequences

Let $\mathbb{Q}^+_{\mathcal{N}}$ denote the positive rational numbers, i.e., those $p \in \mathbb{Q}_{\mathcal{N}}$ that satisfy $p > 0$. We define a sequence (a_n) of rational numbers to be a **Cauchy sequence**, if for every $\varepsilon \in \mathbb{Q}^+_{\mathcal{N}}$ there is an $N \in \mathcal{N}$ such that for all $m, n \in \mathcal{N}$ with $m, n \geq N$, $|a_n - a_m| < \varepsilon$. We denote the set of all Cauchy sequences of rationals by $\mathscr{C}$. Two Cauchy sequences $(a_n), (b_n) \in \mathscr{C}$ are said to be **equivalent**, denoted $(a_n) \approx (b_n)$, if for each positive rational number $\varepsilon \in \mathbb{Q}^+_{\mathcal{N}}$ there is an $N \in \mathcal{N}$ such that for all $n \in \mathcal{N}$ with $n \geq N$, $|a_n - b_n| < \varepsilon$. In order to simplify the notation, we shall write $\lim_{n\to\infty}(a_n - b_n) = 0$. Notice that the meaning of $\lim_{n\to\infty}$ depends on whether $\mathcal{N} = \omega$ or $\mathcal{N} = \omega^*$.

It is obvious that the relation $\approx$ is reflexive and symmetric. Moreover, it is also transitive. To see this, let (a_n), (b_n), (c_n) be Cauchy sequences with $(a_n) \approx (b_n)$ and $(b_n) \approx (c_n)$. Then for each positive $\varepsilon \in \mathbb{Q}^+_{\mathcal{N}}$ there are

$N_1, N_2 \in \mathcal{N}$ such that for all $n_1, n_2 \in \mathcal{N}$ with $n_1 \geq N_1$ and $n_2 \geq N_2$ we have

$$|a_{n_1} - b_{n_1}| < \tfrac{\varepsilon}{2} \quad \text{and} \quad |b_{n_2} - c_{n_2}| < \tfrac{\varepsilon}{2}\,.$$

Consequently, for all $n \geq \max\{N_1, N_2\}$ we have

$$|a_n - c_n| \leq |a_n - b_n| + |b_n - c_n| < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon\,,$$

which shows that $(a_n) \approx (c_n)$. Thus, $\approx$ is an equivalence relation on $\mathscr{C}$ and the equivalence classes with respect to $\approx$ are given by

$$[(a_n)] := \big\{(b_n) \in \mathscr{C} : (b_n) \approx (a_n)\big\}.$$

Let $\mathbb{R}^{\mathscr{C}}_{\mathcal{N}}$ denote the set of all equivalence classes of rational Cauchy sequences, i.e.,

$$\mathbb{R}^{\mathscr{C}}_{\mathcal{N}} := \{[(a_n)] : (a_n) \in \mathscr{C}\}.$$

The elements of $\mathbb{R}^{\mathscr{C}}_{\mathcal{N}}$ are called **real numbers**.

In order to obtain a model of the real numbers, we need to define the functions addition $+$ and multiplication $\cdot$ on $\mathbb{R}^{\mathscr{C}}_{\mathcal{N}}$, including the neutral elements $0_{[\mathscr{C}]}$ and $1_{[\mathscr{C}]}$, respectively; then we have to define a linear ordering $<$ on $\mathbb{R}^{\mathscr{C}}_{\mathcal{N}}$, and finally, we check that the structure $\mathbb{R}^{\mathscr{C}}_{\mathcal{N}} = (\mathbb{R}^{\mathscr{C}}_{\mathcal{N}}, 0_{[\mathscr{C}]}, 1_{[\mathscr{C}]}, +, \cdot, <)$ thus obtained satisfies all axioms of the real numbers.

Addition and multiplication: For $r, s \in \mathbb{R}^{\mathscr{C}}_{\mathcal{N}}$, represented by $(a_n), (b_n) \in \mathscr{C}$ we define:

$$r + s := [(a_n + b_n)]$$
$$r \cdot s := [(a_n \cdot b_n)]$$

Lemma 17.0. *Addition and multiplication of reals are well-defined, i.e., if $r, s \in \mathbb{R}^{\mathscr{C}}_{\mathcal{N}}$ such that r is represented by $(a_n), (a'_n)$ and s is represented by $(b_n), (b'_n)$, then $(a_n + b_n)$ and $(a_n b_n)$ are again Cauchy sequences such that $(a_n + b_n) \approx (a'_n + b'_n)$ and $(a_n b_n) \approx (a'_n b'_n)$.*

Proof. In order to verify that $(a_n + b_n)$ is a Cauchy sequence, let $\varepsilon \in \mathbb{Q}^+_{\mathcal{N}}$. By assumption, there are $N_1, N_2 \in \mathcal{N}$ such that for all $n, m \geq N := \max\{N_1, N_2\}$ we have $|a_m - a_n| < \frac{\varepsilon}{2}$ and $|b_m - b_n| < \frac{\varepsilon}{2}$. Then it follows

$$|(a_n + b_n) - (a_m + b_m)| \leq |a_n - a_m| + |b_n - b_m| < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon$$

for all $n, m \geq N$.

Next, we prove that $(a_n b_n)$ is a Cauchy sequence. Since Cauchy sequences are bounded, there is a $C \in \mathcal{N}$ such that $|a_n|, |b_n| \leq C$ for all $n \in \mathcal{N}$. Now we can choose $M_1, M_2 \in \mathcal{N}$ such that for all $n, m \geq M := \max\{M_1, M_2\}$ we have $|a_n - a_m| < \frac{\varepsilon}{2C}$ and $|b_n - b_m| < \frac{\varepsilon}{2C}$ for all $n, m \geq M$. Consequently, we obtain:

$$\begin{aligned}|a_n b_n - a_m b_m| &= |a_n(b_n - b_m) + b_m(a_n - a_m)| \\ &\leq |a_n| \cdot |b_n - b_m| + |b_m| \cdot |a_n - a_m| \\ &\leq C \cdot \tfrac{\varepsilon}{2C} + C \cdot \tfrac{\varepsilon}{2C} \\ &= \varepsilon\end{aligned}$$

Hence, $(a_n b_n)$ is a Cauchy sequence. The second part uses similar arguments. ⊣

Furthermore, we define the neutral elements $0_{[\mathscr{C}]}$ and $1_{[\mathscr{C}]}$ in the following way:

$$\begin{aligned}0_{[\mathscr{C}]} &:= [(0_n)] \quad \text{where } 0_n = 0 \text{ for all } n \in \mathcal{N} \\ 1_{[\mathscr{C}]} &:= [(1_n)] \quad \text{where } 1_n = 1 \text{ for all } n \in \mathcal{N}\end{aligned}$$

Linear ordering: Let $r, s \in \mathbb{R}_{\mathcal{N}}^{\mathscr{C}}$ such that $r = [(a_n)]$ and $s = [(b_n)]$. Then we define:

$$r < s :\Longleftrightarrow \exists \varepsilon \in \mathbb{Q}_{\mathcal{N}}^{+}\, \exists N \in \mathcal{N}\, \forall n \geq N (b_n - a_n > \varepsilon)$$

Again, we have to verify that this definition is well-defined.

THEOREM 17.1. *The structure* $\mathbb{R}_{\mathcal{N}}^{\mathscr{C}} = (\mathbb{R}_{\mathcal{N}}^{\mathscr{C}}, 0_{[\mathscr{C}]}, 1_{[\mathscr{C}]}, +, \cdot, <)$ *is a model of the axioms of the real numbers.*

Proof. The only non-trivial axioms are the existence of a multiplicative inverse and the completeness axiom. Suppose that $r \neq 0_{[\mathscr{C}]}$ is a real number represented by (a_n). Then we define $r^{-1} = [(\tilde{a}_n)]$, where

$$\tilde{a}_n := \begin{cases} \frac{1}{a_n} & \text{if } a_n \neq 0, \\ 1 & \text{otherwise.} \end{cases}$$

Since (a_n) is a Cauchy sequence such that $[(a_n)] \neq 0_{[\mathscr{C}]}$, only for finitely many $n \in \mathcal{N}$ we have $a_n = 0$. Thus, $\lim_{n\to\infty}(a_n \tilde{a}_n - 1) = 0$ and hence $r \cdot r^{-1} = 1_{[\mathscr{C}]}$.

In order to prove that $\mathbb{R}_{\mathcal{N}}^{\mathscr{C}}$ is complete, we first verify R_{15}. Suppose that (r_n) is a Cauchy sequence of real numbers and let r_n be represented by (a_k^n), where (a_k^n) is a Cauchy sequence of rational numbers. Since (a_k^n) is a Cauchy sequence, for every $n \in \mathcal{N}$ there is $N_n \in \mathcal{N}$ such that

$$\forall k, l \geq N_n \big(|a_k^n - a_l^n| < \tfrac{1}{n}\big).$$

Now we consider the diagonal sequence (d_n) with $d_n := a_{N_n}^n$ for every $n \in \mathcal{N}$.

CLAIM. (d_n) *is a Cauchy sequence of rationals which represents a real number* $r = [(d_n)]$ *such that* $\lim_{n\to\infty} r_n = r$.

Proof of Claim. First we show that (d_n) is a Cauchy sequence, and then we prove that it represents a limit of the sequence (r_n) of reals.

(d_n) *is a Cauchy sequence*: Suppose that $\varepsilon \in \mathbb{Q}^+_{\mathcal{N}}$. Note that since (r_n) is a Cauchy sequence of reals, there exists $N \in \mathcal{N}$ with $N \geq \frac{3}{\varepsilon}$ such that

$$\forall m, n \geq N\big(|r_m - r_n| < \tfrac{\varepsilon}{3}\big).$$

In particular, this implies that for all $m, n \in \mathcal{N}$ with $m, n \geq N$, there is $N_{m,n}$ such that

$$\forall k \geq N_{m,n}\big(|a^m_k - a^n_k| < \tfrac{\varepsilon}{3}\big).$$

Now let $m, n \in \mathcal{N}$ such that $n, m \geq N$. We have to verify that $|d_m - d_n| < \varepsilon$. Choose $k \in \mathcal{N}$ with $k \geq N_{m,n}$. We have

$$\begin{aligned} |d_m - d_n| &= |a^m_{N_m} - a^n_{N_n}| \\ &\leq \underbrace{|a^m_{N_m} - a^m_k|}_{<\frac{1}{m}\leq\frac{1}{N}<\frac{\varepsilon}{3}} + \underbrace{|a^m_k - a^n_k|}_{<\frac{\varepsilon}{3}} + \underbrace{|a^n_k - a^n_{N_n}|}_{<\frac{1}{n}\leq\frac{1}{N}<\frac{\varepsilon}{3}}. \end{aligned}$$

Hence we have $|d_n - d_n| < \varepsilon$, which proves that (d_n) is a Cauchy sequence.

(r_n) *converges to* $r = [(d_n)]$: Suppose that $e \in \mathbb{R}^{\mathscr{C}}_{\mathcal{N}}$ is a positive real number, i.e. $e > 0_{[\mathscr{C}]}$. We need to find $N \in \mathcal{N}$ and $\varepsilon \in \mathbb{Q}^+_{\mathcal{N}}$ such that $|r_n - r| < e$ for all $n \geq N$. Choose a rational Cauchy sequence (b_n) representing e. Since $e > 0_{[\mathscr{C}]}$ the definition of our linear ordering yields $\delta \in \mathbb{Q}^+_{\mathcal{N}}$ and $N_0 \in \mathcal{N}$ such that for all $n \geq N_0$ we have $b_n > \delta$. Moreover, since (d_n) is a Cauchy sequence, there is $N_1 \in \mathcal{N}$ such that for all $m, n \geq N_1$ we have $|d_m - d_n| < \frac{\delta}{3}$. Now let $N \in \mathcal{N}$ be defined by $N := \max\{N_0, N_1, \lceil \frac{3}{\delta} \rceil\}$, where $\lceil \frac{3}{\delta} \rceil$ is the least integer bigger than or equal to $\frac{3}{\delta}$, and let $n \geq N$. We prove that $|r_n - r| < e$, i.e. we show that there is $\varepsilon \in \mathbb{Q}^+_{\mathcal{N}}$ and $N' \in \mathcal{N}$ such that

$$\forall k \geq N' \ \big(b_k - |a^n_k - d_k| > \varepsilon\big).$$

Let $\varepsilon := \frac{\delta}{3}$ and $N' = \max\{N, N_n\}$. Then for each $k \geq N'$ we have

$$\begin{aligned} |a^n_k - d_k| &\leq |a^n_k - d_n| + |d_n - d_k| \\ &= \underbrace{|a^n_k - a^n_{N_n}|}_{<\frac{1}{n}\leq\frac{1}{N}\leq\frac{\delta}{3}} + \underbrace{|d_n - d_k|}_{<\frac{\delta}{3}}, \end{aligned}$$

and hence $|a^n_k - d_k| < \frac{2\delta}{3}$. Since $b_k > \delta$ we further obtain

$$b_k - |a^n_k - d_k| > \delta - \tfrac{2\delta}{3} = \varepsilon.$$

Therefore, we have $\lim_{n\to\infty} r_n = r$. $\dashv_{\text{Claim}}$

Moreover, the Archimedian Axiom R_{16} holds as a consequence of the fact that Cauchy sequences are always bounded: If $r, s \in \mathbb{R}^{\mathscr{C}}_{\mathcal{N}}$ such that $0_{[\mathscr{C}]} < r < s$,

then $\frac{s}{r}$ is again a real number represented by some Cauchy sequence (a_n). Since (a_n) is bounded (as a sequence of rational numbers), there is a natural number $N \in \mathcal{N}$ such that $|a_n| < N$ for all $n \in \mathcal{N}$. Hence, $\frac{s}{r} < N + 1$ and $r(N+1) > s$. ⊣

A Natural Construction of the Real Numbers

In this section, we construct a model of the real numbers in which the real numbers are equivalence classes of certain functions, so-called *slopes*, from $\mathbb{Z}_{\mathcal{N}}$ to $\mathbb{Z}_{\mathcal{N}}$, i.e., from $\mathbb{Z}_{\mathcal{N}}$ we will directly construct a model $\mathbb{R}_{\mathcal{N}}^{\mathscr{S}}$ of the real numbers without first constructing a model of the rational numbers. It will turn out that the models $\mathbb{R}_{\mathcal{N}}^{\mathscr{S}}$ and $\mathbb{R}_{\mathcal{N}}^{\mathscr{C}}$ are isomorphic, no matter whether $\mathcal{N} = \omega$ or $\mathcal{N} = \omega^*$.

A **slope** is a function

$$\lambda : \mathbb{Z}_{\mathcal{N}} \to \mathbb{Z}_{\mathcal{N}}$$

for which there exists an $M_\lambda \in \mathcal{N}$, such that for all $n, m \in \mathbb{Z}_{\mathcal{N}}$ we have

$$\big|\lambda(n+m) - \big(\lambda(n) + \lambda(m)\big)\big| \;\leq\; M_\lambda\,.$$

Roughly speaking, a slope is an almost linear function from $\mathbb{Z}_{\mathcal{N}}$ to $\mathbb{Z}_{\mathcal{N}}$. Let $\mathscr{S}$ denote the set of all slopes. We say that two slopes $\lambda, \lambda' \in \mathscr{S}$ are **equivalent**, denoted by $\lambda \sim \lambda'$, if there exists an $M \in \mathcal{N}$ such that for all $n \in \mathbb{Z}_{\mathcal{N}}$ we have

$$\big|\lambda(n) - \lambda'(n)\big| \;\leq\; M\,.$$

Obviously, the relation $\sim$ is reflexive and symmetric, and if for all $n \in \mathbb{Z}_{\mathcal{N}}$,

$$\big|\lambda(n) - \lambda'(n)\big| \;\leq\; M_{\lambda,\lambda'}$$

and

$$\big|\lambda'(n) - \lambda''(n)\big| \;\leq\; M_{\lambda',\lambda''},$$

then

$$\big|\lambda(n) - \lambda''(n)\big| \;\leq\; (M_{\lambda,\lambda'} + M_{\lambda',\lambda''}),$$

which shows that $\sim$ is also transitive. Therefore, $\sim$ is an equivalence relation on $\mathscr{S}$. For a slope $\lambda \in \mathscr{S}$, let

$$[\lambda] := \big\{\lambda' \in \mathscr{S} : \lambda' \sim \lambda\big\}.$$

Let $\mathbb{R}_{\mathcal{N}}^{\mathscr{S}}$ denote the set of equivalence classes of slopes, i.e.,

$$\mathbb{R}_{\mathcal{N}}^{\mathscr{S}} = \big\{[\lambda] : \lambda \in \mathscr{S}\big\}.$$

The elements of $\mathbb{R}_{\mathcal{N}}^{\mathscr{S}}$ are denoted by letters like $r, s, t, \ldots$ and are called **real numbers**.

In what follows, we shall first define two binary functions addition $+$ and multiplication $\cdot$ on $\mathbb{R}_{\mathcal{N}}^{\mathscr{S}}$, including the neutral elements $0_{[\mathscr{S}]}$ and $1_{[\mathscr{S}]}$, respectively; then we introduce the binary relation $<$; and finally, we shall define an isomorphism between the structures $\mathbb{R}_{\mathcal{N}}^{\mathscr{C}} = \big(\mathbb{R}_{\mathcal{N}}^{\mathscr{C}}, 0_{[\mathscr{C}]}, 1_{[\mathscr{C}]}, +, \cdot, <\big)$ and $\mathbb{R}_{\mathcal{N}}^{\mathscr{S}} = \big(\mathbb{R}_{\mathcal{N}}^{\mathscr{S}}, 0_{[\mathscr{S}]}, 1_{[\mathscr{S}]}, +, \cdot, <\big)$, which implies that $\mathbb{R}_{\mathcal{N}}^{\mathscr{S}}$ is a model of the reals.

Addition: Let $r, s \in \mathbb{R}_{\mathcal{N}}^{\mathscr{S}}$ be two reals. Then there are slopes $\lambda, \lambda' \in \mathscr{S}$, such that $r = [\lambda]$ and $s = [\lambda']$. We define $r + s$ by stipulating

$$r + s := [\lambda + \lambda'] \,,$$

where

$$\begin{array}{rccc} \lambda + \lambda' : & \mathbb{Z}_{\mathcal{N}} & \to & \mathbb{Z}_{\mathcal{N}} \\ & n & \mapsto & \lambda(n) + \lambda'(n) \,. \end{array}$$

It is easy to see that $\lambda + \lambda'$ is a slope and that $r + s$ is independent of the choice of the representatives λ and λ'. Furthermore, we define

$$0_{[\mathscr{S}]} := [\lambda_0] \quad \text{where } \lambda_0(n) := 0 \text{ for all } n \in \mathbb{Z}_{\mathcal{N}}.$$

We obviously have that $0_{[\mathscr{S}]}$ is a neutral element with respect to addition. For a real $r = [\lambda]$, let

$$-r := [-\lambda] \quad \text{where for all } n \in \mathbb{Z}_{\mathcal{N}},\ (-\lambda)(n) := -\lambda(n) \,.$$

For all reals r, we obviously have $r + (-r) = 0_{[\mathscr{S}]}$, where $r + (-r)$ is usually written as $r - r$.

Multiplication: Let $r, s \in \mathbb{R}_{\mathcal{N}}^{\mathscr{S}}$ be two reals, and let $\lambda, \lambda' \in \mathscr{S}$ be the corresponding slopes. We define $r \cdot s$ by stipulating

$$r \cdot s := [\lambda \circ \lambda'] \,,$$

where

$$\begin{array}{rccc} \lambda \circ \lambda' : & \mathbb{Z}_{\mathcal{N}} & \to & \mathbb{Z}_{\mathcal{N}} \\ & n & \mapsto & \lambda\big(\lambda'(n)\big) \,. \end{array}$$

Furthermore, we define

$$1_{[\mathscr{S}]} := [\lambda_1] \quad \text{where for all } n \in \mathbb{Z}_{\mathcal{N}},\ \lambda_1(n) = n \,.$$

We obviously have that $1_{[\mathscr{S}]}$ is a neutral element with respect to multiplication. However, we have to show that the composition $\lambda \circ \lambda'$ of two slopes is again a slope, and that $r \cdot s$ is independent of the choice of the representatives λ and λ'.

In order to simplify the notation, for a slope $\lambda \in \mathscr{S}$ and any $n, m \in \mathbb{Z}_{\mathcal{N}}$, we say that $\lambda(n+m)$ and $\lambda(n)+\lambda(m)$ are **similar (with respect to λ)**, denoted by

$$\lambda(n+m) \underset{\lambda}{\approx} \lambda(n)+\lambda(m)\,.$$

In fact, $\lambda(n+m) \underset{\lambda}{\approx} \lambda(n)+\lambda(m)$ just means that the absolute value of the difference of $\lambda(n+m)$ and $\lambda(n)+\lambda(m)$ is uniformly bounded (i.e., independently of $n, m \in \mathbb{Z}_{\mathcal{N}}$). Notice that by definition, for each $u \in \mathbb{Z}_{\mathcal{N}}$ we have

$$\lambda(n+m+u) \underset{\lambda}{\approx} \lambda(n)+\lambda(m)\,. \qquad (*)$$

Lemma 17.2. *Let the slopes λ, λ' represent $r \in \mathbb{R}_{\mathcal{N}}^{\mathscr{S}}$, and let the slopes μ, μ' represent $s \in \mathbb{R}_{\mathcal{N}}^{\mathscr{S}}$. Then the compositions $\lambda \circ \mu$ and $\lambda' \circ \mu'$ are equivalent slopes.*

Proof. We first show that $\lambda \circ \mu$ is a slope, i.e., there exists an $M_{\lambda\circ\mu}$ such that for all $n, m \in \mathbb{Z}_{\mathcal{N}}$,

$$\big|\lambda\circ\mu(n+m) - \big(\lambda\circ\mu(n)+\lambda\circ\mu(m)\big)\big| \;\le\; M_{\lambda\circ\mu}\,.$$

Since μ and λ are both slopes, we have the following two relations:

$$\begin{aligned} \mu(n+m) &\underset{\mu}{\approx} \mu(n)+\mu(m) \\ \lambda\big(\underbrace{\mu(n)}_{n'}+\underbrace{\mu(m)}_{m'}\big) &\underset{\lambda}{\approx} \lambda\circ\underbrace{\mu(n)}_{n'}+\lambda\circ\underbrace{\mu(m)}_{m'} \end{aligned}$$

By the former relation, for all $n, m \in \mathbb{Z}_{\mathcal{N}}$ there exists a $u_{n,m}$ with

$$|u_{n,m}| \le M_\mu\,,$$

such that

$$\lambda\big(\mu(n+m)\big) = \lambda\big(\mu(n)+\mu(m)+u_{n,m}\big)\,,$$

and therefore, by $(*)$ we obtain

$$\lambda\big(\mu(n+m)\big) \underset{\lambda}{\approx} \lambda\big(\mu(n)+\mu(m)\big)\,.$$

Thus, by the latter relation we obtain

$$\lambda\circ\mu(n+m) \underset{\lambda}{\approx} \lambda\circ\mu(n)+\lambda\circ\mu(m)\,,$$

which shows that $\lambda\circ\mu$ (as well as $\lambda'\circ\mu'$) is a slope.

In order to see that $\lambda\circ\mu$ and $\lambda'\circ\mu'$ are equivalent slopes, first notice that

$$\lambda\circ\mu(n+m) \underset{\lambda}{\approx} \lambda\big(\mu(n)+\mu(m)\big) \underset{\lambda}{\approx} \lambda\big(\mu'(n)+\mu'(m)\big)\,.$$

Similarly, we have $\lambda' \circ \mu(n+m) \underset{\lambda'}{\approx} \lambda'\big(\mu'(n) + \mu'(m)\big)$, and since $\lambda \sim \lambda'$, there is an $M_{\lambda,\lambda'} \in \mathcal{N}$ such that for all $n \in \mathbb{Z}_{\mathcal{N}}$,

$$\big|\lambda \circ \mu(n) - \lambda' \circ \mu'(n)\big| \leq M_{\lambda,\lambda'}\,,$$

which shows that the slopes $\lambda \circ \mu$ and $\lambda' \circ \mu'$ are equivalent. ⊣

Linear ordering: In order to define the binary relation $<$, we first define the unary relation $\mathrm{pos}(\cdot)$ on $\mathscr{S}$ by stipulating

$$\mathrm{pos}(\lambda) \;:\Longleftrightarrow\; \forall N \in \mathcal{N}\, \exists m \in \mathcal{N}\big(\lambda(m) > N\big)\,.$$

Now, for any slopes $\lambda, \mu \in \mathscr{S}$, we define

$$\lambda < \mu \;:\Longleftrightarrow\; \mathrm{pos}(\mu - \lambda)\,.$$

Notice that the relation $<$ on $\mathscr{S}$ is transitive and that $\mathrm{pos}(\lambda)$ is equivalent to $0_{[\mathscr{S}]} < \lambda$. Finally, we define the relation $<$ on $\mathbb{R}_{\mathcal{N}}^{\mathscr{S}}$ by stipulating

$$[\lambda] < [\mu] \;:\Longleftrightarrow\; \lambda < \mu\,.$$

By the Solution to Exercise 17.0, the relation $<$ is well-defined.

In order to show that the structures $\mathbb{R}_{\mathcal{N}}^{\mathscr{C}}$ and $\mathbb{R}_{\mathcal{N}}^{\mathscr{S}}$ are isomorphic, which implies that $\mathbb{R}_{\mathcal{N}}^{\mathscr{S}}$ is a model of the reals, we first prove the following

Fact 17.3. *Let $\lambda \in \mathscr{S}$ and $M_\lambda \in \mathcal{N}$ be such that for all $n, m \in \mathbb{Z}_{\mathcal{N}}$ we have*

$$\big|\lambda(n+m) - \lambda(n) - \lambda(m)\big| \;\leq\; M_\lambda\,.$$

Then for all $n, m \in \mathbb{Z}_{\mathcal{N}}$ we have

$$\big|\lambda(n) \cdot m - \lambda(n \cdot m)\big| \;\leq\; (m+1) \cdot M_\lambda\,.$$

Proof. Notice that for each $n \in \mathbb{Z}_{\mathcal{N}}$ we have $|\lambda(0)| \leq M_\lambda$, since

$$M_\lambda \geq \big|\lambda(n+0) - \lambda(n) - \lambda(0)\big| = |-\lambda(0)| = |\lambda(0)|\,.$$

The proof is now by induction on m. If $m = 0$, then for all $n \in \mathbb{Z}_{\mathcal{N}}$ we have

$$\big|\lambda(n) \cdot 0 - \lambda(n \cdot 0)\big| = |-\lambda(0)| = |\lambda(0)| \leq M_\lambda.$$

Assume that for some $m \geq 0$ and for all $n \in \mathbb{Z}_{\mathcal{N}}$, we have

$$\big|\lambda(n) \cdot m - \lambda(n \cdot m)\big| \;\leq\; (m+1) \cdot M_\lambda\,.$$

Then, for all $n \in \mathbb{Z}_{\mathcal{N}}$ we have:

$$\begin{aligned}\big|\lambda(n)\cdot(m+1)-\lambda(n\cdot(m+1))\big| &= \big|\lambda(n)\cdot m+\lambda(n)-\lambda(n\cdot m+n)\big|\\ &\le \big|\lambda(n)\cdot m+\lambda(n)-\lambda(n\cdot m)-\lambda(n)\big|+M_\lambda\\ &= \big|\lambda(n)\cdot m-\lambda(n\cdot m)\big|+M_\lambda\\ &\le (m+1)\cdot M_\lambda+M_\lambda\\ &= (m+2)\cdot M_\lambda\end{aligned}$$

This obviously completes the proof. ⊣

Now, we are ready to prove that $\mathbb{R}^{\mathscr{S}}_{\mathcal{N}}$ is a model of the reals.

Proposition 17.4. *The two structures*

$$\mathbb{R}^{\mathscr{C}}_{\mathcal{N}}=(\mathbb{R}^{\mathscr{C}}_{\mathcal{N}},0_{[\mathscr{C}]},1_{[\mathscr{C}]},+,\cdot,<)\quad\textit{and}\quad\mathbb{R}^{\mathscr{S}}_{\mathcal{N}}=(\mathbb{R}^{\mathscr{S}}_{\mathcal{N}},0_{[\mathscr{S}]},1_{[\mathscr{S}]},+,\cdot,<)$$

are isomorphic.

Proof. First, we define a mapping $\gamma:\mathscr{S}\to\mathscr{C}$ which maps each slope $\lambda\in\mathscr{S}$ to a Cauchy sequence $\gamma(\lambda)$. Then we show that $\lambda\sim\lambda'$ if and only if $\gamma(\lambda)\approx\gamma(\lambda')$. With γ, we then define a bijection $\Gamma:\mathbb{R}^{\mathscr{S}}_{\mathcal{N}}\to\mathbb{R}^{\mathscr{C}}_{\mathcal{N}}$ which induces an isomorphism between $\mathbb{R}^{\mathscr{S}}_{\mathcal{N}}$ and $\mathbb{R}^{\mathscr{C}}_{\mathcal{N}}$.

Let $\gamma:\mathscr{S}\to\mathscr{C}$ be defined by stipulating $\gamma(\lambda):=(a^\lambda_n)$, where

$$a^\lambda_n:=\begin{cases}0 & \text{if } n=0,\\ \frac{\lambda(n)}{n} & \text{otherwise.}\end{cases}$$

We have to show that γ is well-defined, i.e., (a^λ_n) is a Cauchy sequence. For this, let $\lambda\in\mathscr{S}$ and consider $\frac{\lambda(n)}{n}-\frac{\lambda(m)}{m}$ for some $n,m\in\mathcal{N}^+$. Notice that

$$\frac{\lambda(n)}{n}-\frac{\lambda(m)}{m}=\frac{\lambda(n)\cdot m-\lambda(m)\cdot n}{n\cdot m},$$

and that by Fact 17.3 we have

$$\big|\lambda(n)\cdot m-\lambda(n\cdot m)\big|\ \le\ (m+1)\cdot M_\lambda\quad\text{and}\quad\big|\lambda(m)\cdot n-\lambda(m\cdot n)\big|\ \le\ (n+1)\cdot M_\lambda\,.$$

Hence,

$$\left|\frac{\lambda(n)\cdot m-\lambda(m)\cdot n}{n\cdot m}\right|\ \le\ \frac{n+m+2}{n\cdot m}\cdot M_\lambda\,,$$

and since M_λ is fixed, for every $\varepsilon\in\mathbb{Q}^+_{\mathcal{N}}$ we find an $N\in\mathcal{N}$ such that for all $n,m\in\mathcal{N}$ with $m,n\ge N$,

$$\left|\frac{\lambda(n)}{n}-\frac{\lambda(m)}{m}\right|\le\frac{n+m+2}{n\cdot m}\cdot M_\lambda\le\varepsilon\,,$$

which shows that (a^λ_n) is a Cauchy sequence.

Now we show that for any slopes $\lambda, \lambda' \in \mathscr{S}$, if $\lambda \sim \lambda'$ then $(a_n^{\lambda}) \approx (a_n^{\lambda'})$. For this purpose, recall that $\lambda \sim \lambda'$ if and only if there exists an $M \in \mathcal{N}$ such that for all $n \in \mathbb{Z}_{\mathcal{N}}$ we have $|\lambda(n) - \lambda'(n)| \leq M$. With respect to the corresponding Cauchy sequences (a_n^{λ}) and $(a_n^{\lambda'})$, this gives us

$$\left|(a_n^{\lambda}) - (a_n^{\lambda'})\right| = \left|\frac{\lambda(n)}{n} - \frac{\lambda'(n)}{n}\right| = \left|\frac{\lambda(n) - \lambda'(n)}{n}\right| \;\leq\; \frac{M}{n}\,,$$

which shows that $(a_n^{\lambda}) \approx (a_n^{\lambda'})$.

Let us define the function $\Gamma : \mathbb{R}_{\mathcal{N}}^{\mathscr{S}} \to \mathbb{R}_{\mathcal{N}}^{\mathscr{C}}$ by stipulating

$$\Gamma\big([\lambda]\big) := \big[\gamma(\lambda)\big].$$

By the above result, the function Γ is well-defined. In order to show that the structures $\mathbb{R}_{\mathcal{N}}^{\mathscr{S}}$ and $\mathbb{R}_{\mathcal{N}}^{\mathscr{C}}$ are isomorphic, we have to show that Γ is a bijection. For this, we show that Γ is surjective and injective.

Γ is surjective: Let $(a_n) \in \mathscr{C}$ be a Cauchy sequence. With respect to (a_n), let $k_1 < k_2 < \ldots$ be a strictly increasing sequence in $\mathcal{N}$ such that for every $n \in \mathcal{N}^+$ we have

$$\forall m_1, m_2 \geq k_n \left(\left|\lfloor n \cdot a_{m_1} \rfloor - \lfloor n \cdot a_{m_2} \rfloor\right| \;\leq\; 1\right),$$

where for a rational $\frac{p}{q}$, $\lfloor \frac{p}{q} \rfloor := \max\{z \in \mathbb{Z}_{\mathcal{N}} : z \leq \frac{p}{q}\}$. In order to see that such a sequence $k_1 < k_2 < \ldots$ exists, notice that since $(a_n) \in \mathscr{C}$, for every $n \in \mathcal{N}^+$ we find a $k_n \in \mathcal{N}$ such that

$$\forall m_1, m_2 \geq k_n \left(\left|a_{m_1} - a_{m_2}\right| \;<\; \frac{1}{n^2}\right).$$

Hence, for $n \in \mathcal{N}^+$ and all $m_1, m_2 \geq k_n$ we obtain

$$\left|n \cdot a_{m_1} - n \cdot a_{m_2}\right| \;<\; \frac{1}{n}$$

and therefore

$$\left|\lfloor n \cdot a_{m_1} \rfloor - \lfloor n \cdot a_{m_2} \rfloor\right| \;\leq\; 1\,.$$

Now, we define $\lambda : \mathbb{Z}_{\mathcal{N}} \to \mathbb{Z}_{\mathcal{N}}$ with respect to (a_n) by stipulating

$$\lambda(n) = \begin{cases} \lfloor n \cdot a_{k_n} \rfloor & \text{for } n \in \mathcal{N}^+, \\ 0 & \text{for } n = 0, \\ -\lfloor -n \cdot a_{k_{-n}} \rfloor & \text{otherwise.} \end{cases}$$

Notice that for all $n \in \mathbb{Z}_{\mathcal{N}}$, we have $\lambda(-n) = -\lambda(n)$. Therefore, in order to show that $\lambda \in \mathscr{S}$ is a slope, it is enough to show that $u_{n,m} := |\lambda(n+m) - \lambda(n) - \lambda(m)|$ is bounded for $n, m \in \mathcal{N}$. Now, for all $n, m \in \mathcal{N}$ we have:

$$\begin{aligned}
u_{n,m} &= |\lambda(n+m) - \lambda(n) - \lambda(m)| \\
&= \left|\lfloor (n+m) \cdot a_{k_{n+m}} \rfloor - \lfloor n \cdot a_{k_n} \rfloor - \lfloor m \cdot a_{k_m} \rfloor\right| \\
&\leq \left|\lfloor n \cdot a_{k_{n+m}} \rfloor + \lfloor m \cdot a_{k_{n+m}} \rfloor - \lfloor n \cdot a_{k_n} \rfloor - \lfloor m \cdot a_{k_m} \rfloor\right| + 1 \\
&= \left|\lfloor n \cdot a_{k_{n+m}} \rfloor - \lfloor n \cdot a_{k_n} \rfloor + \lfloor m \cdot a_{k_{n+m}} \rfloor - \lfloor m \cdot a_{k_m} \rfloor\right| + 1 \\
&\leq \left|\lfloor n \cdot a_{k_{n+m}} \rfloor - \lfloor n \cdot a_{k_n} \rfloor\right| + \left|\lfloor m \cdot a_{k_{n+m}} \rfloor - \lfloor m \cdot a_{k_m} \rfloor\right| + 1 \\
&\leq 1 + 1 + 1 \\
&= 3
\end{aligned}$$

This shows that λ is a slope. Moreover, for $k_0 := 0$ and $a_0 := 0$, we obtain $\gamma(\lambda) \approx (a_{k_n}) \approx (a_n)$, and since $(a_n) \in \mathscr{C}$ was arbitrary, this implies that Γ is surjective.

Γ is injective: We have to show that for any slopes $\lambda, \lambda' \in \mathscr{S}$, if $\lambda \nsim \lambda'$ then $\gamma(\lambda) \not\approx \gamma(\lambda')$. Let $(a_n) = \gamma(\lambda)$ and $(b_n) = \gamma(\lambda')$. Then $a_0 = b_0 = 0$ and for all $n \in \mathcal{N}^+$, $a_n = \frac{\lambda(n)}{n}$ and $b_n = \frac{\lambda'(n)}{n}$. Since $\lambda, \lambda' \in \mathscr{S}$, there are $M_\lambda, M_{\lambda'} \in \mathcal{N}$ such that for all $n, m \in \mathcal{N}$,

$$|\lambda(2n) - 2\lambda(n)| \leq M_\lambda \quad \text{and} \quad |\lambda'(2n) - 2\lambda'(n)| \leq M_{\lambda'}\,.$$

Assume that $\lambda \nsim \lambda'$. Then for each $M \in \mathcal{N}$, there is an $n \in \mathcal{N}$ such that $|\lambda(n) - \lambda'(n)| > M$. Let

$$M_0 := M_\lambda + M_{\lambda'} + 1$$

and let $n_0 \in \mathcal{N}^+$ be such that

$$|\lambda(n_0) - \lambda'(n_0)| > M_0\,.$$

Now, since

$$\left|\big(\lambda(2n_0) - 2\lambda(n_0)\big) - \big(\lambda'(2n_0) - 2\lambda'(n_0)\big)\right| \leq M_\lambda + M_{\lambda'},$$

we obtain

$$\left|\lambda(2n_0) - \lambda'(2n_0)\right| > 2M_0 - (M_\lambda + M_{\lambda'}) = M_\lambda + M_{\lambda'} + 2\,.$$

Similarly, we obtain

$$\left|\lambda(4n_0) - \lambda'(4n_0)\right| > 2(M_\lambda + M_{\lambda'} + 2) - (M_\lambda + M_{\lambda'}) = M_\lambda + M_{\lambda'} + 4\,,$$

and in general, we have

$$\left|\lambda(2^k n_0) - \lambda'(2^k n_0)\right| > M_\lambda + M_{\lambda'} + 2^k\,.$$

For the corresponding Cauchy sequences (a_n) and (b_n), we therefore have

$$|a_{2^k n_0} - b_{2^k n_0}| = \left| \frac{\lambda(2^k n_0)}{2^k n_0} - \frac{\lambda'(2^k n_0)}{2^k n_0} \right| > \frac{M_\lambda + M_{\lambda'} + 2^k}{2^k n_0} \geq \frac{1}{n_0},$$

which shows that $(a_n) \not\approx (b_n)$ and completes the proof that $\Gamma : \mathbb{R}^{\mathscr{S}}_{\mathcal{N}} \to \mathbb{R}^{\mathscr{C}}_{\mathcal{N}}$ is a bijection.

It remains to show that the structures $\mathbb{R}^{\mathscr{S}}_{\mathcal{N}}$ and $\mathbb{R}^{\mathscr{C}}_{\mathcal{N}}$ are isomorphic, which is done in the Solution to Exercise 17.1. ⊣

Non-Standard Models of the Reals

In the previous section, starting with either $\mathcal{N} = \omega$ or $\mathcal{N} = \omega^*$, we have constructed four models of the real numbers, namely $\mathbb{R}^{\mathscr{C}}_{\omega}$, $\mathbb{R}^{\mathscr{C}}_{\omega^*}$, $\mathbb{R}^{\mathscr{S}}_{\omega}$, $\mathbb{R}^{\mathscr{S}}_{\omega^*}$, and we have shown in Proposition 17.4 that $\mathbb{R}^{\mathscr{C}}_{\omega} \cong \mathbb{R}^{\mathscr{S}}_{\omega}$ and $\mathbb{R}^{\mathscr{C}}_{\omega^*} \cong \mathbb{R}^{\mathscr{S}}_{\omega^*}$.

The models $\mathbb{R}^{\mathscr{C}}_{\omega}$ and $\mathbb{R}^{\mathscr{S}}_{\omega}$ correspond to the standard model $\mathbb{R}$ (with respect to some model of ZF), whereas $\mathbb{R}^{\mathscr{C}}_{\omega^*}$ and $\mathbb{R}^{\mathscr{S}}_{\omega^*}$ are isomorphic non-standard models of the reals (constructed in some model of ZFC), denoted by $\mathbb{R}_{\omega^*}$.

Other non-standard models of the reals are obtained by an ultrapower of the standard model $\mathbb{R}$ with respect to some non-trivial ultrafilters $\mathscr{U} \subseteq \mathscr{P}(\omega)$. The models which we obtain with this construction are denoted by $\mathbb{R}^*$. By Łoś's Theorem 15.2 we know that all these models $\mathbb{R}^*$ are elementarily equivalent to $\mathbb{R}$, independent of the choice of the ultrafilter $\mathscr{U}$. Therefore, beside the non-standard models $\mathbb{R}_{\omega^*}$ as constructed above, we also have the non-standard models $\mathbb{R}^*$. It is natural to ask whether the models $\mathbb{R}_{\omega^*}$ are also elementarily equivalent to the standard model $\mathbb{R}$. This is indeed the case. Moreover, if we use the same ultrafilter to construct ω^* (from ω) and $\mathbb{R}^*$ (from $\mathbb{R}$), then the models $\mathbb{R}_{\omega^*}$ and $\mathbb{R}^*$ are isomorphic (see the Solution to Exercise 17.2 for a bijection between $\mathbb{R}^*_{\omega}$ and $\mathbb{R}_{\omega^*}$).

A Brief Introduction to Non-Standard Analysis

The idea of Non-Standard Analysis is that we work simultaneously with two models of the real numbers. One model, let us call it the *ground model*, takes the role of the standard model $\mathbb{R}$, and the other model, which is an ultrapower of $\mathbb{R}$ with respect to an ultrafilter $\mathscr{U}$ over ω, denoted by $\mathbb{R}^*$, is in the view of a non-standard model which is elementarily equivalent to $\mathbb{R}$. Now, we take the standpoint that proper Analysis takes place in the model $\mathbb{R}^*$, but — as people living in $\mathbb{R}$ — we can only "see" the standard part of the reals in $\mathbb{R}^*$. Even though we have quite a restricted view to proper Analysis from the model $\mathbb{R}$, by the fact that the models $\mathbb{R}$ and $\mathbb{R}^*$ are elementarily equivalent, we cannot detect any difference between the two models on the formal level.

In fact, each sentence which is valid in one model is also valid in the other model. For example, in order to solve a problem in $\mathbb{R}$, we can carry out our calculations in $\mathbb{R}^*$, where we can use reals in $\mathbb{R}^*$ which do not exist in $\mathbb{R}$, and at the end we simply "project" the result to $\mathbb{R}$ again.

Let us now have a closer look at the models $\mathbb{R}$ and $\mathbb{R}^*$, and let us fix some notation: The domain of $\mathbb{R}$ is denoted by $\mathbb{R}$ with the natural numbers $\mathbb{N}$, and the domain of $\mathbb{R}^*$ is denoted by $\mathbb{R}^*$ with the natural numbers $\mathbb{N}^*$. The elements of $\mathbb{R}^*$ are equivalence classes $[f]$ of functions $f : \omega \to \mathbb{R}$. For such equivalence classes we usually just write r^*. With the embedding

$$\begin{aligned} \mathbb{R} &\to \mathbb{R}^* \\ r &\mapsto [c_r] \quad \text{where } c_r : \omega \to \{r\} \end{aligned}$$

we obtain that $\mathbb{R}$ is a subset of $\mathbb{R}^*$, and that $\mathbb{N}$ is a subset of $\mathbb{N}^*$. Furthermore, we see that the equivalence class $[d]$, where $d(n) := n$ for all $n \in \mathbb{N}$, is an element of $\mathbb{N}^*$ which is bigger than all elements of $\mathbb{N}$. Thus, $N := [d]$ is an element in $\mathbb{N}^*$ which does not belong to $\mathbb{N}$. On the other hand, the equivalence class $\delta_0 := [d^{-1}]$, where $d^{-1}(n) := \frac{1}{n}$ for all $n \in \mathbb{N} \setminus \{0\}$ and $d^{-1}(0) := 0$, is an element in $\bar{\mathbb{R}}$, for which we have $0 < \delta_0 < \frac{1}{n}$ (for all $n \in \mathbb{N}$). From the viewpoint of $\mathbb{R}^*$, δ_0 is just a positive real, in fact a positive rational. However, from the viewpoint of $\mathbb{R}$, δ_0 does not exist, since it would be an infinitely small real number, a so-called **infinitesimal** (i.e., a non-zero real number whose absolute value is smaller than $\frac{1}{n}$ for any $n \in \mathbb{N} \setminus \{0\}$).

We say that $r^*, s^* \in \mathbb{R}^*$ are **infinitely close**, denoted by $r^* \approx s^*$, if $r^* - s^*$ is infinitesimal. Note that $\approx$ defines an equivalence relation on $\mathbb{R}^*$. Furthermore, let $\bar{\mathbb{R}}$ be the set of all reals $r^* \in \mathbb{R}^*$, such that for some $s_1, s_2 \in \mathbb{R}$ we have $s_1 \leq r^* \leq s_2$. Obviously, we have $\mathbb{R} \subseteq \bar{\mathbb{R}} \subseteq \mathbb{R}^*$.

The following result states that for each real $r^* \in \bar{\mathbb{R}}$ there is a unique real $r \in \mathbb{R}$ which is infinitely close to r^*. This fact allows us to "project" the reals in $\bar{\mathbb{R}}$ to $\mathbb{R}$.

PROPOSITION 17.5. *For each real $r^* \in \bar{\mathbb{R}}$, there is a unique real $r \in \mathbb{R}$ such that $r^* \approx r$.*

Proof. Uniqueness is obvious, since if there are $r_1, r_2 \in \mathbb{R}$ such that $r^* \approx r_1$ and $r^* \approx r_2$, then by transitivity we have $r_1 \approx r_2$. Since $r_1, r_2 \in \mathbb{R}$, it follows that $r_1 - r_2 = 0$ and thus $r_1 = r_2$.

For the existence, we proceed as follows: Let $r^* = [f]$ for some $f : \omega \to \mathbb{R}$ and let $s, t \in \mathbb{R}$ be such that $[c_s] \leq [f] \leq [c_t]$, which implies

$$\big\{n \in \omega : s \leq f(n) \leq t\big\} \in \mathscr{U} \,.$$

We construct sequences (s_n) and (t_n) in $\mathbb{R}$ as follows. Let $s_0 := s$ and $t_0 := t$. Assume that s_n and t_n are already defined and that we have

$$x_n := \big\{n \in \omega : s_n \leq f(n) \leq t_n\big\} \in \mathscr{U} \,.$$

Let

$$y_n := \left\{ n \in \omega : s_n \leq f(n) \leq \tfrac{s_n + t_n}{2} \right\}.$$

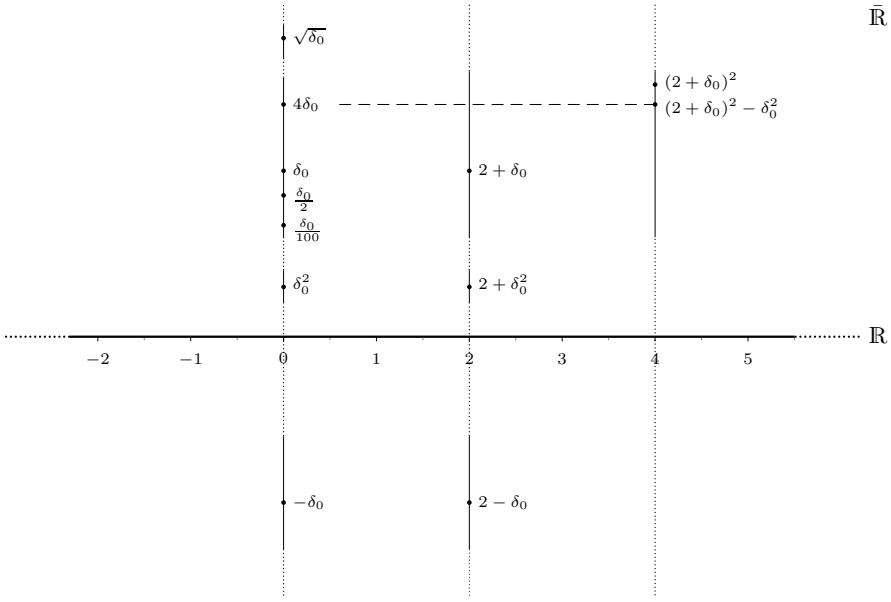
Since $\mathscr{U}$ is an ultrafilter, we have either $y_n \in \mathscr{U}$ or $(\omega \setminus y_n) \in \mathscr{U}$. In the former case, let $s_{n+1} = s_n$ and $t_{n+1} = \frac{s_n + t_n}{2}$, and in the latter case, let $s_{n+1} = \frac{s_n + t_n}{2}$ and $t_{n+1} = t_n$. Notice that since in the former case we have $x_n \cap y_n \in \mathscr{U}$, and in the latter case we have $x_n \cap (\omega \setminus y_n) \in \mathscr{U}$, in both cases we have

$$x_{n+1} := \left\{ n \in \omega : s_{n+1} \leq f(n) \leq t_{n+1} \right\} \in \mathscr{U}.$$

By construction, (s_n) and (t_n) are monotone sequences, where for all $n \in \mathbb{N}$ we have $[c_{s_n}] \leq [f] \leq [c_{t_n}]$. Furthermore, since $\lim_{n\to\infty}(t_n - s_n) = 0$, the supremum $r \in \mathbb{R}$ of (s_n) is equal to the infimum of (t_n), which shows that $r^* \approx r$. ⊣

For $r^* \in \bar{\mathbb{R}}$, the unique $r \in \mathbb{R}$ such that $r \approx r^*$ is called the **standard part** of r^*, denoted by $\mathrm{st}(r^*)$. As mentioned above, we can consider the standard part of a real $r^* \in \bar{\mathbb{R}}$ as a projection of r^* to $\mathbb{R}$, similar to the real part of a complex number. Moreover, Proposition 17.5 shows that every $r^* \in \bar{\mathbb{R}}$ is of the form $r^* = [c_r] + [f_\delta]$, where $r = \mathrm{st}(r^*)$ and $f_\delta : \omega \to \mathbb{R}$ is such that $\lim_{n\to\infty} f_\delta(n) = 0$, i.e., f_δ is a zero sequence. For the sake of simplicity, we just write $r + \delta$ instead of $r^* = [c_r] + [f_\delta]$.

The following figure, in which δ_0 is an arbitrary infinitesimal, visualises the ordering of $\bar{\mathbb{R}}$, how $\mathbb{R}$ is embedded in $\bar{\mathbb{R}}$, and shows some simple calculations.



For example, let $N \in \mathbb{N}^* \setminus \mathbb{N}$ and let $\delta_0 := \frac{1}{N}$. Then the standard part of δ_0 is 0, i.e., for people living in $\mathbb{R}$, $\delta_0 \approx 0$. Similarly, $2 + \delta_0 \approx 2$, or more formally, $\mathrm{st}(2 + \delta_0) = 2$. For example, $2 + \delta_0$ belongs to $\bar{\mathbb{R}}$, but N does not belong to $\bar{\mathbb{R}}$, since there is no $s \in \mathbb{R}$ such that $N \leq s$. The set $\bar{\mathbb{R}}$, as a subset of $\mathbb{R}^*$, is linearly ordered by $<^*_{\mathbb{R}}$. Notice that $<^*_{\mathbb{R}}$ restricted to $\mathbb{R}$ is just the usual linear ordering on $\mathbb{R}$ (which follows from the fact that $\mathbb{R}$ is a submodel of $\mathbb{R}^*$ and that $\mathbb{R}$ and $\mathbb{R}^*$ are elementarily equivalent).

The next result shows that in Non-Standard Analysis, one can compute, for example, definite integrals without using limits.

Proposition 17.6. (a) *Let $a \in \mathbb{R}$, let f be a real-valued function which is continuous at a, and let $a^* \in \mathbb{R}^*$ be such that $a^* \approx a$. Then we have:*

$$f(a^*) \approx f(a)$$

(For a proof see Robert [44, Ch. 4].)

(b) *Let $a \in \mathbb{R}$ and let f be a real-valued function. If there exists an $r \in \mathbb{R}$ such that*

$$\mathrm{st}\left(\frac{f(a+\delta) - f(a)}{\delta}\right) = r \quad \text{for all } \delta \approx 0 \text{ with } \delta \neq 0,$$

then f is differentiable at a and we have $f'(a) = r$.

(For a proof see Robert [44, Ch. 5].)

(c) *Let $b \in \mathbb{R}$, let f be a real-valued function which is continuous on the interval $[0, b]$, and let $N \in \mathbb{N}^* \setminus \mathbb{N}$. Then in $\mathbb{R}$ we have:*

$$\int_0^b f(x)\,dx \;=\; \mathrm{st}\Big(\frac{b}{N} \sum_{k=0}^{N-1} f\big(\tfrac{kb}{N}\big)\Big)$$

(For a proof see Robert [44, Ch. 6].)

Other applications of Non-Standard Analysis are given by the following two examples.

Example 17.7. As a first example, we prove **L'Hospital's Rule**: Let f and g be two real-valued functions which are derivable at $x_0 \in \mathbb{R}$, where $f(x_0) = g(x_0) = 0$. Furthermore, let ε be an infinitesimal. Then, by Proposition 17.6.(a), we have

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \mathrm{st}\Big(\frac{f(x_0 + \varepsilon)}{g(x_0 + \varepsilon)}\Big).$$

Since $f(x_0) = 0$, for f we have

$$\mathrm{st}\big(f(x_0+\varepsilon)\big) = \mathrm{st}\big(f(x_0+\varepsilon) - f(x_0)\big)\,,$$

and similarly for g. Thus, we get:

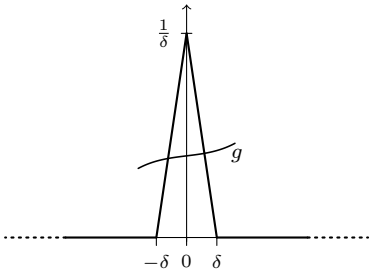
$$\begin{aligned}
\mathrm{st}\Big(\frac{f(x_0+\varepsilon)}{g(x_0+\varepsilon)}\Big) &= \mathrm{st}\Big(\frac{f(x_0+\varepsilon)-f(x_0)}{g(x_0+\varepsilon)-g(x_0)}\Big)\\
&= \mathrm{st}\Big(\frac{f(x_0+\varepsilon)-f(x_0)}{g(x_0+\varepsilon)-g(x_0)}\cdot\frac{\varepsilon}{\varepsilon}\Big)\\
&= \mathrm{st}\Big(\frac{f(x_0+\varepsilon)-f(x_0)}{\varepsilon}\cdot\frac{\varepsilon}{g(x_0+\varepsilon)-g(x_0)}\Big)\\
&= \mathrm{st}\Big(\frac{f(x_0+\varepsilon)-f(x_0)}{\varepsilon}\Big)\ \cdot\ \mathrm{st}\Big(\frac{\varepsilon}{g(x_0+\varepsilon)-g(x_0)}\Big)
\end{aligned}$$

Therefore, by Proposition 17.6.(b), we finally have

$$\lim_{x\to x_0}\frac{f(x)}{g(x)} = \frac{f'(x_0)}{g'(x_0)}\,.$$

Example 17.8. Let us now consider the **Dirac Delta Function**: For this, let g be a real-valued function which is continuous at 0, and let δ be a positive infinitesimal. With respect to δ, we define the function

$$f_\delta(x) = \begin{cases} \frac{x}{\delta^2}+\frac{1}{\delta} & \text{for } -\delta < x \le 0,\\ -\frac{x}{\delta^2}+\frac{1}{\delta} & \text{for } 0 < x < \delta,\\ 0 & \text{otherwise.}\end{cases}$$



Then we obtain

$$\int_{\mathbb{R}} g(x)\cdot f_\delta(x)\,dx \;=\; \int_{-\delta}^{\delta} g(x)\cdot f_\delta(x)\,dx\,.$$

Since g is continuous at 0, by Proposition 17.6.(a), there is an infinitesimal $\varepsilon > 0$, such that for all $x^* \in [-\delta,\delta]$, $g(0)-\varepsilon < g(x^*) < g(0)+\varepsilon$, which implies

$$\big(g(0)-\varepsilon\big)\cdot\int_{-\delta}^{\delta} f_\delta(x)\,dx \;<\; \int_{-\delta}^{\delta} g(x)\cdot f_\delta(x)\,dx \;<\; \big(g(0)+\varepsilon\big)\cdot\int_{-\delta}^{\delta} f_\delta(x)\,dx\,.$$

Now, since $\int_{-\delta}^{\delta} f_\delta(x)\,dx = 1$ and $\mathrm{st}\big(g(0) \pm \varepsilon\big) = g(0)$, this shows that

$$\mathrm{st}\left(\int_{-\delta}^{\delta} g(x) \cdot f_\delta(x)\,dx \right) \;=\; g(0)\,.$$

Computing the Fourier coefficients of f_δ, we obtain $a_0 := \frac{1}{\pi}$, and for all positive $n \in \mathbb{N}^*$ we have $b_n = 0$ and

$$a_n := \frac{2}{\pi n^2 \delta^2}\big(1 - \cos(n\delta)\big).$$

Notice that for all $n \in \mathbb{N}$ we have $\mathrm{st}(a_n) = \frac{1}{\pi}$.

Notes

The natural construction of a model of the real numbers is due to A'Campo [1], who proved all results directly from the properties of slopes — without using Cauchy sequences. The structure of non-standard models of A'Campo's construction of the reals was first studied by Mizrahi [36].

Non-Standard Analysis was developed in the early 1960s by Robinson. Even though the idea of working with infinitesimals can be traced back to Leibniz and L'Hospital, it was Robinson who laid the logical foundations for infinitesimals and infinite numbers (see, for example, [45]).

Exercises

17.0 Let $\lambda, \mu \in \mathscr{S}$, and let $\lambda', \mu' \in \mathscr{S}$ be such that $\lambda' \sim \lambda$ and $\mu' \sim \mu$. Show that $[\lambda] < [\mu]$ if and only if $[\lambda'] < [\mu']$.

17.1 (a) Show that for any slopes $\lambda, \mu \in \mathscr{S}$ we have $\Gamma\big([\lambda] + [\mu]\big) = \Gamma\big([\lambda]\big) + \Gamma\big([\mu]\big)$.
(b) Show that $\Gamma(0_{[\mathscr{S}]}) = 0_{[\mathscr{C}]}$.
(c) Show that for any slopes $\lambda, \mu \in \mathscr{S}$ we have $\Gamma\big([\lambda] \cdot [\mu]\big) = \Gamma\big([\lambda]\big) \cdot \Gamma\big([\mu]\big)$.
(d) Show that $\Gamma(1_{[\mathscr{S}]}) = 1_{[\mathscr{C}]}$.

17.2 Find a bijection between the sets $\mathbb{R}_\omega^*$ and $\mathbb{R}_{\omega^*}$, where these sets are constructed over the same non-trivial ultrafilter. More precisely, construct a bijection between the sets $(\mathbb{R}_\omega^{\mathscr{C}})^*$ and $\mathbb{R}_{\omega^*}^{\mathscr{C}}$.

17.3 Show that for $a^2 \neq 1$,

$$\int_{x=0}^{\pi} \log\big(1 - 2a\cos(x) + a^2\big)dx \;=\; \begin{cases} 0 & \text{if } |a| < 1, \\ \pi \cdot \log a^2 & \text{if } |a| > 1. \end{cases}$$

This exercise is taken from Robert [44], where one can find many more applications of Non-Standard Analysis.

Tautologies

In this section we give a list of some of the most important tautologies. Many of them have been used explicitly and implicitly in several formal proofs.

(A.0) $\vdash \varphi \to \varphi$
(A.1) $\vdash \varphi \leftrightarrow \varphi$

(B) $\{\psi, \varphi\} \vdash \varphi \wedge \psi$

(C) $\vdash (\psi \to \varphi) \to (\psi \to \forall \nu \varphi)$ [for $\nu \notin \text{free}(\psi)$]

(D.0) $\{\varphi_1 \to \varphi_2, \varphi_2 \to \varphi_3\} \vdash \varphi_1 \to \varphi_3$
(D.1) $\{\varphi_1 \to \psi, \varphi_2 \to \psi\} \vdash (\varphi_1 \vee \varphi_2) \to \psi$
(D.2) $\{\psi \to \varphi_1, \psi \to \varphi_2\} \vdash \psi \to (\varphi_1 \wedge \varphi_2)$

(E) $\vdash \varphi \to \big(\psi \to (\varphi \wedge \psi)\big)$

(F) $\vdash \varphi \leftrightarrow \neg\neg\varphi$

(G) $\vdash (\varphi \to \psi) \leftrightarrow (\neg\psi \to \neg\varphi)$

(H.0) $\{\varphi \leftrightarrow \psi\} \vdash \neg\varphi \leftrightarrow \neg\psi$
(H.1) $\{\varphi \leftrightarrow \varphi', \psi \leftrightarrow \psi'\} \vdash (\varphi \to \psi) \leftrightarrow (\varphi' \to \psi')$
(H.2) $\{\varphi \leftrightarrow \varphi', \psi \leftrightarrow \psi'\} \vdash (\varphi \vee \psi) \leftrightarrow (\varphi' \vee \psi')$
(H.3) $\{\varphi \leftrightarrow \varphi', \psi \leftrightarrow \psi'\} \vdash (\varphi \wedge \psi) \leftrightarrow (\varphi' \wedge \psi')$

(I.1) $\vdash (\varphi_1 \wedge \varphi_2) \leftrightarrow (\varphi_2 \wedge \varphi_1)$
(I.2) $\vdash \big((\varphi_1 \wedge \varphi_2) \wedge \varphi_3\big) \leftrightarrow \big(\varphi_1 \wedge (\varphi_2 \wedge \varphi_3)\big)$

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3

(J.0) $\vdash (\varphi_1 \vee \varphi_2) \leftrightarrow (\varphi_2 \vee \varphi_1)$
(J.1) $\vdash \big((\varphi_1 \vee \varphi_2) \vee \varphi_3\big) \leftrightarrow \big(\varphi_1 \vee (\varphi_2 \vee \varphi_3)\big)$

(K) $\vdash (\varphi \to \psi) \leftrightarrow (\neg\varphi \vee \psi)$

(L.0) $\vdash \neg(\varphi \wedge \psi) \leftrightarrow (\neg\varphi \vee \neg\psi)$
(L.1) $\vdash \neg(\varphi \vee \psi) \leftrightarrow (\neg\varphi \wedge \neg\psi)$

(M.0) $\vdash (\varphi_1 \wedge \varphi_2) \vee \varphi_3 \leftrightarrow (\varphi_1 \vee \varphi_3) \wedge (\varphi_2 \vee \varphi_3)$
(M.1) $\vdash (\varphi_1 \vee \varphi_2) \wedge \varphi_3 \leftrightarrow (\varphi_1 \wedge \varphi_3) \vee (\varphi_2 \wedge \varphi_3)$

(N.0) $\vdash \nu = \nu' \leftrightarrow \nu' = \nu$
(N.1) $\vdash (\nu = \nu' \wedge \nu' = \nu'') \to \nu = \nu''$

(O.0) $\vdash \varphi(\nu) \leftrightarrow \varphi(\nu')$ [if ν' does not appear in $\varphi(\nu)$]
(O.1) $\vdash \exists\nu\varphi(\nu) \leftrightarrow \exists\nu'\varphi(\nu')$ [if ν' does not appear in $\varphi(\nu)$]
(O.2) $\vdash \forall\nu\varphi(\nu) \leftrightarrow \forall\nu'\varphi(\nu')$ [if ν' does not appear in $\varphi(\nu)$]

(P.0) $\{\varphi \leftrightarrow \psi\} \vdash \forall\nu\varphi \leftrightarrow \forall\nu\psi$
(P.1) $\{\varphi \leftrightarrow \psi\} \vdash \exists\nu\varphi \leftrightarrow \exists\nu\psi$

(Q.0) $\vdash \neg\exists\nu\varphi \leftrightarrow \forall\nu\neg\varphi$
(Q.1) $\vdash \neg\forall\nu\varphi \leftrightarrow \exists\nu\neg\varphi$

(R) $\vdash \forall\nu\varphi \leftrightarrow \neg\exists\nu\neg\varphi$

(S.0) $\vdash \exists\nu\exists\nu'\varphi \leftrightarrow \exists\nu'\exists\nu\varphi$
(S.1) $\vdash \forall\nu\forall\nu'\varphi \leftrightarrow \forall\nu'\forall\nu\varphi$
(S.2) $\vdash \nexists\nu\exists\nu\varphi \leftrightarrow \exists\nu\varphi$
(S.3) $\vdash \nexists\nu\forall\nu\varphi \leftrightarrow \forall\nu\varphi$

(T.0) $\vdash \exists\nu\varphi \wedge \exists\nu'\psi \leftrightarrow \exists\nu\exists\nu'(\varphi \wedge \psi)$ [for $\nu \notin \mathrm{free}(\psi)$, $\nu' \notin \mathrm{free}(\varphi)$]
(T.1) $\vdash \forall\nu\varphi \wedge \forall\nu'\psi \leftrightarrow \forall\nu\forall\nu'(\varphi \wedge \psi)$ [for $\nu \notin \mathrm{free}(\psi)$, $\nu' \notin \mathrm{free}(\varphi)$]
(T.2) $\vdash \exists\nu\varphi \wedge \psi \leftrightarrow \exists\nu(\varphi \wedge \psi)$ [for $\nu \notin \mathrm{free}(\psi)$]
(T.3) $\vdash \forall\nu\varphi \wedge \psi \leftrightarrow \forall\nu(\varphi \wedge \psi)$ [for $\nu \notin \mathrm{free}(\psi)$]

(U.0) $\vdash \exists\nu\varphi \vee \exists\nu'\psi \leftrightarrow \exists\nu\exists\nu'(\varphi \vee \psi)$ [for $\nu \notin \mathrm{free}(\psi)$, $\nu' \notin \mathrm{free}(\varphi)$]
(U.1) $\vdash \forall\nu\varphi \vee \forall\nu'\psi \leftrightarrow \forall\nu\forall\nu'(\varphi \vee \psi)$ [for $\nu \notin \mathrm{free}(\psi)$, $\nu' \notin \mathrm{free}(\varphi)$]
(U.2) $\vdash \exists\nu\varphi \vee \psi \leftrightarrow \exists\nu(\varphi \vee \psi)$ [for $\nu \notin \mathrm{free}(\psi)$]
(U.3) $\vdash \forall\nu\varphi \vee \psi \leftrightarrow \forall\nu(\varphi \vee \psi)$ [for $\nu \notin \mathrm{free}(\psi)$]

Solutions

Chapter 1

1.0 (a) A formal proof of $\varphi \wedge \psi$ from $\mathbf{\Phi} = \{\varphi, \psi\}$ is given by:

φ_0: φ	$\varphi \in \mathbf{\Phi}$
φ_1: ψ	$\psi \in \mathbf{\Phi}$
φ_2: $\psi \to (\varphi \to (\varphi \wedge \psi))$	instance of L_5
φ_3: $\varphi \to (\varphi \wedge \psi)$	from φ_2 and φ_1 by (MP)
φ_4: $\varphi \wedge \psi$	from φ_3 and φ_0 by (MP)

(b) A formal proof of $\psi \wedge \varphi$ from $\mathbf{\Phi} = \{\varphi \wedge \psi\}$ is given by:

φ_0: $\varphi \wedge \psi$	$\varphi \wedge \psi \in \mathbf{\Phi}$
φ_1: $(\varphi \wedge \psi) \to \varphi$	instance of L_3
φ_2: $(\varphi \wedge \psi) \to \psi$	instance of L_4
φ_3: ψ	from φ_2 and φ_0 by (MP)
φ_4: φ	from φ_1 and φ_0 by (MP)
φ_5: $\psi \wedge \varphi$	by 1.0.(a) with $\tilde{\mathbf{\Phi}} = \{\varphi_3, \varphi_4\}$

which means that φ_5 follows by the formal proof of Exercise 1.0.(a) with respect to $\tilde{\mathbf{\Phi}} = \{\varphi_3, \varphi_4\}$, in more detail:

φ_{5a}: $\varphi \to (\psi \to (\psi \wedge \varphi))$	instance of L_5
φ_{5b}: $\psi \to (\psi \wedge \varphi)$	from φ_{5a} and φ_4 by (MP)
φ_{5c}: $\psi \wedge \varphi$	from φ_{5b} and φ_3 by (MP)

1.1 (a) A formal proof of $(\varphi \vee \psi) \to (\psi \vee \varphi)$, where we set $\vartheta :\equiv \psi \vee \varphi$, is given by:

φ_0: $\psi \to \vartheta$	instance of L_6
φ_1: $\varphi \to \vartheta$	instance of L_7
φ_2: $(\varphi \to \vartheta) \to ((\psi \to \vartheta) \to ((\varphi \vee \psi) \to \vartheta))$	instance of L_8
φ_3: $(\psi \to \vartheta) \to ((\varphi \vee \psi) \to \vartheta)$	from φ_2 and φ_1 by (MP)
φ_4: $(\varphi \vee \psi) \to \vartheta$	from φ_3 and φ_0 by (MP)

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3

(b) A formal proof of $(\varphi \wedge \psi) \to (\psi \wedge \varphi)$, where we set $\vartheta :\equiv \varphi \wedge \psi$, is given by:

φ_0:	$\vartheta \to \varphi$	instance of L_3
φ_1:	$\vartheta \to \psi$	instance of L_4
φ_2:	$\varphi \to (\psi \to (\psi \wedge \varphi))$	instance of L_5
φ_3:	$\varphi_2 \to (\vartheta \to \varphi_2)$	instance of L_1
φ_4:	$\vartheta \to \varphi_2$	from φ_3 and φ_2 by (MP)
φ_5:	$\vartheta \to \underbrace{(\varphi \to (\psi \to (\psi \wedge \varphi)))}_{\equiv \varphi_2} \to ((\vartheta \to \varphi) \to (\vartheta \to (\psi \to (\psi \wedge \varphi))))$	instance of L_2
φ_6:	$(\vartheta \to \varphi) \to (\vartheta \to (\psi \to (\psi \wedge \varphi)))$	from φ_5 and φ_4 by (MP)
φ_7:	$\vartheta \to (\psi \to (\psi \wedge \varphi))$	from φ_6 and φ_0 by (MP)
φ_8:	$(\vartheta \to (\psi \to (\psi \wedge \varphi))) \to ((\vartheta \to \psi) \to (\vartheta \to (\psi \wedge \varphi)))$	instance of L_2
φ_9:	$(\vartheta \to \psi) \to (\vartheta \to (\psi \wedge \varphi))$	from φ_8 and φ_7 by (MP)
φ_{10}:	$\vartheta \to (\psi \wedge \varphi)$	from φ_9 and φ_1 by (MP)

1.2 (a) A formal proof of $(\psi_1 \wedge \psi_2) \wedge \psi_3$ from $\mathbf{\Phi} = \{\psi_1 \wedge (\psi_2 \wedge \psi_3)\}$ is given by:

φ_0:	$\psi_1 \wedge (\psi_2 \wedge \psi_3)$	$\psi_1 \wedge (\psi_2 \wedge \psi_3) \in \mathbf{\Phi}$
φ_1:	$(\psi_1 \wedge (\psi_2 \wedge \psi_3)) \to \psi_1$	instance of L_3
φ_2:	$(\psi_1 \wedge (\psi_2 \wedge \psi_3)) \to (\psi_2 \wedge \psi_3)$	instance of L_4
φ_3:	ψ_1	from φ_1 and φ_0 by (MP)
φ_4:	$\psi_2 \wedge \psi_3$	from φ_2 and φ_0 by (MP)
φ_5:	$(\psi_2 \wedge \psi_3) \to \psi_2$	instance of L_3
φ_6:	$(\psi_2 \wedge \psi_3) \to \psi_3$	instance of L_4
φ_7:	ψ_2	from φ_5 and φ_4 by (MP)
φ_8:	ψ_3	from φ_6 and φ_4 by (MP)
φ_9:	$\psi_2 \to (\psi_1 \to (\psi_1 \wedge \psi_2))$	instance of L_5
φ_{10}:	$\psi_1 \to (\psi_1 \wedge \psi_2)$	from φ_9 and φ_7 by (MP)
φ_{11}:	$\psi_1 \wedge \psi_2$	from φ_{10} and φ_3 by (MP)
φ_{12}:	$\psi_3 \to ((\psi_1 \wedge \psi_2) \to ((\psi_1 \wedge \psi_2) \wedge \psi_3))$	instance of L_5
φ_{13}:	$(\psi_1 \wedge \psi_2) \to ((\psi_1 \wedge \psi_2) \wedge \psi_3)$	from φ_{12} and φ_8 by (MP)
φ_{14}:	$(\psi_1 \wedge \psi_2) \wedge \psi_3$	from φ_{13} and φ_{11} by (MP)

(b) A formal proof of $\psi_1 \wedge (\psi_2 \wedge \psi_3)$ from $\mathbf{\Phi} = \{(\psi_1 \wedge \psi_2) \wedge \psi_3\}$ is given by:

φ_0:	$(\psi_1 \wedge \psi_2) \wedge \psi_3$	$(\psi_1 \wedge \psi_2) \wedge \psi_3 \in \mathbf{\Phi}$
φ_1:	$((\psi_1 \wedge \psi_2) \wedge \psi_3) \to (\psi_1 \wedge \psi_2)$	instance of L_3
φ_2:	$((\psi_1 \wedge \psi_2) \wedge \psi_3) \to \psi_3$	instance of L_4
φ_3:	$\psi_1 \wedge \psi_2$	from φ_1 and φ_0 by (MP)
φ_4:	ψ_3	from φ_2 and φ_0 by (MP)
φ_5:	$(\psi_1 \wedge \psi_2) \to \psi_1$	instance of L_3
φ_6:	$(\psi_1 \wedge \psi_2) \to \psi_2$	instance of L_4
φ_7:	ψ_1	from φ_5 and φ_3 by (MP)
φ_8:	ψ_2	from φ_6 and φ_3 by (MP)
φ_9:	$\psi_3 \to (\psi_2 \to (\psi_2 \wedge \psi_3))$	instance of L_5
φ_{10}:	$\psi_2 \to (\psi_2 \wedge \psi_3)$	from φ_9 and φ_4 by (MP)
φ_{11}:	$\psi_2 \wedge \psi_3$	from φ_{10} and φ_8 by (MP)
φ_{12}:	$(\psi_2 \wedge \psi_3) \to (\psi_1 \to (\psi_1 \wedge (\psi_2 \wedge \psi_3)))$	instance of L_5
φ_{13}:	$\psi_1 \to (\psi_1 \wedge (\psi_2 \wedge \psi_3))$	from φ_{12} and φ_{11} by (MP)
φ_{14}:	$\psi_1 \wedge (\psi_2 \wedge \psi_3)$	from φ_{13} and φ_7 by (MP)

1.3 A formal proof of $\neg\psi \to \neg\varphi$ from $\mathbf{\Phi} = \{\varphi \to \psi\}$ is given by:

φ_0:	$\varphi \to \psi$	$\varphi \to \psi \in \mathbf{\Phi}$
φ_1:	$\varphi_0 \to (\neg\psi \to \varphi_0)$	instance of L_1
φ_2:	$\neg\psi \to \varphi_0$	from φ_1 and φ_0 by (MP)
φ_3:	$(\varphi \to \psi) \to \underbrace{((\varphi \to \neg\psi) \to \neg\varphi)}_{\equiv:\vartheta}$	instance of L_9
φ_4:	$\varphi_3 \to (\neg\psi \to \varphi_3)$	instance of L_1
φ_5:	$\neg\psi \to \varphi_3$	from φ_4 and φ_3 by (MP)
φ_6:	$(\neg\psi \to \underbrace{(\varphi_0 \to \vartheta)}_{\equiv\varphi_3}) \to ((\neg\psi \to \varphi_0) \to (\neg\psi \to \vartheta))$	instance of L_2
φ_7:	$(\neg\psi \to \varphi_0) \to (\neg\psi \to \vartheta)$	from φ_6 and φ_5 by (MP)
φ_8:	$\neg\psi \to \vartheta$	from φ_7 and φ_2 by (MP)
φ_9:	$(\neg\psi \to \underbrace{((\varphi \to \neg\psi) \to \neg\varphi)}_{\equiv\vartheta}) \to ((\neg\psi \to (\varphi \to \neg\psi)) \to (\neg\psi \to \neg\varphi))$	instance of L_2
φ_{10}:	$(\neg\psi \to (\varphi \to \neg\psi)) \to (\neg\psi \to \neg\varphi)$	from φ_9 and φ_8 by (MP)
φ_{11}:	$\neg\psi \to (\varphi \to \neg\psi)$	instance of L_1
φ_{12}:	$\neg\psi \to \neg\varphi$	from φ_{10} and φ_{11} by (MP)

1.4 (a) A formal proof of $\psi_0 \to \psi_2$ from $\mathbf{\Phi} = \{\psi_0 \to \psi_1, \psi_1 \to \psi_2\}$ is given by:

φ_0:	$(\psi_1 \to \psi_2) \to (\psi_0 \to (\psi_1 \to \psi_2))$	instance of L_1
φ_1:	$\psi_1 \to \psi_2$	$\psi_1 \to \psi_2 \in \mathbf{\Phi}$
φ_2:	$\psi_0 \to (\psi_1 \to \psi_2)$	from φ_0 and φ_1 by (MP)
φ_3:	$(\psi_0 \to (\psi_1 \to \psi_2)) \to ((\psi_0 \to \psi_1) \to (\psi_0 \to \psi_2))$	instance of L_2
φ_4:	$(\psi_0 \to \psi_1) \to (\psi_0 \to \psi_2)$	from φ_3 and φ_2 by (MP)
φ_5:	$\psi_0 \to \psi_1$	$\psi_0 \to \psi_1 \in \mathbf{\Phi}$
φ_6:	$\psi_0 \to \psi_2$	from φ_4 and φ_5 by (MP)

(b) A formal proof of $(\psi_0 \vee \psi_1) \to \varphi$ from $\mathbf{\Phi} = \{\psi_0 \to \varphi, \psi_1 \to \varphi\}$ is given by:

φ_0:	$(\psi_0 \to \varphi) \to ((\psi_1 \to \varphi) \to ((\psi_0 \vee \psi_1) \to \varphi))$	instance of L_8
φ_1:	$\psi_0 \to \varphi$	$\psi_0 \to \varphi \in \mathbf{\Phi}$
φ_2:	$(\psi_1 \to \varphi) \to ((\psi_0 \vee \psi_1) \to \varphi)$	from φ_0 and φ_1 by (MP)
φ_3:	$\psi_1 \to \varphi$	$\psi_1 \to \varphi \in \mathbf{\Phi}$
φ_4:	$(\psi_0 \vee \psi_1) \to \varphi$	from φ_2 and φ_3 by (MP)

(c) A formal proof of $\varphi \to (\psi_0 \wedge \psi_1)$ from $\mathbf{\Phi} = \{\varphi \to \psi_0, \varphi \to \psi_1\}$ is given by:

φ_0:	$\varphi \to \psi_0$	$\varphi \to \psi_0 \in \mathbf{\Phi}$
φ_1:	$\varphi \to \psi_1$	$\varphi \to \psi_1 \in \mathbf{\Phi}$
φ_2:	$(\varphi \to \psi_0) \to (\varphi \to (\varphi \to \psi_0))$	instance of L_1
φ_3:	$(\varphi \to \psi_1) \to (\varphi \to (\varphi \to \psi_1))$	instance of L_1
φ_4:	$\varphi \to (\varphi \to \psi_0)$	from φ_2 and φ_0 by (MP)
φ_5:	$\varphi \to (\varphi \to \psi_1)$	from φ_3 and φ_1 by (MP)
φ_6:	$\psi_1 \to \underbrace{(\psi_0 \to (\psi_0 \wedge \psi_1))}_{\equiv:\vartheta}$	instance of L_5
φ_7:	$\varphi_6 \to (\varphi \to \varphi_6)$	instance of L_1
φ_8:	$\varphi \to \varphi_6$	from φ_7 and φ_6 by (MP)

φ_9: $(\varphi \to \underbrace{(\psi_1 \to \vartheta)}_{\equiv\varphi_6}) \to ((\varphi \to \psi_1) \to (\varphi \to \vartheta))$ instance of L_2
φ_{10}: $(\varphi \to \psi_1) \to (\varphi \to \vartheta)$ from φ_9 and φ_8 by (MP)
φ_{11}: $\varphi \to (\psi_0 \to (\psi_0 \wedge \psi_1))$ from φ_{10} and φ_1 by (MP)
φ_{12}: $(\varphi \to (\psi_0 \to (\psi_0 \wedge \psi_1))) \to ((\varphi \to \psi_0) \to (\varphi \to (\psi_0 \wedge \psi_1)))$ instance of L_2
φ_{13}: $(\varphi \to \psi_0) \to (\varphi \to (\psi_0 \wedge \psi_1))$ from φ_{12} and φ_{11} by (MP)
φ_{14}: $\varphi \to (\psi_0 \wedge \psi_1)$ from φ_{13} and φ_0 by (MP)

1.5 A formal proof of $((\psi_1 \wedge \psi_2) \vee \psi_3) \to ((\psi_1 \vee \psi_3) \wedge (\psi_2 \vee \psi_3))$ is given by:

φ_0: $(\psi_1 \wedge \psi_2) \to \psi_1$ instance of L_3
φ_1: $\psi_1 \to (\psi_1 \vee \psi_3)$ instance of L_6
φ_2: $(\psi_1 \wedge \psi_2) \to (\psi_1 \vee \psi_3)$ by 1.4.(a) with $\tilde{\mathbf{\Phi}} = \{\varphi_0, \varphi_1\}$
φ_3: $(\psi_1 \wedge \psi_2) \to \psi_2$ instance of L_4
φ_4: $\psi_2 \to (\psi_2 \vee \psi_3)$ instance of L_6
φ_5: $(\psi_1 \wedge \psi_2) \to (\psi_2 \vee \psi_3)$ by 1.4.(a) with $\tilde{\mathbf{\Phi}} = \{\varphi_3, \varphi_4\}$
φ_6: $(\psi_1 \wedge \psi_2) \to ((\psi_1 \vee \psi_3) \wedge (\psi_2 \vee \psi_3))$ by 1.4.(c) with $\tilde{\mathbf{\Phi}} = \{\varphi_2, \varphi_5\}$
φ_7: $\psi_3 \to (\psi_1 \vee \psi_3)$ instance of L_7
φ_8: $\psi_3 \to (\psi_2 \vee \psi_3)$ instance of L_7
φ_9: $\psi_3 \to ((\psi_1 \vee \psi_3) \wedge (\psi_2 \vee \psi_3))$ by 1.4.(c) with $\tilde{\mathbf{\Phi}} = \{\varphi_7, \varphi_8\}$
φ_{10}: $((\psi_1 \wedge \psi_2) \vee \psi_3) \to ((\psi_1 \vee \psi_3) \wedge (\psi_2 \vee \psi_3))$ by 1.4.(b) with $\tilde{\mathbf{\Phi}} = \{\varphi_5, \varphi_9\}$

1.6 A formal proof of $\forall x \forall y\, (x = y \to y = x)$ is given by:

φ_0: $((x = y \wedge x = x) \to (x = x \to y = x))$ instance of L_{15}
φ_1: $((x = y \wedge x = x) \to (x = x \to y = x)) \to (((x = y \wedge x = x) \to x = x) \to ((x = y \wedge x = x) \to y = x))$ instance of L_2
φ_2: $((x = y \wedge x = x) \to x = x) \to ((x = y \wedge x = x) \to y = x)$ from φ_1 and φ_0 by (MP)
φ_3: $(x = y \wedge x = x) \to x = x$ instance of L_4
φ_4: $(x = y \wedge x = x) \to y = x$ from φ_2 and φ_3 by (MP)
φ_5: $((x = y \wedge x = x) \to y = x) \to (x = y \to ((x = y \wedge x = x) \to y = x))$ instance of L_1
φ_6: $x = y \to ((x = y \wedge x = x) \to y = x)$ from φ_5 and φ_4 by (MP)
φ_7: $(x = y \to ((x = y \wedge x = x) \to y = x)) \to ((x = y \to (x = y \wedge x = x)) \to (x = y \to y = x))$ instance of L_2
φ_8: $(x = y \to (x = y \wedge x = x)) \to (x = y \to y = x)$ from φ_7 and φ_6 by (MP)
φ_9: $x = x$ instance of L_{14}
φ_{10}: $x = x \to (x = y \to (x = y \wedge x = x))$ instance of L_5
φ_{11}: $x = y \to (x = y \wedge x = x)$ from φ_{10} and φ_9 by (MP)
φ_{12}: $x = y \to y = x$ from φ_8 and φ_{11} by (MP)
φ_{13}: $\forall y\, (x = y \to y = x)$ from φ_{12} by $(\forall)$
φ_{14}: $\forall x \forall y\, (x = y \to y = x)$ from φ_{13} by $(\forall)$

Chapter 2

2.0 (a) Using THEOREM 1.7, we can reduce the problem to the following claim: For every formula containing only $\neg$ and $\wedge$ as logical operators, there is an equivalent formula using only $\tilde{\wedge}$. To show this claim, we use THEOREM 1.6 with the following equivalences:

(i) $\neg\varphi \Leftrightarrow \varphi \,\tilde{\wedge}\, \varphi$
(ii) $\varphi \wedge \psi \Leftrightarrow (\varphi \,\tilde{\wedge}\, \psi) \,\tilde{\wedge}\, (\varphi \,\tilde{\wedge}\, \psi)$

For (i), notice that with (DT), (MP), L_3 and L_5 we can easily obtain $\varphi \Leftrightarrow \varphi \wedge \varphi$. Now, with THEOREM 1.6 we then find

$$\neg\varphi \Leftrightarrow \neg(\varphi \wedge \varphi) \Leftrightarrow \varphi \,\tilde{\wedge}\, \varphi.$$

For (ii) we again use THEOREM 1.6 which gives us the following chain of equivalences:

$$\begin{aligned}\varphi \wedge \psi &\Leftrightarrow \neg\neg(\varphi \wedge \psi) \Leftrightarrow (\neg(\varphi \wedge \psi)) \,\tilde{\wedge}\, (\neg(\varphi \wedge \psi)) \\ &\Leftrightarrow (\varphi \,\tilde{\wedge}\, \psi) \,\tilde{\wedge}\, (\varphi \,\tilde{\wedge}\, \psi).\end{aligned}$$

where we used TAUTOLOGY (F) and (i).

(b) Just as above we only have to show:

(iii) $\neg\varphi \Leftrightarrow \varphi \,\tilde{\vee}\, \varphi$
(iv) $\varphi \wedge \psi \Leftrightarrow (\varphi \,\tilde{\vee}\, \varphi) \,\tilde{\vee}\, (\psi \,\tilde{\vee}\, \psi)$

The first equivalence is immediate from $\varphi \Leftrightarrow \varphi \vee \varphi$ which requires L_6, L_8, and TAUTOLOGY (A.1).
For (iv) we again use THEOREM 1.6 which gives us the following chain of equivalences:

$$\begin{aligned}\varphi \wedge \psi &\Leftrightarrow \neg\neg(\varphi \wedge \psi) \Leftrightarrow \neg(\neg\varphi \vee \neg\psi) \\ &\Leftrightarrow \neg((\varphi \,\tilde{\vee}\, \varphi) \vee (\psi \,\tilde{\vee}\, \psi)) \Leftrightarrow (\varphi \,\tilde{\vee}\, \varphi) \,\tilde{\vee}\, (\psi \,\tilde{\vee}\, \psi).\end{aligned}$$

where we used TAUTOLOGY (F) and (L.0).

2.1 By TAUTOLOGY (A.1) we have reflexivity. For symmetry, consider two formulae φ, ψ with $\varphi \Leftrightarrow \psi$. We want to prove $\{\varphi \leftrightarrow \psi\} \vdash \psi \leftrightarrow \varphi$. Now, since $\varphi \leftrightarrow \psi$ is just an abbreviation for $(\varphi \to \psi) \wedge (\psi \to \varphi)$, this follows directly from TAUTOLOGY (I.1).
In order to show transitivity, suppose $\varphi_1 \Leftrightarrow \varphi_2$, $\varphi_2 \Leftrightarrow \varphi_3$. As we can obtain $\varphi_1 \leftrightarrow \varphi_3$ from $\varphi_1 \leftrightarrow \varphi_2$ by replacing an occurrence of φ_2 by φ_3, we apply THEOREM 1.6 to find:

$$\varphi_2 \Leftrightarrow \varphi_3 \quad \Longrightarrow \quad \varphi_1 \leftrightarrow \varphi_2 \Leftrightarrow \varphi_1 \leftrightarrow \varphi_3$$

Alternatively, one can also prove $\{\varphi_1 \leftrightarrow \varphi_2, \varphi_2 \leftrightarrow \varphi_3\} \vdash \varphi_1 \leftrightarrow \varphi_3$ directly with L_1, L_2, and L_5.

2.2 For the direction that $\mathbf{\Phi} \vdash \varphi$ implies $\mathbf{\Phi} \models \varphi$ for every formula φ, there are six rules left to check. Namely $(\mathsf{I}\wedge)$, $(\mathsf{E}\wedge)$, $(\mathsf{I}\forall)$, $(\mathsf{E}\forall)$, $(\mathsf{I}=)$ and $(\mathsf{E}=)$.

For $(\mathsf{I}\wedge)$, suppose that $\mathbf{\Phi} \vdash \varphi$ and $\mathbf{\Phi} \vdash \psi$. We verify that $\mathbf{\Phi} \vdash \varphi \wedge \psi$:

φ_0:	$\psi \to (\varphi \to (\varphi \wedge \psi))$	instance of L_5
φ_1:	ψ	by assumption
φ_2:	$\varphi \to (\varphi \wedge \psi)$	from φ_0 and φ_1 by (MP)
φ_3:	φ	by assumption
φ_4:	$\varphi \wedge \psi$	from φ_2 and φ_3 by (MP)

For $(\mathsf{E}\wedge)$, suppose that $\mathbf{\Phi} \vdash \varphi \wedge \psi$. We verify that $\mathbf{\Phi} \vdash \varphi$ as well as $\mathbf{\Phi} \vdash \psi$.

φ_0:	$\varphi \wedge \psi$	by assumption
φ_1:	$(\varphi \wedge \psi) \to \varphi$	instance of L_3
φ_2:	$(\varphi \wedge \psi) \to \psi$	instance of L_4
φ_3:	φ	from φ_1 and φ_0 by (MP)
φ_4:	ψ	from φ_2 and φ_0 by (MP)

For $(\mathsf{I}\forall)$, let ν be a variable such that $\nu \notin \mathrm{free}(\chi)$ for any $\chi \in \mathbf{\Phi}$ and suppose that $\mathbf{\Phi} \vdash \varphi(\nu)$. Then by $(\forall)$ it follows that $\mathbf{\Phi} \vdash \forall\nu\varphi(\nu)$.

For $(\mathsf{E}\forall)$, let ν be a variable and τ a term such that $\nu \notin \mathrm{free}(\chi)$ for any $\chi \in \mathbf{\Phi}$ and the substitution $\varphi(\nu/\tau)$ is admissible. Suppose that $\mathbf{\Phi} \vdash \forall\nu\varphi(\nu)$. We verify that $\mathbf{\Phi} \vdash \varphi(\tau)$:

φ_0:	$\forall\nu\varphi(\nu) \to \varphi(\tau)$	instance of L_{10}
φ_1:	$\forall\nu\varphi(\nu)$	by assumption
φ_2:	$\varphi(\tau)$	from φ_0 and φ_1 by (MP)

The rule $(\mathsf{I}=)$ follows directly from L_{14}.

In order to check $(\mathsf{E}=)$, we first prove

$$\{\tau_1 = \tau_2\} \vdash \sigma(\nu/\tau_1) = \sigma(\nu/\tau_2)$$

by induction on term construction. If σ is a variable or a constant symbol this is obvious. Suppose that σ is the term $F\sigma_1, \cdots, \sigma_n$. Since we may assume that $\{\tau_1 = \tau_2\} \vdash \sigma_i(\nu/\tau_1) = \sigma_i(\nu/\tau_2)$ by induction, we obtain $\{\tau_1 = \tau_2\} \vdash \sigma(\nu/\tau_1) = \sigma(\nu/\tau_2)$ by L_{16}.
Now note that an atomic formula φ is either of the form $\sigma_1 = \sigma_2$ or $R\sigma_1, \cdots \sigma_n$ for terms $\sigma_1, \ldots, \sigma_n$ and a relation symbol R. We consider only the case that φ is $R\sigma_1, \cdots, \sigma_n$. If we assume $\tau_1 = \tau_2$ then by the observation above we have $\sigma_i(\nu/\tau_1) = \sigma_i(\nu/\tau_2)$ for all i. Since $\varphi(\nu/\tau_i)$ is simply the formula $R\sigma(\nu/\tau_i), \cdots, \sigma(\nu/\tau_i)$, $(\mathsf{E}=)$ follows from L_{15}. This completes the first part of the proof, namely that $\mathbf{\Phi} \vdash \varphi$ implies $\mathbf{\Phi} \models \varphi$ for every formula φ.

Conversely, we still have to check that every logical axiom from L_0-L_{16} can be deduced by the calculus of natural deduction.

Let us start with L_0, namely checking that $\vdash \varphi \vee \neg\varphi$:

$$\begin{array}{rll}
\{\varphi, \neg(\varphi \vee \neg\varphi)\} & \vdash \varphi & \text{by } (\mathsf{IR}) \\
\{\varphi, \neg(\varphi \vee \neg\varphi)\} & \vdash \varphi \vee \neg\varphi & \text{by } (\mathsf{I}\vee) \\
\{\varphi, \neg(\varphi \vee \neg\varphi)\} & \vdash \neg(\varphi \vee \neg\varphi) & \text{by } (\mathsf{IR}) \\
\{\varphi, \neg(\varphi \vee \neg\varphi)\} & \vdash (\varphi \vee \neg\varphi) \wedge \neg(\varphi \vee \neg\varphi) & \text{by } (\mathsf{I}\wedge) \\
\{\neg(\varphi \vee \neg\varphi)\} & \vdash \neg\varphi & \text{by } (\mathsf{I}\neg) \\
\{\neg\varphi, \neg(\varphi \vee \neg\varphi)\} & \vdash \neg\varphi & \text{by } (\mathsf{IR}) \\
\{\neg\varphi, \neg(\varphi \vee \neg\varphi)\} & \vdash \varphi \vee \neg\varphi & \text{by } (\mathsf{I}\vee) \\
\{\neg\varphi, \neg(\varphi \vee \neg\varphi)\} & \vdash \neg(\varphi \vee \neg\varphi) & \text{by } (\mathsf{IR}) \\
\{\neg\varphi, \neg(\varphi \vee \neg\varphi)\} & \vdash (\varphi \vee \neg\varphi) \wedge \neg(\varphi \vee \neg\varphi) & \text{by } (\mathsf{I}\wedge) \\
\{\neg(\varphi \vee \neg\varphi)\} & \vdash \neg\neg\varphi & \text{by } (\mathsf{I}\neg) \\
\{\neg(\varphi \vee \neg\varphi)\} & \vdash \neg\varphi \wedge \neg\neg\varphi & \text{by } (\mathsf{I}\wedge) \\
 & \vdash \neg\neg(\varphi \vee \neg\varphi) & \text{by } (\mathsf{I}\neg) \\
 & \vdash \varphi \vee \neg\varphi & \text{by } (\mathsf{E}\neg)
\end{array}$$

For L_1, we need to show $\vdash \varphi \to (\psi \to \varphi)$.

$$\begin{array}{rll}
\{\varphi, \psi\} & \vdash \varphi & \text{by } (\mathsf{IR}) \\
\{\varphi\} & \vdash \psi \to \varphi & \text{by } (\mathsf{I}\to) \\
 & \vdash \varphi \to (\psi \to \varphi) & \text{by } (\mathsf{I}\to)
\end{array}$$

For L_2, we need to show

$$\vdash (\psi \to (\varphi_1 \to \varphi_2)) \to ((\psi \to \varphi_1) \to (\psi \to \varphi_2))\,.$$

$$\begin{array}{rll}
\{\psi \to (\varphi_1 \to \varphi_2), \psi \to \varphi_1, \psi\} & \vdash \psi & \text{by } (\mathsf{IR}) \\
\{\psi \to (\varphi_1 \to \varphi_2), \psi \to \varphi_1, \psi\} & \vdash \psi \to \varphi_1 & \text{by } (\mathsf{IR}) \\
\{\psi \to (\varphi_1 \to \varphi_2), \psi \to \varphi_1, \psi\} & \vdash \varphi_1 & \text{by } (\mathsf{E}\to) \\
\{\psi \to (\varphi_1 \to \varphi_2), \psi \to \varphi_1, \psi\} & \vdash \psi \to (\varphi_1 \to \varphi_2) & \text{by } (\mathsf{IR}) \\
\{\psi \to (\varphi_1 \to \varphi_2), \psi \to \varphi_1, \psi\} & \vdash \varphi_1 \to \varphi_2 & \text{by } (\mathsf{E}\to) \\
\{\psi \to (\varphi_1 \to \varphi_2), \psi \to \varphi_1, \psi\} & \vdash \varphi_2 & \text{by } (\mathsf{E}\to) \\
\{\psi \to (\varphi_1 \to \varphi_2), \psi \to \varphi_1\} & \vdash \psi \to \varphi_2 & \text{by } (\mathsf{I}\to) \\
\{\psi \to (\varphi_1 \to \varphi_2)\} & \vdash (\psi \to \varphi_1) \to (\psi \to \varphi_2) & \text{by } (\mathsf{I}\to) \\
 & \vdash (\psi \to (\varphi_1 \to \varphi_2)) \to & \\
 & \quad ((\psi \to \varphi_1) \to (\psi \to \varphi_2)) & \text{by } (\mathsf{I}\to)
\end{array}$$

For L_3, we need to show $\vdash (\varphi \wedge \psi) \to \varphi$.

$$\begin{array}{rll}
\{\varphi \wedge \psi\} & \vdash \varphi \wedge \psi & \text{by } (\mathsf{IR}) \\
\{\varphi \wedge \psi\} & \vdash \varphi & \text{by } (\mathsf{E}\wedge) \\
 & \vdash (\varphi \wedge \psi) \to \varphi & \text{by } (\mathsf{I}\to)
\end{array}$$

For L_4, we need to show $\vdash (\varphi \wedge \psi) \to \psi$.

$$\begin{array}{rll}
\{\varphi \wedge \psi\} & \vdash \varphi \wedge \psi & \text{by } (\mathsf{IR}) \\
\{\varphi \wedge \psi\} & \vdash \psi & \text{by } (\mathsf{E}\wedge) \\
 & \vdash (\varphi \wedge \psi) \to \psi & \text{by } (\mathsf{I}\to)
\end{array}$$

For L_5, we need to show $\vdash \varphi \to (\psi \to (\psi \wedge \varphi))$.

$$\begin{array}{rll}
\{\varphi,\psi\} & \vdash \varphi & \text{by } (\mathsf{IR}) \\
\{\varphi,\psi\} & \vdash \psi & \text{by } (\mathsf{IR}) \\
\{\varphi,\psi\} & \vdash \psi \wedge \varphi & \text{by } (\mathsf{I}\wedge) \\
\{\varphi\} & \vdash \psi \to (\psi \wedge \varphi) & \text{by } (\mathsf{I}\to) \\
 & \vdash \varphi \to (\psi \to (\psi \wedge \varphi)) & \text{by } (\mathsf{I}\to)
\end{array}$$

For L_6, we need to show $\vdash \varphi \to (\varphi \vee \psi)$.

$$\begin{array}{rll}
\{\varphi\} & \vdash \varphi & \text{by } (\mathsf{IR}) \\
\{\varphi\} & \vdash \varphi \vee \psi & \text{by } (\mathsf{I}\vee) \\
 & \vdash \varphi \to (\varphi \vee \psi) & \text{by } (\mathsf{I}\to)
\end{array}$$

For L_7, we need to show $\vdash \psi \to (\varphi \vee \psi)$.

$$\begin{array}{rll}
\{\psi\} & \vdash \varphi & \text{by } (\mathsf{IR}) \\
\{\psi\} & \vdash \varphi \vee \psi & \text{by } (\mathsf{I}\vee) \\
 & \vdash \psi \to (\varphi \vee \psi) & \text{by } (\mathsf{I}\to)
\end{array}$$

For L_8, we need to show

$$\vdash (\varphi_1 \to \varphi_3) \to ((\varphi_2 \to \varphi_3) \to ((\varphi_1 \vee \varphi_2) \to \varphi_3))\,.$$

$$\begin{array}{rll}
\{\varphi_1 \to \varphi_3, \varphi_2 \to \varphi_3, \varphi_1 \vee \varphi_2\} & \vdash \varphi_1 \vee \varphi_2 & \text{by } (\mathsf{IR}) \\
\{\varphi_1 \to \varphi_3, \varphi_2 \to \varphi_3, \varphi_1 \vee \varphi_2, \varphi_1\} & \vdash \varphi_1 & \text{by } (\mathsf{IR}) \\
\{\varphi_1 \to \varphi_3, \varphi_2 \to \varphi_3, \varphi_1 \vee \varphi_2, \varphi_1\} & \vdash \varphi_1 \to \varphi_3 & \text{by } (\mathsf{IR}) \\
\{\varphi_1 \to \varphi_3, \varphi_2 \to \varphi_3, \varphi_1 \vee \varphi_2, \varphi_1\} & \vdash \varphi_3 & \text{by } (\mathsf{E}\to) \\
\{\varphi_1 \to \varphi_3, \varphi_2 \to \varphi_3, \varphi_1 \vee \varphi_2, \varphi_2\} & \vdash \varphi_2 & \text{by } (\mathsf{IR}) \\
\{\varphi_1 \to \varphi_3, \varphi_2 \to \varphi_3, \varphi_1 \vee \varphi_2, \varphi_2\} & \vdash \varphi_2 \to \varphi_3 & \text{by } (\mathsf{IR}) \\
\{\varphi_1 \to \varphi_3, \varphi_2 \to \varphi_3, \varphi_1 \vee \varphi_2, \varphi_2\} & \vdash \varphi_3 & \text{by } (\mathsf{E}\to) \\
\{\varphi_1 \to \varphi_3, \varphi_2 \to \varphi_3, \varphi_1 \vee \varphi_2\} & \vdash \varphi_3 & \text{by } (\mathsf{E}\vee) \\
\{\varphi_1 \to \varphi_3, \varphi_2 \to \varphi_3\} & \vdash (\varphi_1 \vee \varphi_2) \to \varphi_3 & \text{by } (\mathsf{I}\to) \\
\{\varphi_1 \to \varphi_3\} & \vdash (\varphi_2 \to \varphi_3) \to ((\varphi_1 \vee \varphi_2) \to \varphi_3) & \text{by } (\mathsf{I}\to) \\
 & \vdash (\varphi_1 \to \varphi_3) \to & \\
 & \quad ((\varphi_2 \to \varphi_3) \to ((\varphi_1 \vee \varphi_2) \to \varphi_3)) & \text{by } (\mathsf{I}\to)
\end{array}$$

Note that L_9 and L_{13} are already verified. Furthermore, $(\mathsf{E}\forall)$ implies L_{10} and $(\mathsf{I}\exists)$ implies L_{11}.

For L_{12}, we need to verify

$$\forall\nu(\psi \to \varphi(\nu)) \to (\psi \to \forall\nu\varphi(\nu))$$

for any variable $\nu \notin \mathrm{free}(\psi)$.

$$\begin{array}{rll}
\{\forall\nu(\psi \to \varphi(\nu)), \psi\} & \vdash \forall\nu(\psi \to \varphi(\nu)) & \text{by } (\mathsf{IR}) \\
\{\forall\nu(\psi \to \varphi(\nu)), \psi\} & \vdash \psi \to \varphi(\nu) & \text{by } (\mathsf{E}\forall) \\
\{\forall\nu(\psi \to \varphi(\nu)), \psi\} & \vdash \psi & \text{by } (\mathsf{IR}) \\
\{\forall\nu(\psi \to \varphi(\nu)), \psi\} & \vdash \varphi(\nu) & \text{by } (\mathsf{E}\to) \\
\{\forall\nu(\psi \to \varphi(\nu)), \psi\} & \vdash \forall\nu\varphi(\nu) & \text{by } (\mathsf{I}\forall) \\
\{\forall\nu(\psi \to \varphi(\nu))\} & \vdash \psi \to \forall\nu\varphi(\nu) & \text{by } (\mathsf{I}\to) \\
 & \vdash \forall\nu(\psi \to \varphi(\nu)) \to (\psi \to \forall\nu\varphi(\nu)) & \text{by } (\mathsf{I}\to)
\end{array}$$

L_{14} is immediate from $(\mathsf{I}=)$.

For L_{15}, we need to check that

$$(\tau_1 = \tau_1' \wedge \cdots \wedge \tau_n = \tau_n') \to (R(\tau_1, \ldots, \tau_n) \to R(\tau_1', \ldots, \tau_n')) .$$

$\{\tau_1 = \tau_1' \wedge \cdots \wedge \tau_n = \tau_n', R(\tau_1, \ldots, \tau_n)\} \vdash \tau_i = \tau_i'$ (for each i) by (IR) and (E$\wedge$)
$\{\tau_1 = \tau_1' \wedge \cdots \wedge \tau_n = \tau_n', R(\tau_1, \ldots, \tau_n)\} \vdash R(\tau_1, \ldots, \tau_n)$ by (IR)
$\{\tau_1 = \tau_1' \wedge \cdots \wedge \tau_n = \tau_n', R(\tau_1, \ldots, \tau_n)\} \vdash R(\tau_1', \ldots, \tau_n')$ by (E=)

Note that in the last step we have applied (E=) multiple times. Applying (I$\to$) twice yields the desired result. N

For L_{16}, we need to check that

$$(\tau_1 = \tau_1' \wedge \cdots \wedge \tau_n = \tau_n') \to (F(\tau_1, \ldots, \tau_n) = F(\tau_1', \ldots, \tau_n')) .$$

$\{\tau_1 = \tau_1' \wedge \cdots \wedge \tau_n = \tau_n'\} \vdash \tau_i = \tau_i'$ (for each i) by (IR) and (E$\wedge$)
$\{\tau_1 = \tau_1' \wedge \cdots \wedge \tau_n = \tau_n'\} \vdash F(\tau_1, \ldots, \tau_n) = F(\tau_1, \ldots, \tau_n)$ by (I=)
$\{\tau_1 = \tau_1' \wedge \cdots \wedge \tau_n = \tau_n'\} \vdash F(\tau_1, \ldots, \tau_n) = F(\tau_1', \ldots, \tau_n')$ by (E=)

As in the case of L_{15} we have to apply (E=) multiple times. The result follows by (I$\to$).

2.3 Let $\mathbf{\Phi}$ be a set of formulae, φ a formula, τ a term and ν a variable. We will prove that the following two statements hold:

$$\mathbf{\Phi} \vdash \neg\varphi(\tau) \quad \Longrightarrow \quad \mathbf{\Phi} \vdash \neg\forall\nu\varphi(\nu)$$
$$\mathbf{\Phi} \vdash \varphi(\tau) \quad \Longrightarrow \quad \mathbf{\Phi} \vdash \neg\forall\nu\neg\varphi(\nu)$$

For the first statement, assume $\neg\varphi(\tau)$. Then, by (I$\exists$), we get $\exists\nu\neg\varphi(\nu)$, which is equivalent to $\neg\forall\nu\varphi(\nu)$ by Tautology (Q.1).
For the second statement, assume $\varphi(\tau)$. Then, again by (I$\exists$), we get $\exists\nu\varphi(\nu)$, which is equivalent to $\neg\neg\exists\nu\varphi(\nu)$ by Tautology (F). Thus, applying Tautology (Q.0) and (G), we obtain $\neg\forall\nu\neg\varphi(\nu)$.

2.4 For the first direction $\vdash \neg(\varphi \wedge \psi) \to (\neg\varphi \vee \neg\psi)$ of Tautology (L.0) we first show $\{\neg(\varphi \wedge \psi), \varphi, \psi\} \vdash \lightning$:

φ_0:	φ	assumption
φ_1:	ψ	assumption
φ_2:	$\psi \to (\varphi \to (\varphi \wedge \psi))$	instance of L_5
φ_3:	$\varphi \to (\varphi \wedge \psi)$	from φ_2 and φ_1 by (MP)
φ_4:	$\varphi \wedge \psi$	from φ_3 and φ_0 by (MP)

The contradiction then follows with L_5. Using Corollary 2.8 we obtain $\{\neg(\varphi \wedge \psi), \varphi\} \vdash \neg\psi$ and a simple application of L_7 gives $\{\neg(\varphi \wedge \psi), \varphi\} \vdash \neg\varphi \vee \neg\psi$.

With L_6 we can show $\{\neg(\varphi \wedge \psi), \neg\varphi\} \vdash \neg\varphi \vee \neg\psi$. Finally with ($\vee$1) we obtain $\{\neg(\varphi \wedge \psi)\} \vdash \neg\varphi \vee \neg\psi$, and we can thus conclude using (DT).

Towards $\vdash (\neg\varphi \vee \neg\psi) \to \neg(\varphi \wedge \psi)$ we first show $\{\varphi \wedge \psi, \neg\varphi \vee \neg\psi\} \vdash \lightning$:

φ_0:	$\varphi \wedge \psi$	assumption
φ_1:	$(\varphi \wedge \psi) \to \varphi$	instance of L_3
φ_2:	φ	from φ_0 and φ_1 by (MP)
φ_3:	$(\varphi \wedge \psi) \to \psi$	instance of L_4
φ_4:	ψ	from φ_0 and φ_3 by (MP)
φ_5:	$\neg\varphi \vee \neg\psi$	assumption
φ_6:	$(\neg\varphi \vee \neg\psi) \to (\varphi \to \neg\psi)$	instance of Tautology (K)
φ_7:	$\varphi \to \neg\psi$	from φ_5 and φ_6 by (MP)
φ_8:	$\neg\psi$	from φ_2 and φ_7 by (MP)

The contradiction follows again with L_5. From Corollary 2.8 we obtain $\{\varphi \wedge \psi\} \vdash \neg(\neg\varphi \vee \neg\psi)$, by Proposition 2.9 we get $\{\neg\varphi \vee \neg\psi\} \vdash \neg(\varphi \wedge \psi)$ and finally we conclude with (DT).

The proof of Tautology (R) requires (G). Notice that with Theorem 1.6 and Tautology (K) and (F), we obtain

$$\varphi \to \psi \Leftrightarrow \neg\varphi \vee \psi \qquad \neg\psi \to \neg\varphi \Leftrightarrow \psi \vee \neg\varphi$$

and therefore, Tautology (G) follows directly from commutativity of $\vee$ which in turn is proved easily with L_6, L_7 and L_8.

We begin with $\vdash \neg\neg\forall\nu\varphi \to \neg\exists\nu\neg\varphi$:

φ_0:	$\forall\nu\varphi \to \varphi$	instance of L_{10}
φ_1:	$(\forall\nu\varphi \to \varphi) \to (\neg\varphi \to \neg\forall\nu\varphi)$	instance of (G)
φ_2:	$\neg\varphi \to \neg\forall\nu\varphi$	from φ_0 and φ_1 by (MP)
φ_3:	$\forall\nu(\neg\varphi \to \neg\forall\nu\varphi)$	from φ_2 by ($\forall$)
φ_4:	$\forall\nu(\neg\varphi \to \neg\forall\nu\varphi) \to (\exists\nu\neg\varphi \to \neg\forall\nu\varphi)$	instance of L_{13}
φ_5:	$\exists\nu\neg\varphi \to \neg\forall\nu\varphi$	from φ_3 and φ_4 by (MP)
φ_6:	$(\exists\nu\neg\varphi \to \neg\forall\nu\varphi) \to (\neg\neg\forall\nu\varphi \to \neg\exists\nu\neg\varphi)$	instance of (G)
φ_7:	$\neg\neg\forall\nu\varphi \to \neg\exists\nu\neg\varphi$	from φ_5 and φ_6 by (MP)

From here we obtain $\vdash \forall\nu\varphi \to \neg\exists\nu\neg\varphi$ again with Tautology (F) and Theorem 1.6.

For the other direction note that by Theorem 1.6 we have

$$\begin{aligned}\neg(\neg\exists\nu\neg\varphi \to \forall\nu\varphi) &\Leftrightarrow \neg(\neg\neg\exists\nu\neg\varphi \vee \forall\nu\varphi) \Leftrightarrow \neg(\neg\neg\exists\nu\neg\varphi \vee \neg\neg\forall\nu\varphi)\\ &\Leftrightarrow \neg\neg(\neg\exists\nu\neg\varphi \wedge \neg\forall\nu\varphi) \Leftrightarrow (\neg\exists\nu\neg\varphi \wedge \neg\forall\nu\varphi)\end{aligned}$$

where we used Tautology (K), (F) and (L.0). Thus if $(\neg\exists\nu\neg\varphi \wedge \neg\forall\nu\varphi) \vdash \lightning$ then $\neg(\neg\exists\nu\neg\varphi \to \forall\nu\varphi) \vdash \lightning$ as well and therefore by Corollary 2.8 we obtain $\vdash \neg\exists\nu\neg\varphi \to \forall\nu\varphi$.

We can now show $(\neg\exists\nu\neg\varphi \wedge \neg\forall\nu\varphi) \vdash \lightning$:

φ_0:	$(\neg\exists\nu\neg\varphi \wedge \neg\forall\nu\varphi)$	assumption
φ_1:	$(\neg\exists\nu\neg\varphi \wedge \neg\forall\nu\varphi) \to (\neg\exists\nu\neg\varphi)$	instance of L_3
φ_2:	$\neg\exists\nu\neg\varphi$	from φ_0 and φ_1 by (MP)
φ_3:	$\neg\varphi \to \exists\nu\neg\varphi$	instance of L_{11}
φ_4:	$(\neg\varphi \to \exists\nu\neg\varphi) \to (\neg\exists\nu\neg\varphi \to \neg\neg\varphi)$	instance of (G)
φ_5:	$\neg\exists\nu\neg\varphi \to \neg\neg\varphi$	from φ_3 and φ_4 by (MP)
φ_6:	$\neg\neg\varphi$	from φ_3 and φ_5 by (MP)

φ_7:	$\neg\neg\varphi \to \varphi$	instance of (F)
φ_8:	φ	from φ_6 and φ_7 by (MP)
φ_9:	$\forall\nu\varphi$	from φ_8 by ($\forall$)
φ_{10}:	$(\neg\exists\nu\neg\varphi \wedge \neg\forall\nu\varphi) \to (\neg\forall\nu\varphi)$	instance of L_4
φ_{11}:	$\neg\forall\nu\varphi$	from φ_0 and φ_{10} by (MP)

from which a contradiction is reached as above.

2.5 To complete the proof of Theorem 1.7, it now only remains to prove the equivalence

$$\varphi \vee \psi \Leftrightarrow \neg(\neg\varphi \wedge \neg\psi)\,.$$

However, by Theorem 1.6, this follows immediately from Tautology (F) and (L.0).

2.6 For part (a), let us first consider the case $n = 1$. We prove each implication seperately and start with $\vdash (\neg\varphi \vee \neg\psi) \to \neg(\varphi \wedge \psi)$.

φ_0:	$(\neg\varphi \to \neg(\varphi \wedge \psi)) \to ((\neg\psi \to \neg(\varphi \wedge \psi)) \to ((\neg\varphi \vee \neg\psi) \to \neg(\varphi \wedge \psi)))$	instance of L_8
φ_1:	$(\varphi \wedge \psi) \to \varphi$	instance of L_3
φ_2:	$\neg\varphi \to \neg(\varphi \wedge \psi)$	by Proposition 2.9 and (DT)
φ_3:	$(\neg\psi \to \neg(\varphi \wedge \psi)) \to ((\neg\varphi \vee \neg\psi) \to \neg(\varphi \wedge \psi))$	from φ_0 and φ_2 by (MP)
φ_4:	$\varphi \wedge \psi \to \psi$	instance of L_4
φ_5:	$\neg\psi \to \neg(\varphi \wedge \psi)$	by Proposition 2.9 and (DT)
φ_6:	$(\neg\varphi \vee \neg\psi) \to \neg(\varphi \wedge \psi)$	from φ_3 and φ_5 by (MP)

For the other direction, namely $\vdash \neg(\varphi \wedge \psi) \to (\neg\varphi \vee \neg\psi)$, we will show $\varphi \vdash \neg(\varphi \wedge \psi) \to (\neg\varphi \vee \neg\psi)$ and $\neg\varphi \vdash \neg(\varphi \wedge \psi) \to (\neg\varphi \vee \neg\psi)$ separately, which leads to the desired conclusion by case distinction ($\vee$1). We start with the latter, assuming $\neg\varphi$.

φ_0:	$\neg\varphi \to (\neg\varphi \vee \neg\psi)$	instance of L_6
φ_1:	$\neg\varphi$	by assumption
φ_2:	$\neg\varphi \vee \neg\psi$	from φ_0 and φ_1 by (MP)
φ_3:	$(\neg\varphi \vee \neg\psi) \to (\neg(\varphi \wedge \psi) \to (\neg\varphi \vee \neg\psi))$	instance of L_1
φ_4:	$\neg(\varphi \wedge \psi) \to (\neg\varphi \vee \neg\psi)$	from φ_3 and φ_2 by (MP)

To show $\{\varphi, \neg(\varphi \wedge \psi)\} \vdash (\neg\varphi \vee \neg\psi)$, we argue by contradiction. Assume φ, $\neg(\varphi \wedge \psi)$ and ψ.

φ_0:	$\psi \to (\varphi \to (\varphi \wedge \psi))$	instance of L_5
φ_1:	ψ	by assumption
φ_2:	$\varphi \to (\varphi \wedge \psi)$	from φ_0 and φ_1 by (MP)
φ_3:	φ	by assumption
φ_4:	$\varphi \wedge \psi$	from φ_2 and φ_3 by (MP)
φ_5:	$\neg(\varphi \wedge \psi) \to ((\varphi \wedge \psi) \to ((\varphi \wedge \psi) \wedge \neg(\varphi \wedge \psi)))$	instance of L_5
φ_6:	$\neg(\varphi \wedge \psi)$	by assumption
φ_7:	$(\varphi \wedge \psi) \to ((\varphi \wedge \psi) \wedge \neg(\varphi \wedge \psi))$	from φ_5 and φ_6 by (MP)
φ_8:	$(\varphi \wedge \psi) \wedge \neg(\varphi \wedge \psi)$	from φ_7 and φ_4 by (MP)

Which is a contradiction, hence, by Corollary 2.8 $\{\varphi, \neg(\varphi \wedge \psi)\} \vdash \neg\psi$ and thus, using L_7, $\{\varphi, \neg(\varphi \wedge \psi)\} \vdash \neg\varphi \vee \neg\psi$. So, as mentioned above, we

conclude $\vdash \neg(\varphi \wedge \psi) \to (\neg\varphi \vee \neg\psi)$ by case distinction ($\vee$1). This finishes the case $n = 1$.

The general case follows inductively by the following observation:

$$\begin{aligned}
&(\neg\varphi_0 \vee \neg\varphi_1) \vee \neg\varphi_2 \\
&\qquad \Leftrightarrow \neg(\varphi_0 \wedge \varphi_1) \vee \neg\varphi_2 && \text{by case } n = 1 \text{ applied to } \varphi_0 \text{ and } \varphi_1 \\
&\qquad \Leftrightarrow \neg((\varphi_0 \wedge \varphi_1) \wedge \varphi_2) && \text{by case } n = 1 \text{ applied to } (\varphi_0 \wedge \varphi_1) \text{ and } \varphi_2
\end{aligned}$$

For part (b), we argue as follows:

$$\begin{aligned}
&\neg(\neg\varphi_0 \wedge \cdots \wedge \neg\varphi_n) \\
&\qquad \Leftrightarrow \neg\neg\varphi_0 \vee \cdots \vee \neg\neg\varphi_n && \text{by part (a) applied to } \neg\varphi_i \\
&\qquad \Leftrightarrow \varphi_0 \vee \cdots \vee \varphi_n && \text{by Tautology (F) and (H.2)}
\end{aligned}$$

Which leads to the conclusion, using Proposition 2.9.

For part (c), note that by Tautology (K) we have $\varphi \to \psi \Leftrightarrow \neg\varphi \vee \psi$ for any formulae φ and ψ. Similar to the argument concerning part (a) we iteratively apply the following:

$$\begin{aligned}
&\varphi_0 \to (\varphi_1 \to \varphi_2) \\
&\qquad \Leftrightarrow \neg\varphi_0 \vee (\varphi_1 \to \varphi_2) && \text{by Tautology (K) applied to } \varphi_0 \text{ and } \varphi_1 \to \varphi_2 \\
&\qquad \Leftrightarrow \neg\varphi_0 \vee (\neg\varphi_1 \vee \varphi_2) && \text{by Tautology (K) applied to } \varphi_1 \text{ and } \varphi_2
\end{aligned}$$

2.7 The proof is by induction on the construction of φ. First, note that by the proof of Theorem 1.7, we may assume that φ only contains the symbols $\neg$ and $\wedge$ as logical operators and $\forall$ as quantifier.
If φ is an atomic formula, L_{15} gives the assertion. So, suppose $\varphi \equiv \neg\psi$ for some formula ψ, $(\tau_1 = \tau_1' \wedge \cdots \wedge \tau_n = \tau_n')$, and that the statement already holds for ψ. We verify that $\neg\psi(\tau_1, \ldots, \tau_n) \to \neg\psi(\tau_1', \ldots, \tau_n')$, which proves by (DT) that φ satisfies the assertion. Because of Tautology (G), it is equivalent to check $\psi(\tau_1', \ldots, \tau_n') \to \psi(\tau_1, \ldots, \tau_n)$.
By Tautology (N.0), we have $(\tau_1' = \tau_1 \wedge \cdots \wedge \tau_n' = \tau_n)$ from our assumption. Thus, by applying the assertion to ψ and using (DT), we obtain the desired implication $\psi(\tau_1', \ldots, \tau_n') \to \psi(\tau_1, \ldots, \tau_n)$.

Now, suppose $\varphi \equiv \varphi_1 \wedge \varphi_2$ for some formulae φ_1 and φ_2 which satisfy the assertion. Further, assume the following formulae:

$$\begin{aligned}
&(\tau_1 = \tau_1' \wedge \cdots \wedge \tau_n = \tau_n') \\
&\varphi_1(\tau_1, \ldots, \tau_n) \to \varphi_1(\tau_1', \ldots, \tau_n') \\
&\varphi_2(\tau_1, \ldots, \tau_n) \to \varphi_2(\tau_1', \ldots, \tau_n') \\
&\varphi_1(\tau_1, \ldots, \tau_n) \wedge \varphi_2(\tau_1, \ldots, \tau_n)
\end{aligned}$$

We verify that $\varphi_1(\tau_1', \ldots, \tau_n') \wedge \varphi_2(\tau_1', \ldots, \tau_n')$. By applying L_3 and L_4 we obtain $\varphi_1(\tau_1, \ldots, \tau_n)$ and $\varphi_2(\tau_1, \ldots, \tau_n)$. Since φ_1 and φ_2 satisfy the assertion we obtain $\varphi_1(\tau_1', \ldots, \tau_n')$ and $\varphi_2(\tau_1', \ldots, \tau_n')$ by (MP). Then, L_5 and (MP) lead to $\varphi_1(\tau_1', \ldots, \tau_n') \wedge \varphi_2(\tau_1', \ldots, \tau_n')$.

It remains to check the statement for $\varphi \equiv \forall\nu\psi$ where ν is some variable and ψ is a formula. So, suppose $\tau_2 = \tau_2' \wedge \cdots \wedge \tau_n = \tau_n'$ and

$\forall\nu\,\psi(\nu,\tau_2,\ldots,\tau_n)$. By L_{10}, we get $\psi(\nu,\tau_2,\ldots,\tau_n)$. By our assumptions, this leads to $\psi(\nu,\tau_2',\ldots,\tau_n')$, and applying $(\forall)$ gives the assertion.

2.8 The proof is by induction on the construction of φ. If φ is atomic, there is nothing to prove. So, let $\varphi \equiv \neg\psi$ for some formula ψ in CNF. Using the equivalence $\neg(\xi_1 \wedge \xi_2) \Leftrightarrow (\neg\xi_1 \vee \neg\xi_2)$ (cf. TAUTOLOGY (L.0)), ψ is equivalent to some formula of the form

$$(\psi_{1,1} \wedge \cdots \wedge \psi_{1,k_1}) \vee \cdots \vee (\psi_{m,1} \wedge \cdots \wedge \psi_{m,k_m})$$

for some quantifier-free formulae $\psi_{i,j}$ which are either atomic or negations of atomic formulae. The assertion follows by iteratively applying TAUTOLOGY (M.0).

Similarly, if $\varphi \equiv \varphi_1 \vee \varphi_2$ for some formulae φ_1 and φ_2 in CNF TAUTOLOGY (M.0) does the job.

If now $\varphi \equiv \varphi_1 \wedge \varphi_2$ for some formulae φ_1 and φ_2 in CNF, then φ is in CNF by construction. So, it only remains to check the case where $\varphi \equiv \psi \to \psi'$ for some formulae ψ and ψ' in CNF. Using TAUTOLOGY (K), φ is equivalent to some formula of the form

$$\begin{aligned}\neg\big((\psi_{1,1} \vee \cdots \vee \psi_{1,k_1}) \wedge \cdots \wedge (\psi_{m,1} \vee \cdots \vee \psi_{m,k_m})\big)\\ \vee \big((\psi'_{1,1} \vee \cdots \vee \psi'_{1,k_1}) \wedge \cdots \wedge (\psi'_{m,1} \vee \cdots \vee \psi'_{m,k_m})\big)\end{aligned}$$

for quantifier-free formulae $\psi_{i,j}$ and $\psi'_{i,j}$, which are either atomic or negations of atomic formulae, and this is equivalent to some formula in CNF by the above discussion.

2.9 (a) Define $\chi :\equiv (\varphi \to \psi) \to ((\varphi \to \neg\psi) \to \neg\varphi)$. If we assume that $\vdash \varphi \to \chi$ and $\vdash \neg\varphi \to \chi$ we can conclude with:

φ_0:	$\varphi \to \chi$	assumption
φ_1:	$\neg\varphi \to \chi$	assumption
φ_2:	$(\varphi \to \chi) \to ((\neg\varphi \to \chi) \to ((\varphi \vee \neg\varphi) \to \chi))$	instance of L_8
φ_3:	$(\neg\varphi \to \chi) \to (\varphi \vee \neg\varphi) \to \chi)$	from φ_2 and φ_0 by (MP)
φ_4:	$(\varphi \vee \neg\varphi) \to \chi$	from φ_3 and φ_1 by (MP)
φ_5:	$\varphi \vee \neg\varphi$	instance of L_0
φ_6:	χ	from φ_4 and φ_5 by (MP)

It thus remains to prove the two assumptions above. We first show $\{\varphi \to \psi, \varphi \to \neg\psi, \varphi\} \vdash \neg\varphi$:

φ_0:	φ	assumption
φ_1:	$\varphi \to \psi$	assumption
φ_2:	ψ	from φ_1 and φ_0 by (MP)
φ_3:	$\varphi \to \neg\psi$	assumption
φ_4:	$\neg\psi$	from φ_3 and φ_0 by (MP)
φ_5:	$\neg\psi \to (\psi \to \neg\varphi)$	instance of L_9
φ_6:	$\psi \to \neg\varphi$	from φ_5 and φ_4 by (MP)
φ_7:	$\neg\varphi$	from φ_6 and φ_2 by (MP)

Trivially, we have $\{\varphi \to \psi, \varphi \to \neg\psi, \neg\varphi\} \vdash \neg\varphi$ and with (DT) we then obtain $\vdash \varphi \to \chi$ and $\vdash \neg\varphi \to \chi$.

Note that while we have only used L_0, L_8, and L_9, the axiom schemes L_1 and L_2 are required in the proof of (DT).

(b) We can again apply (DT) and show $\{\mathsf{L}_1, \mathsf{L}_2, \mathsf{L}_{9^3/_4}, \varphi\} \vdash \neg\neg\varphi$:

φ_0:	φ	assumption
φ_1:	$\varphi \to (\neg\varphi \to \varphi)$	instance of L_1
φ_2:	$\neg\varphi \to \varphi$	from φ_1 and φ_0 by (MP)
φ_3:	$\neg\varphi \to (\varphi \to \neg\varphi)$	instance of L_1
φ_4:	$(\neg\varphi \to (\varphi \to \neg\varphi)) \to ((\neg\varphi \to \varphi) \to (\neg\varphi \to \neg\varphi))$	instance of L_2
φ_5:	$(\neg\varphi \to \varphi) \to (\neg\varphi \to \neg\varphi)$	from φ_4 and φ_3 by (MP)
φ_6:	$\neg\varphi \to \neg\varphi$	from φ_5 and φ_2 by (MP)
φ_7:	$(\neg\varphi \to \varphi) \to ((\neg\varphi \to \neg\varphi) \to \neg\neg\varphi)$	instance of $\mathsf{L}_{9^3/_4}$
φ_8:	$(\neg\varphi \to \neg\varphi) \to \neg\neg\varphi$	from φ_7 and φ_2 by (MP)
φ_9:	$\neg\neg\varphi$	from φ_8 and φ_6 by (MP)

(c) We define $|\cdot|$ as in the hint and first show that every instance of axioms L_1–L_9 has value 1:

L_1: Assume $|\mathsf{L}_1| \neq 1$. Then $|\varphi| = 1$ and $|\psi \to \varphi| \neq 1$, and therefore $|\psi| = 1$ and $|\varphi| \neq 1$, which is a contradiction.

L_2: Assume $|\mathsf{L}_2| \neq 1$. Then $|\psi \to (\varphi_1 \to \varphi_2)| = 1$ and $|(\psi \to \varphi_1) \to (\psi \to \varphi_2)| \neq 1$. Thus $|\psi \to \varphi_1| = 1$ and $|\psi \to \varphi_2| \neq 1$, which implies $|\psi| = 1$, $|\varphi_2| \neq 1$ and $|\varphi_1| = 1$. Hence, $|\varphi_1 \to \varphi_2| \neq 1$ and $|\psi \to (\varphi_1 \to \varphi_2)| \neq 1$, which is a contradiction.

L_3: Notice that if we have $|\chi_1| \leq |\chi_2|$ for any formulae χ_1, χ_2, then necessarily $|\chi_1 \to \chi_2| = 1$. As $|\varphi \wedge \psi| \leq |\varphi|$, we clearly have $|\mathsf{L}_1| = 1$.

L_4: Same argument as for L_3.

L_5: Assume $|\mathsf{L}_5| \neq 1$. Then $|\varphi| = 1$ and $|\psi \to (\psi \to \varphi)| \neq 1$. From here we conclude again that $|\psi \to \varphi| \neq 1$ and $|\varphi| \neq 1$, which is a contradiction.

L_6: As $|\varphi \vee \psi| \geq |\varphi|$, we get $|\mathsf{L}_6| = 1$.

L_7: Same argument as for L_6.

L_8: Assume $|\mathsf{L}_8| \neq 1$. Then $|\varphi_1 \to \varphi_3| = 1$, $|(\varphi_2 \to \varphi_3) \to ((\varphi_1 \vee \varphi_2) \to \varphi_3)| \neq 1$, $|(\varphi_2 \to \varphi_3)| = 1$, $|(\varphi_1 \vee \varphi_2) \to \varphi_3| \neq 1$, and therefore $|\varphi_3| \neq 1$ and $|(\varphi_1 \vee \varphi_2)| = 1$. We then obtain $|\varphi_1| = 1$ or $|\varphi_2| = 1$, which in turn implies $|\varphi_1 \to \varphi_3| \neq 1$ or $|\varphi_2 \to \varphi_3| \neq 1$.

L_9: Assume $|\mathsf{L}_9| \neq 1$. Then $|\neg\varphi| = 1$, $|\varphi \to \psi| \neq 1$, $|\varphi| = 1$ and $|\psi| \neq 1$, and since $1 = |\neg\varphi| = -|\varphi| = -1$, we get a contradiction.

Notice that because from $|\varphi| = 1$ and $|\varphi \to \psi| = 1$ it follows that $|\psi| = 1$, any application of (MP) to formulas with value 1 gives us another formula with value 1. In that sense, (MP) is compatible with $|\cdot|$ and any formula that can be proven using only instances of L_1–L_9 and (MP) must have value 1. However, this is not always the case for an instance of $\mathsf{L}_{9^{3}/_{4}}$: If $|\varphi| = 0$, then $|\varphi \to \psi| = 1$, $|\varphi \to \neg\psi| = 1$ and $|(\varphi \to \neg\psi) \to \neg\varphi| = 0$. Thus, we obtain $|\mathsf{L}_{9^{3}/_{4}}| = 0$, which shows that $\mathsf{L}_{9^{3}/_{4}}$ cannot be proven from the axioms L_1–L_9, i.e., $\{\mathsf{L}_1\text{–}\mathsf{L}_9\} \nvdash \mathsf{L}_{9^{3}/_{4}}$.

2.10 For part (a), assume $\neg\neg\varphi$ and argue as follows:

φ_0:	$\neg\neg\varphi \to (\neg\varphi \to \varphi)$	instance of L_9
φ_1:	$\neg\neg\varphi$	by assumption
φ_2:	$\neg\varphi \to \varphi$	from φ_0 and φ_1 by (MP)
φ_3:	$\varphi \vee \neg\varphi$	instance of L_0
φ_4:	$\varphi \to \varphi$	by Example 1.1
φ_5:	$(\varphi \to \varphi) \to ((\neg\varphi \to \varphi) \to ((\varphi \vee \neg\varphi) \to \varphi))$	instance of L_8
φ_6:	$(\neg\varphi \to \varphi) \to ((\varphi \vee \neg\varphi) \to \varphi)$	from φ_5 and φ_4 by (MP)
φ_7:	$(\varphi \vee \neg\varphi) \to \varphi$	from φ_6 and φ_2 by (MP)
φ_8:	φ	from φ_7 and φ_3 by (MP)

For part (b), we first verify that

$$\{\mathsf{L}_1, \mathsf{L}_2, \mathsf{L}_{9^{3}/_{4}}\} \vdash (\varphi \to \psi) \to (\neg\psi \to \neg\varphi) \qquad (*)$$

for any formulae φ and ψ. So, let us assume $\varphi \to \psi$ and $\neg\psi$. We verify that $\neg\varphi$ holds, then $(*)$ follows by (DT).

φ_0:	$\neg\psi \to (\varphi \to \neg\psi)$	instance of L_1
φ_1:	$\neg\psi$	by assumption
φ_2:	$\varphi \to \neg\psi$	from φ_0 and φ_1 by (MP)
φ_3:	$(\varphi \to \psi) \to ((\varphi \to \neg\psi) \to \neg\varphi)$	instance of $\mathsf{L}_{9^{3}/_{4}}$
φ_4:	$\varphi \to \psi$	by assumption
φ_5:	$(\varphi \to \neg\psi) \to \neg\varphi$	from φ_3 and φ_4 by (MP)
φ_6:	$\neg\varphi$	from φ_5 and φ_2 by (MP)

Now, let us prove $\{\mathsf{L}_1, \mathsf{L}_2, \mathsf{L}_6, \mathsf{L}_7, \mathsf{L}_{9^{3}/_{4}}\} \vdash \neg\neg(\varphi \vee \neg\varphi)$.

φ_0:	$\varphi \to (\varphi \vee \neg\varphi)$	instance of L_6
φ_1:	$\neg(\varphi \vee \neg\varphi) \to \neg\varphi$	from φ_0 by $(*)$
φ_2:	$\neg\varphi \to (\varphi \vee \neg\varphi)$	instance of L_7
φ_3:	$\neg(\varphi \vee \neg\varphi) \to \neg\neg\varphi$	from φ_2 by $(*)$
φ_4:	$(\neg(\varphi \vee \neg\varphi) \to \neg\varphi) \to$ $((\neg(\varphi \vee \neg\varphi) \to \neg\neg\varphi) \to \neg\neg(\varphi \vee \neg\varphi))$	instance of $\mathsf{L}_{9^{3}/_{4}}$
φ_5:	$(\neg(\varphi \vee \neg\varphi) \to \neg\neg\varphi) \to \neg\neg(\varphi \vee \neg\varphi)$	from φ_4 and φ_1 by (MP)
φ_6:	$\neg\neg(\varphi \vee \neg\varphi)$	from φ_5 and φ_3 by (MP)

Since the axioms of part (b) are also given in part (c), we may use the formula $\neg\neg(\varphi \vee \neg\varphi)$ to prove the latter.

φ_0:	$\neg\neg(\varphi \vee \neg\varphi) \to \varphi \vee \neg\varphi$	instance of $\mathsf{L}_{9^{1}/_{4}}$
φ_1:	$\neg\neg(\varphi \vee \neg\varphi)$	by part (b)
φ_2:	$\varphi \vee \neg\varphi$	from φ_0 and φ_1 by (MP)

For parts (d) and (e) we define a mapping $|\cdot|$ as suggested in the hint by assigning every formula a value in $\{-1, 0, 1\}$. For atomic formulae we may choose any value. For every other formula the value is fixed by induction on formula construction, as illustrated in the following table:

$\lvert\varphi\rvert$	$\lvert\psi\rvert$	$\lvert\neg\varphi\rvert$	$\lvert\varphi\to\psi\rvert$	$\lvert\varphi\wedge\psi\rvert$	$\lvert\varphi\vee\psi\rvert$
-1	-1	1	1	-1	-1
-1	0	1	1	-1	0
-1	1	1	1	-1	1
0	-1	-1	-1	-1	0
0	0	-1	1	0	0
0	1	-1	1	0	1
1	-1	-1	-1	-1	1
1	0	-1	0	0	1
1	1	-1	1	1	1

One can now check with the above table that every instance of any axiom from L_1–L_9 and $\mathsf{L}_{9^3/_4}$ has value 1 under $|\cdot|$. Furthermore, with respect to (MP) we have that if $|\varphi| = 1$ and $|\varphi \to \psi| = 1$, then also $|\psi| = 1$. Thus, for every formula θ with $\{\mathsf{L}_1\text{–}\mathsf{L}_9, \mathsf{L}_{9^3/_4}\} \vdash \theta$ we have $|\theta| = 1$.

Now, if φ is any formula with $|\varphi| = 0$, then $|\neg\varphi| = -1$, $|\neg\neg\varphi| = 1$, and $|\neg\neg\varphi \to \varphi| = 0$, and since $\neg\neg\varphi \to \varphi$ is an instance of $\mathsf{L}_{9^1/_4}$, this implies that $\{\mathsf{L}_1\text{–}\mathsf{L}_9, \mathsf{L}_{9^3/_4}\} \nvdash \mathsf{L}_{9^1/_4}$.

Similarly, $\varphi \vee \neg\varphi$ is an instance of L_0 with $|\varphi \vee \neg\varphi| = 0$ for $|\varphi| = 0$. Hence, $\{\mathsf{L}_1\text{–}\mathsf{L}_9, \mathsf{L}_{9^3/_4}\} \nvdash \mathsf{L}_0$.

2.11 Let $\mathbf{\Phi}_0 := \{\mathsf{L}_0\text{–}\mathsf{L}_9\}$ and $\mathbf{\Phi}_1 := \{\mathsf{L}_1\text{–}\mathsf{L}_9, \mathsf{L}_{9^3/_4}\}$, for simplified notation. We prove each direction separately. First, we verify that every proof of $\neg\neg\varphi$ from $\mathbf{\Phi}_1$ can be transformed in a proof for $\mathbf{\Phi}_0 \vdash \neg\neg\varphi$. First, note that it follows from Exercise 2.9.(a) that $\mathbf{\Phi}_0 \vdash \mathsf{L}_{9^3/_4}$. Now, since $\mathbf{\Phi}_0 \vdash \mathsf{L}_{9^1/_4}$ (see Exercise 2.10.(a)), $\mathbf{\Phi}_0 \vdash \neg\neg\varphi$ then implies $\mathbf{\Phi}_0 \vdash \varphi$.

For the other direction, suppose $\mathbf{\Phi}_0 \vdash \varphi$. Since φ is quantifier-free, we may assume that its proof from $\mathbf{\Phi}_0$ does not involve $(\forall)$. We will now check that $\mathbf{\Phi}_1 \vdash \neg\neg\psi$ for every instance ψ of any axiom scheme in $\mathbf{\Phi}_0$.

First let us prove $\mathbf{\Phi}_1 + \psi \vdash \neg\neg\psi$:

φ_0:	ψ	by assumption
φ_1:	$(\neg\psi \to \psi) \to ((\neg\psi \to \neg\psi) \to \neg\neg\psi)$	instance of $\mathsf{L}_{9^3/_4}$
φ_2:	$\psi \to (\neg\psi \to \psi)$	instance of L_1
φ_3:	$\neg\psi \to \psi$	from φ_2 and φ_0 by (MP)
φ_4:	$(\neg\psi \to \neg\psi) \to \neg\neg\psi$	from φ_1 and φ_3 by (MP)
φ_5:	$\neg\psi \to \neg\psi$	by Example 1.1
φ_6:	$\neg\neg\psi$	from φ_4 and φ_5 by (MP)

Together with EXERCISE 2.10.(b), we thus have $\mathbf{\Phi}_1 \vdash \neg\neg\psi$ for every instance ψ of any axiom scheme in $\mathbf{\Phi}_0$. It remains to check that a rule similar to (MP) holds. For this we claim that

$$\mathbf{\Phi}_1 + \neg\neg(\psi \to \chi) \vdash \neg\neg\psi \to \neg\neg\chi\,. \tag{$**$}$$

In order to show this, we prove that $\mathbf{\Phi}_1 + \neg(\psi \wedge \chi) \vdash \psi \to \neg\chi$.

φ_0:	$\neg(\psi \wedge \chi)$	by assumption
φ_1:	$(\psi \to (\neg(\psi \wedge \chi) \to \neg\chi)) \to ((\psi \to \neg(\psi \wedge \chi)) \to (\psi \to \neg\chi))$	instance of L_2
φ_2:	$\chi \to (\psi \to (\psi \wedge \chi))$	instance of L_5
φ_3:	$\psi \to (\chi \to (\psi \wedge \chi))$	from φ_2 by (DT)
φ_4:	$\psi \to (\neg(\psi \wedge \chi) \to \neg\chi)$	from φ_3 by ($*$) and (DT)
φ_5:	$(\psi \to \neg(\psi \wedge \chi)) \to (\psi \to \neg\chi)$	from φ_1 and φ_4 by (MP)
φ_6:	$\neg(\psi \wedge \chi) \to (\psi \to \neg(\psi \wedge \chi))$	instance of L_1
φ_7:	$\psi \to \neg(\psi \wedge \chi)$	from φ_6 and φ_0 by (MP)
φ_8:	$\psi \to \neg\chi$	from φ_5 and φ_7 by (MP)

Note that in order to obtain φ_3 from φ_2 we have to apply (DT) multiple times.

Hence, we have $\mathbf{\Phi}_1 + \neg(\neg\neg\psi \wedge \neg\chi) \vdash \neg\neg\psi \to \neg\neg\chi$. In order to obtain ($**$), it suffices to show $\mathbf{\Phi}_1 \vdash (\neg\neg\psi \wedge \neg\chi) \to \neg(\psi \to \chi)$, since then ($**$) follows from ($*$) and (MP). Using (DT) we show instead

$$\mathbf{\Phi}_1 + \neg\neg\psi \wedge \neg\chi \vdash \neg(\psi \to \chi)\,.$$

φ_0:	$\neg\neg\psi \wedge \neg\chi$	by assumption
φ_1:	$(\neg\neg\psi \wedge \neg\chi) \to \neg\neg\psi$	instance of L_3
φ_2:	$\neg\neg\psi$	from φ_1 and φ_0 by (MP)
φ_3:	$(\neg\neg\psi \wedge \neg\chi) \to \neg\chi$	instance of L_4
φ_4:	$\neg\chi$	from φ_3 and φ_0 by (MP)
φ_5:	$((\neg\chi \to \neg\psi) \to \neg\psi) \to (((\neg\chi \to \neg\psi) \to \neg\neg\psi) \to \neg(\neg\chi \to \neg\psi))$	instance of $\mathsf{L}_{93/4}$
φ_6:	$(\neg\chi \to \neg\psi) \to \neg\psi$	from φ_4 by (DT) and (MP)
φ_7:	$((\neg\chi \to \neg\psi) \to \neg\neg\psi) \to \neg(\neg\chi \to \neg\psi)$	from φ_5 and φ_6 by (MP)
φ_8:	$\neg\neg\psi \to ((\neg\chi \to \neg\psi) \to \neg\neg\psi)$	instance of L_1
φ_9:	$(\neg\chi \to \neg\psi) \to \neg\neg\psi$	from φ_8 ad φ_2 by (MP)
φ_{10}:	$\neg(\neg\chi \to \neg\psi)$	from φ_7 and φ_9 by (MP)
φ_{11}:	$((\psi \to \chi) \to (\neg\chi \to \neg\psi)) \to (((\psi \to \chi) \to \neg(\neg\chi \to \neg\psi)) \to \neg(\psi \to \chi))$	instance of $\mathsf{L}_{93/4}$
φ_{12}:	$(\psi \to \chi) \to (\neg\chi \to \neg\psi)$	instance of ($*$)
φ_{13}:	$((\psi \to \chi) \to \neg(\neg\chi \to \neg\psi)) \to \neg(\psi \to \chi)$	from φ_{11} and φ_{12} by (MP)
φ_{14}:	$\neg(\neg\chi \to \neg\psi) \to ((\psi \to \chi) \to \neg(\neg\chi \to \neg\psi))$	instance of L_1
φ_{15}:	$(\psi \to \chi) \to \neg(\neg\chi \to \neg\psi)$	from φ_{14} and φ_{10} by (MP)
φ_{16}:	$\neg(\psi \to \chi)$	from φ_{13} and φ_{15} by (MP)

Hence, we have $\mathbf{\Phi}_1 + \neg\neg(\psi \to \chi) \vdash \neg\neg\psi \to \neg\neg\chi$. So, if χ follows from ψ and $\psi \to \chi$ by (MP) in a proof of $\mathbf{\Phi}_0 \vdash \varphi$ and we have $\mathbf{\Phi}_1 \vdash \neg\neg\psi$, as well as $\mathbf{\Phi}_1 \vdash \neg\neg(\psi \to \chi)$, we can prove $\neg\neg\psi \to \neg\neg\chi$ from $\mathbf{\Phi}_1$ as above and obtain a proof for $\neg\neg\chi$ by (MP). Hence, every proof of φ from $\mathbf{\Phi}_0$ leads to a proof of $\neg\neg\varphi$ from $\mathbf{\Phi}_1$.

Chapter 3

3.0 By Example 1.0 we know $\vdash \exists x(x = x)$. From the Soundness Theorem, in particular Fact 3.8, for any model $\mathbf{M}$ we have $\mathbf{M} \vDash \exists x(x = x)$. Therefore, by the definition of "$\vDash$" there must exist some element in the domain of $\mathbf{M}$, which shows that the domain of $\mathbf{M}$ is non-empty. Thus, it is not necessary to require that the domain of an $\mathscr{L}$-structure to be non-empty.

3.1 We define three domains

$$A_1 := \{a\}, \qquad A_2 := \{a, b\}, \qquad A_3 := \{a, b, c\}$$

and the relations

$$R^{\mathbf{M}_1} = \emptyset, \quad R^{\mathbf{M}_2} = \{\langle a, a\rangle, \langle b, b\rangle, \langle a, b\rangle\}, \quad R^{\mathbf{M}_3} = A_3^2 \setminus \{\langle a, c\rangle, \langle c, a\rangle\}.$$

It can be checked that the corresponding models $\mathbf{M}_1$, $\mathbf{M}_2$ and $\mathbf{M}_3$ satisfy the requirements.
It remains to show that the chosen domains are minimal: By the previous exercise it is clear that no domain can have fewer elements than A_1, and any model with just one element will always satisfy φ_2 and φ_3. Finally, it can be checked that any model $\mathbf{M}$ with domain A_2 such that $\mathbf{M} \vDash \neg\varphi_3$ cannot be a model of φ_1. Therefore, all three domains are minimal.

3.2 Let a_0 be an arbitrary but fixed element of the domain A of $\mathbf{M}'$. We construct a model $\mathbf{M}$ as follows: For each constant symbol $c \in \mathscr{L}$ which does not belong to $\mathscr{L}'$, let $c^{\mathbf{M}} := a_0$. Similarly, for each n-ary function symbol $F \in \mathscr{L}$ which does not belong to $\mathscr{L}'$, let $F^{\mathbf{M}} : A^n \to A$ be such that $F^{\mathbf{M}}$ maps each element of A^n to a_0, and for each n-ary relation symbol $R \in \mathscr{L}$ which does not belong to $\mathscr{L}'$, let $R^{\mathbf{M}} := A^n$. Finally, interpret the symbols in $\mathscr{L}'$ in the $\mathscr{L}$-structure $\mathbf{M}$ as they are interpreted in the $\mathscr{L}'$-structure $\mathbf{M}'$.
We now show that $\mathbf{M} \vDash \mathsf{T}$ by induction on the construction of the $\mathscr{L}'$-sentences $\sigma \in \mathsf{T}$: Let φ be an $\mathscr{L}'$-formula which appears in the construction of some $\sigma \in \mathsf{T}$. If φ is an atomic formula, then, since all symbols contained in φ are interpreted in $\mathbf{M}$ as in $\mathbf{M}'$, we obtain $\mathbf{M} \vDash \varphi$. If $\varphi \equiv \neg\psi$, then $\mathbf{M}' \vDash \varphi$ gives $\mathbf{M}' \nvDash \psi$ and by induction we find $\mathbf{M} \nvDash \psi$ and therefore $\mathbf{M} \vDash \neg\varphi$. The other cases follow similarly.

3.3 The solution to this exercise follows directly from Fact 3.4 which in turn can be proven by induction on the construction of formula just as in the solution to Exercise 3.2.

3.4 Let $\mathbf{M} \vDash \mathsf{DLO}$ be a countable model of DLO with domain $\{d_j : j \in \mathbb{N}\}$, and let $\{q_i : i \in \mathbb{N}\}$ be an enumeration of $\mathbb{Q}$. Furthermore, define the function $g : \mathbb{N} \to \mathbb{N}$ as follows: Let $g(0) := 0$ and for each $i \in \mathbb{N}$, let $g(i+1)$ be the least integer $j \in \mathbb{N}$ which is different from $g(0), \ldots, g(i)$, such that for all $k \leq i$ we have

$$q_k < q_{i+1} \;\leftrightarrow\; d_{g(k)} < d_j \,.$$

Now, let $\bar{g}$ map the rationals to the domain of $\mathbf{M}$ by stipulating $\bar{g}(q_i) = d_{g(i)}$. To show that $\mathbf{M}$ is isomorphic to $(\mathbb{Q}, <)$, it is enough to show that $\bar{g}$ is an order-preserving bijection. Since by definition of g, for all $i, j \in \mathbb{N}$ we have $q_i < q_j \;\leftrightarrow\; d_{g(i)} < d_{g(j)}$, the function $\bar{g}$ is an order-preserving injection. So, it remains to show that $\bar{g}$ is surjective which in turn can be shown by proving that g is surjective: Let j_0 be a positive integer and assume that for all $j < j_0$ there is a $k_j \in \mathbb{N}$ such that $g(k_j) = j$. Now, let $i_0 \in \mathbb{N}$ be the least integer such for all $j < j_0$ we have

$$q_{k_j} < q_{i_0} \;\leftrightarrow\; d_j < d_{j_0} \,.$$

Then, $g(i_0) = j_0$, which implies that g is surjective.

3.5 (a) Let $\varphi_{11}, \varphi_{12}$ and φ_{13} be instances of L_{11}-L_{13}, let j be an arbitrary assignment and let $\mathbf{I} = (\mathbf{M}, j)$ be an interpretation with domain A.

We first consider L_{11}: By definition we have

$$\begin{aligned} \mathbf{I} \vDash \varphi_{11} \;&\Longleftrightarrow\; \text{IF } \mathbf{I} \vDash \varphi(\tau) \text{ THEN } \mathbf{I} \vDash \exists\nu\varphi(\nu)\,, \\ \mathbf{I} \vDash \exists\nu\varphi(\nu) \;&\Longleftrightarrow\; \text{THERE EXISTS } a \text{ IN } A\colon \mathbf{I}^a_\nu \vDash \varphi(\nu)\,, \end{aligned}$$

and by Fact 3.0.(b) we have $\mathbf{I}^{\mathbf{I}(\tau)}_{\nu} \vDash \varphi(\nu)$ if and only if $\mathbf{I} \vDash \varphi(\tau)$. Thus, for $a = \mathbf{I}(\tau)$, this shows that L_{11} is valid in $\mathbf{M}$.

For L_{12}, notice that by definition, $\mathbf{I} \vDash \varphi_{12}$ can be stated as

$$\begin{aligned} \mathbf{I} \vDash \varphi_{12} \;&\Longleftrightarrow\; \text{IF FOR ALL } a \text{ IN } A\colon \mathbf{I}^a_\nu \vDash (\psi \to \varphi(\nu)) \\ &\qquad \text{THEN } \mathbf{I} \vDash (\psi \to \forall\nu\varphi(\nu)) \\ &\Longleftrightarrow\; \text{IF } \underbrace{(\text{FOR ALL } a \text{ IN } A\colon \text{ IF } \mathbf{I}^a_\nu \vDash \psi \text{ THEN } \mathbf{I}^a_\nu \vDash \varphi(\nu))}_{(*)} \\ &\qquad \text{THEN } (\text{IF } \mathbf{I} \vDash \psi \text{ THEN FOR ALL } a \text{ IN } A\colon \mathbf{I}^a_\nu \vDash \varphi(\nu)) \end{aligned}$$

By using Fact 3.0 (a) we see that $\mathbf{I}^a_\nu \vDash \psi$ holds if and only if $\mathbf{I} \vDash \psi$. Assume that $(*)$ holds and that $\mathbf{I} \vDash \psi$. Then for all $a \in A$ we obtain $\mathbf{I}^a_\nu \vDash \varphi(\nu)$, which shows that L_{12} is valid in $\mathbf{M}$.

For L_{13}, notice that $\mathbf{I} \vDash \varphi_{13}$ can be stated as

$$\begin{aligned}\mathbf{I} \vDash \varphi_{13} \iff & \text{ IF FOR ALL } a \text{ IN } A\colon \mathbf{I}^{\,a}_{\nu} \vDash (\varphi(\nu) \to \psi) \\ & \text{ THEN } \mathbf{I} \vDash (\exists \nu \varphi(\nu) \to \psi)\end{aligned}$$

$$\begin{aligned}\iff & \text{ IF } \big(\text{FOR ALL } a \text{ IN } A\colon \mathbf{I}^{\,a}_{\nu} \vDash \varphi(\nu) \text{ THEN } \mathbf{I}^{\,a}_{\nu} \vDash \psi\big) \\ & \text{ THEN } \big(\text{ IF THERE EXISTS } a \text{ IN } A\colon \mathbf{I}^{\,a}_{\nu} \vDash \varphi(\nu) \\ & \qquad\qquad \text{THEN } \mathbf{I} \vDash \psi\big)\end{aligned}$$

By similar arguments as above we obtain that L_{13} is valid in $\mathbf{M}$.

(b) Let $\mathbf{I}$ and A be as above and let $\varphi_{14}, \varphi_{15}$ and φ_{16} be instances of the logical axioms L_{14}-L_{16}.

By definition we have

$$\mathbf{I} \vDash \varphi_{14} \iff \mathbf{I}(\tau) \text{ IS THE SAME OBJECT AS } \mathbf{I}(\tau)$$

which is obviously true and therefore L_{14} is valid in $\mathbf{M}$.

For L_{15}, by definition $\mathbf{I} \vDash \varphi_{15}$ can be stated as:

$$\begin{aligned}\mathbf{I} \vDash \varphi_{15} \iff & \text{ IF } \mathbf{I} \vDash (\tau_1 = \tau_1' \wedge \ldots \wedge \tau_n = \tau_n') \\ & \text{ THEN } \mathbf{I} \vDash R(\tau_1, \ldots, \tau_n) \to R(\tau_1', \ldots, \tau_n')\end{aligned}$$

$$\begin{aligned}\iff & \text{ IF } \big(\mathbf{I} \vDash \tau_1 = \tau_1' \text{ AND } \ldots \text{ AND } \mathbf{I} \vDash \tau_n = \tau_n'\big) \\ & \text{ THEN } \big(\text{IF } \mathbf{I} \vDash R(\tau_1, \ldots, \tau_n) \text{ THEN } \mathbf{I} \vDash R(\tau_1', \ldots, \tau_n')\big)\end{aligned}$$

Now, if $\langle \mathbf{I}(\tau_1), \ldots, \mathbf{I}(\tau_n)\rangle$ belongs to $R^{\mathbf{M}}$ and if for all i such that $1 \leq i \leq n$ we have that $\mathbf{I}(\tau_i)$ is the same object as $\mathbf{I}(\tau_i')$, then also $\langle \mathbf{I}(\tau_1'), \ldots, \mathbf{I}(\tau_n')\rangle$ belongs to $R^{\mathbf{M}}$, which shows that L_{15} is valid in $\mathbf{M}$.

Finally, for L_{16} we have:

$$\begin{aligned}\mathbf{I} \vDash \varphi_{16} \iff & \text{ IF } \mathbf{I} \vDash (\tau_1 = \tau_1' \wedge \ldots \wedge \tau_n = \tau_n') \\ & \text{ THEN } \mathbf{I} \vDash F(\tau_1, \ldots, \tau_n) = F(\tau_1', \ldots, \tau_n')\end{aligned}$$

$$\begin{aligned}\iff & \text{ IF } \big(\mathbf{I} \vDash \tau_1 = \tau_1' \text{ AND } \ldots \text{ AND } \mathbf{I} \vDash \tau_n = \tau_n'\big) \\ & \text{ THEN } \mathbf{I}(F(\tau_1, \ldots, \tau_n)) \text{ IS THE} \\ & \qquad \text{SAME OBJECT AS } \mathbf{I}(F(\tau_1', \ldots, \tau_n'))\end{aligned}$$

By definition we have that $\mathbf{I}(F(\tau_1, \ldots, \tau_n))$ is the same object as $F^{\mathbf{M}}(\mathbf{I}(\tau_1), \ldots, \mathbf{I}(\tau_n))$. Now, if we assume that for all i such that $1 \leq i \leq n$, $\mathbf{I}(\tau_i)$ is the same object as $\mathbf{I}(\tau_i')$, then $F^{\mathbf{M}}(\mathbf{I}(\tau_1), \ldots, \mathbf{I}(\tau_n))$

is the same object as $F^{\mathbf{M}}(\mathbf{I}(\tau_1'), \ldots, \mathbf{I}(\tau_n'))$, which implies that

$$\mathbf{I} \vDash F(\tau_1, \ldots, \tau_n) = F(\tau_1', \ldots, \tau_n')$$

and shows that L_{16} is valid in $\mathbf{M}$.

3.6 (a) Let σ be an $\mathscr{L}$-sentence. With THEOREM 2.13 we find an equivalent sentence $\tilde{\sigma}$ satisfying the condition on its variables. Let $\mathbf{M}$ be an arbitrary $\mathscr{L}$-structure and j any assignment. As σ and $\tilde{\sigma}$ are equivalent we have $\vdash \sigma \leftrightarrow \tilde{\sigma}$ and by the SOUNDNESS THEOREM we then obtain $(\mathbf{M}, j) \vDash \sigma \leftrightarrow \tilde{\sigma}$ which is enough to conclude that they are semantically equivalent.

(b) By applying THEOREM 2.14, we can argue as above.

3.7 For part (a), consider the sentence $\sigma :\equiv \forall x \forall y (x = y)$. We construct two models $\mathbf{G}_1$ and $\mathbf{G}_2$ of GT such that $\mathbf{G}_1 \vDash \sigma$ and $\mathbf{G}_2 \vDash \neg\sigma$. Then, $\mathsf{GT} \nvdash \sigma$ and $\mathsf{GT} \nvdash \neg\sigma$, which implies that GT is incomplete.
For $\mathbf{G}_1$ take the domain $A_1 := \{0\}$, let $\mathsf{e}^{\mathbf{G}_1} := 0$ and $\circ^{\mathbf{G}_1}(0,0) := 0$. Then, $\mathbf{G}_1 \vDash \mathsf{GT}$ and $\mathbf{G}_1 \vDash \sigma$.
For $\mathbf{G}_2$ take the domain $A_2 := \{-1, 1\}$ and let $\mathsf{e}^{\mathbf{G}_2} := 1$. Define $\circ^{\mathbf{G}_2}$ like multiplication in $\mathbb{Z}$. Then, $\mathbf{G}_2 \vDash \mathsf{GT}$ and, since $-1 \neq 1$, $\mathbf{G}_2 \vDash \neg\sigma$. This proves part (a).

For part (b), we may use the same sentence σ and the same models $\mathbf{G}_1$ and $\mathbf{G}_2$. One only needs to verify that $\mathbf{G}_1 \vDash \psi$ and $\mathbf{G}_2 \vDash \psi$, which follows from the commutativity of multiplication in $\mathbb{Z}$.

Chapter 4

4.0 The proof of FACT 4.0 makes use of FACT 2.16.(d), which uses L_5. For LEMMA 4.1 we used TAUTOLOGY (F), whose proof uses the axioms L_0, L_1, L_8, L_9. In addition, we used L_2 in the proof of $\vdash \varphi \to \varphi$, which was also used for the proof of TAUTOLOGY (F).
The given proof for PROPOSITION 4.6 makes use of L_3, L_4 and TAUTOLOGY (K), by using THEOREM 1.7, where the proof of TAUTOLOGY (K) uses L_6 and L_7.

4.1 Following the hint, let Λ be the set of all finite lists $\lambda = [\neg\sigma_0, \varrho_1, \ldots, \varrho_n]$ of $\mathscr{L}$-sentences, where for each $1 \leq i \leq n$, either $\varrho_i \equiv \sigma_i$ or $\varrho_i \equiv \neg\sigma_i$. Encode now finite sequences of formulae $s = [\varphi_0, \ldots, \varphi_n]$ by stipulating

$$\#s := \mathbf{3}\#\varphi_n\mathbf{3}\#\varphi_{n-1}\cdots\mathbf{3}\#\varphi_0\,,$$

where each formula is encoded as a string of **0**'s, **1**'s, and **2**'s as it was done for the universal list $\Lambda_{\mathscr{L}}$ of $\mathscr{L}$-sentences. With this encoding of finite sequence of formulae, and with a bijection between these codes and the natural numbers, we can encode each formal proof by a natural number. Finally, let T be the set of all lists $\lambda \in \Lambda$, such that there is no formal proof of an inconsistency from $\mathsf{T} + \lambda$ with a code-number less than the length of λ. In order to obtain an infinite 0-1-tree, to each sequence $\lambda \in T$ we assign a 0-1-sequence by stipulating

$$[\neg\sigma_0, \varrho_1, \ldots, \varrho_n] \;\mapsto\; \langle 1, t_1, \ldots, t_n\rangle\,,$$

where for $1 \leq i \leq n$,

$$t_i := \begin{cases} 0 & \text{if } \varrho_i \equiv \sigma_i, \\ 1 & \text{if } \varrho_i \equiv \neg\sigma_i. \end{cases}$$

Then the set of 0-1-sequence we obtain as images of sequences $\lambda \in T$ is an infinite 0-1-tree with the property that each infinite branch through T corresponds to a maximally consistent set of $\mathscr{L}$-sentences.

4.2 Let $\overline{\mathsf{T}} = \mathbf{Th}(\mathbf{M})$. Note that $\overline{\mathsf{T}}$ is maximally consistent if for every sentence σ we have $\sigma \in \overline{\mathsf{T}}$ or $\neg\,\mathrm{Con}(\overline{\mathsf{T}} + \sigma)$. So, let σ be a sentence with $\sigma \notin \overline{\mathsf{T}}$. We verify that $\neg\,\mathrm{Con}(\overline{\mathsf{T}} + \sigma)$. By construction, $\mathbf{M} \nvDash \sigma$ and thus, $\mathbf{M} \vDash \neg\sigma$, which implies $\neg\sigma \in \overline{\mathsf{T}}$. Now, observe that $\overline{\mathsf{T}} + \sigma \vdash \sigma$ and $\overline{\mathsf{T}} + \sigma \vdash \neg\sigma$, since $\neg\sigma \in \overline{\mathsf{T}}$. Hence, by L_5, $\overline{\mathsf{T}} + \sigma \vdash \sigma \wedge \neg\sigma$, which implies $\neg\,\mathrm{Con}(\overline{\mathsf{T}} + \sigma)$.

Chapter 5

5.0 By contraposition assume that T is not complete. Then there is a sentence σ such that $\mathsf{T} \nvdash \sigma$ and $\mathsf{T} \nvdash \neg\sigma$. In particular, both theories $\mathsf{T} + \sigma$ and $\mathsf{T} + \neg\sigma$ are consistent. Thus, by THEOREM 5.5 we find models $\mathbf{M}_1$ and $\mathbf{M}_2$ such that $\mathbf{M}_1 \models \mathsf{T} + \sigma$ and $\mathbf{M}_2 \models \mathsf{T} + \neg\sigma$. Now, since both models are models of T and since $\mathbf{M}_1 \ncong \mathbf{M}_2$, this shows that T has at least two non-isomorphic models.

Remark: By THEOREM 15.4, the statement also holds for arbitrary signatures $\mathscr{L}$.

5.1 (a) We have to show

(i) $\bigcup_\varphi X_\varphi = \Sigma$, and

(ii) for any $\mathscr{L}$-sentences φ_1 and φ_2, and for every $\mathscr{L}$-structure $\mathbf{M} \in X_{\varphi_1} \cap X_{\varphi_2}$, there is an $\mathscr{L}$-sentence ψ such that

$$\mathbf{M} \in X_\psi \subseteq X_{\varphi_1} \cap X_{\varphi_2}\,.$$

For (i), notice that for any $\mathscr{L}$-sentence φ we have $\Sigma = X_\varphi \cup X_{\neg\varphi}$. For (ii), let $\mathbf{M} \in X_{\varphi_1} \cap X_{\varphi_2}$. Then $\mathbf{M} \models \varphi_1 \wedge \varphi_2$, and therefore, $\mathbf{M} \in X_{\varphi_1 \wedge \varphi_2}$.

(b) Let φ be an $\mathscr{L}$-sentence. Because $\Sigma = X_\varphi \cup X_{\neg\varphi}$ and $X_\varphi \cap X_{\neg\varphi} = \emptyset$, we obtain that $X_\varphi = \Sigma \setminus X_{\neg\varphi}$ is closed. In fact, the basic open sets X_φ are also closed.

(c) Fix any open covering $\bigcup_{i\in I} A_i$ of Σ and assume there is no finite subcovering, i.e., for every finite subset $J \subseteq I$ we have:

$$\bigcup_{j\in J} A_j \neq \Sigma$$

Let Φ be the collection of all $\mathscr{L}$-sentences φ, such that the basic open set X_φ is contained in an open set A_i (for some $i \in I$). By our assumption we have that

$$\bigcup_{\varphi\in\Phi} X_\varphi = \Sigma$$

and for every finite subset $\Psi \subseteq \Phi$ we have $\bigcup_{\varphi\in\Psi} X_\varphi \neq \Sigma$. In other words, for every finite subset $\Psi \subseteq \Phi$ there is an $\mathscr{L}$-structure $\mathbf{M}_{\bar{\Psi}}$, such that

$$\mathbf{M}_{\bar{\Psi}} \in \Sigma \setminus \bigcup_{\varphi\in\Psi} X_\varphi = \bigcap_{\varphi\in\Psi} X_{\neg\varphi}\,,$$

which implies that for each $\varphi \in \Psi$, $\mathbf{M}_{\bar{\Psi}} \models \neg\varphi$. In particular, for every finite subset $\{\varphi_0, \ldots, \varphi_n\} \subseteq \Phi$ we have that $\{\neg\varphi_0, \ldots, \neg\varphi_n\}$

is consistent. Thus, by the Compactness Theorem 2.17, the set $\{\neg\varphi : \varphi \in \Phi\}$ is consistent and by Theorem 5.5 has a model $\mathbf{M}_{\bar{\Phi}}$. Hence, we have

$$\mathbf{M}_{\bar{\Phi}} \in \bigcap_{\varphi\in\Phi} X_{\neg\varphi} = \Sigma \setminus \bigcup_{\varphi\in\Phi} X_{\varphi}\,,$$

which is a contradiction to the fact that $\Sigma = \bigcup_{\varphi\in\Phi} X_{\varphi}$.

5.2 (a) Assume towards a contradiction that there exists an $\mathscr{L}_{\mathsf{DLO}}$-sentence σ, such that $\mathsf{DLO} \nvdash \neg\sigma$ and $\mathsf{DLO} \nvdash \sigma$. Then $\mathsf{DLO} + \sigma$ and $\mathsf{DLO} + \neg\sigma$ are both consistent, and therefore, by Skolem's Paradox 5.6, there are countable models $\mathbf{M}$ and $\mathbf{N}$ such that $\mathbf{M} \models \mathsf{DLO} + \sigma$ and $\mathbf{N} \models \mathsf{DLO} + \neg\sigma$. However, this contradicts the fact that any two countable models of DLO are isomorphic (see Exercise 3.4).

(b) Notice that $(\mathbb{Q}, <)$ and $(\mathbb{R}, <)$ are both models of DLO and are non-isomorphic because of their different cardinality. We claim that they are elementary equivalent: Let σ be an arbitrary $\mathscr{L}$-sentence. By (a) we have either $\mathsf{DLO} \vdash \sigma$ or $\mathsf{DLO} \vdash \neg\sigma$, and therefore, by the Soundness Theorem we have either $\mathbb{Q} \models \sigma$ and $\mathbb{R} \models \sigma$, or $\mathbb{Q} \models \neg\sigma$ and $\mathbb{R} \models \neg\sigma$.

5.3 For each $n \in \mathbb{N}$ define a constant symbol c_n and let

$$\mathscr{L}' := \mathscr{L} \cup \{c_n : n \in \mathbb{N}\}\,.$$

For $n, m \in \mathbb{N}$ with $n < m$ define the $\mathscr{L}$-sentence $\varphi_{n,m} :\equiv c_n \neq c_m$ and let

$$\mathsf{T}' := \mathsf{T} \cup \{\varphi_{n,m} : n, m \in \mathbb{N} \text{ and } n < m\}\,.$$

Since T has arbitrarily large finite models, for each $k \in \mathbb{N}$ there is a model $\mathbf{M}_k \models \mathsf{T}$ such that its domain has at least k elements. Now, by assigning k constants c_n to pairwise different objects in the domain of $\mathbf{M}$, we find a model $\mathbf{M}'_k \models \mathsf{T}$ which is also a model for $\binom{k}{2}$ sentences $\varphi_{n,m}$. In particular, this shows that any finite subset of T' is consistent. We therefore obtain by Compactness Theorem 2.17 and by Theorem 5.5 that T' is consistent and has a model $\mathbf{M}$. By definition of $\varphi_{n,m}$, $\mathbf{M}$ must have an infinite domain, and since $\mathscr{L} \subset \mathscr{L}'$ and $\mathsf{T} \subset \mathsf{T}'$, we have $\mathbf{M} \models \mathsf{T}$.

5.4 Let

$$\underline{n} \equiv \underbrace{\mathsf{1} + \ldots + \mathsf{1}}_{n \text{ times}},$$

where we omit the brackets since by construction $+$ is associative. Now, let $\mathscr{L}'$ be the language $\mathscr{L} \cup \{c\}$ for some constant symbol c, let φ_n be

the formula

$$\mathsf{0} < c \wedge \underline{n} \cdot c < \mathsf{1}\,,$$

and let

$$\mathsf{T}' := \mathsf{T} \cup \left\{\varphi_n : n \in \mathbb{N}\right\}.$$

Now, every finite subset $\mathbf{\Phi}$ of T' is consistent, since if n is the biggest natural number such that $\varphi_n \in \mathbf{\Phi}$, then we can assign c to $\frac{1}{n+1}$ and hence obtain that $(\mathbb{Q}, 0, 1, \frac{1}{n+1}, +, \cdot, <)$ is a model of $\mathbf{\Phi}$. We therefore obtain by the Compactness Theorem 2.17 and by Theorem 5.5 that T' is consistent and has a model $\mathbf{M}$, and $c^{\mathbf{M}}$ is the desired infinitesimal number.

Chapter 6

6.0 A possible axiomatisation for (a) is

GT_0^*: $\forall x \forall y \forall z \big(x \circ (y \circ z) = (x \circ y) \circ z\big)$
GT_1^*: $\exists x \forall y \exists z \big((x \circ y = y) \wedge (z \circ y = x)\big)$

For part (b), let $\psi_{\mathsf{e}}(x) :\equiv \forall y \exists z \big(x \circ y = y \wedge (z \circ y = x)\big)$. We need to check that $\mathsf{GT}^* \vdash \exists! x\, \psi_{\mathsf{e}}(x)$. We will give a semantic proof of this result.
Let $\mathbf{M}$ be an arbitrary model of GT^* and let e_1, e_2 be two witnesses for $\exists x\, \psi_{\mathsf{e}}(x)$ with respect to axiom GT_1^*. Then, for $i \in \{1, 2\}$, we have

$$\mathbf{M} \vDash \forall y (e_i \circ y = y) \tag{1}$$
$$\text{and} \quad \mathbf{M} \vDash \forall y\, \exists z (z \circ y = e_i)\,. \tag{2}$$

So, $\mathbf{M}$ is a model for:

$$\begin{array}{rcll} e_1 & = & e_2 \circ e_1 & \text{from (1) with } i = 2 \text{ and } y = e_1 \\ & = & (z \circ e_1) \circ e_1 & \text{from (2) with } i = 2 \text{ and } y = e_1 \\ & = & z \circ (e_1 \circ e_1) & \text{by } \mathsf{GT}_0^* \\ & = & z \circ e_1 & \text{from (1) with } i = 1 \text{ and } y = e_1 \\ & = & e_2 & \text{from (2) with } i = 2 \text{ and } y = e_1 \end{array}$$

Hence, in $\mathbf{M}$ we have for some z in the domain of $\mathbf{M}$: $\mathbf{M} \vDash \mathsf{GT}^*$ we have $\mathbf{M} \vDash \exists! x \forall y\, \psi_{\mathsf{e}}(x)$, which implies, by Gödel's Completeness Theorem 5.5, $\mathsf{GT}^* \vdash \exists! x\, \psi_{\mathsf{e}}(x)$.

For part (c), let $\psi_{\text{inv}}(x, y) :\equiv y \circ x = \mathsf{e}$. We need to verify that

$$\mathsf{GT}^* \vdash \forall x \exists! y\, \psi_{\text{inv}}(x, y)\,.$$

Similar to the argument above, let $\mathbf{M} \vDash \mathsf{GT}^*$ and $e = \mathsf{e}^{\mathbf{M}} \in \mathbf{M}$ as in part (b). We will first prove the following two results:

$$\mathsf{GT}^* \vdash \forall x (x \circ x = x \to x = \mathsf{e}) \tag{3}$$
$$\mathsf{GT}^* \vdash \forall x \forall y (y \circ x = \mathsf{e} \to x \circ y = \mathsf{e}) \tag{4}$$

For (3), let a be an arbirary element of the domain of $\mathbf{M}$ for which we have $\mathbf{M} \vDash a \circ a = a$. Then $\mathbf{M}$ is a model for:

$$\begin{array}{rcll} a & = & e \circ a & \text{by } \mathsf{GT}_1^* \\ & = & (b \circ a) \circ a & \text{by } \mathsf{GT}_1^*, \text{ for some } b \text{ with } b \circ a = e \\ & = & b \circ (a \circ a) & \text{by } \mathsf{GT}_0^* \\ & = & b \circ a & \text{by assumption} \\ & = & e & \text{by } \mathsf{GT}_1^* \end{array}$$

Thus, since a was arbitrary, $\mathbf{M} \vDash \forall x (x \circ x = x \to x = \mathsf{e})$.

For (4), let b and a be arbitrary elements of the domain of $\mathbf{M}$ for which we have $\mathbf{M} \vDash b \circ a = e$. Then $\mathbf{M}$ is a model for:

$$\begin{aligned}(a \circ b) \circ (a \circ b) &= a \circ ((b \circ a) \circ b) && \text{by } \mathsf{GT}_0^* \\ &= a \circ (e \circ b) && \text{by assumption} \\ &= a \circ b && \text{by } \mathsf{GT}_1^*\end{aligned}$$

Thus, by (3), $\mathbf{M} \vDash a \circ b = e$, which proves (4). Now, let a, b, c in the domain of $\mathbf{M}$ such that

$$\mathbf{M} \vDash b \circ a = e \tag{5}$$

$$\text{and} \quad \mathbf{M} \vDash c \circ a = e\,. \tag{6}$$

Then $\mathbf{M}$ is a model for:

$$\begin{aligned}b &= e \circ b && \text{by } \mathsf{GT}_1^* \\ &= (c \circ a) \circ b && \text{from (6)} \\ &= c \circ (a \circ b) && \text{by } \mathsf{GT}_0^* \\ &= c \circ e && \text{from (5) and (4)} \\ &= c \circ (a \circ c) && \text{from (6) and (4)} \\ &= (c \circ a) \circ c && \text{by } \mathsf{GT}_0^* \\ &= e \circ c && \text{from (6)} \\ &= c && \text{by } \mathsf{GT}_1^*\end{aligned}$$

Thus, since $\mathbf{M}$ was arbitrary, this proves $\mathsf{GT}^* \vdash \forall x \exists! y\, \psi_{\mathrm{inv}}(x, y)$.

6.1 Let $\mathscr{L}$ be a signature, let $c \in \mathscr{L}$ be a constant symbol, and let T be an $\mathscr{L}$-theory. In a first step, we replace the constant symbol $c \in \mathscr{L}$ by a unary relation symbol R_c and denote the corresponding signature by $\mathscr{L}^*$. In a second step, for each $\mathscr{L}$-formula φ we choose a variable ν not occurring in φ and replace φ by the formula

$$\psi_\varphi :\equiv \begin{cases} \exists \nu \big(\varphi(c/\nu) \wedge R_c(\nu)\big) & \text{or} \\ \varphi, \end{cases}$$

depending on whether c occurs in φ or not, where $\varphi(c/\nu)$ is the formula obtained from φ by replacing every instance of c by ν. This way, the theory T transforms to an $\mathscr{L}^*$-theory (i.e., to a theory which does not involve the constant symbol c). To this theory we add the sentence

$$\exists! x \big(R_c(x)\big)$$

and denote the resulting theory by T^*.
Let $\mathbf{M}$ be a an $\mathscr{L}$-structure with domain A, such that $\mathbf{M} \vDash \mathsf{T}$. Now, we extend $\mathbf{M}$ to an $\mathscr{L}^*$-structure $\mathbf{M}^*$ with the same domain by stipulating

$$R_c^{\mathbf{M}^*} := \{c^{\mathbf{M}}\}$$

where $c^{\mathbf{M}^*}$ is undefined in $\mathbf{M}^*$. Then $\mathbf{M}^* \vDash \mathsf{T}^*$ and for each $\mathscr{L}$-formula φ we have $\mathbf{M} \vDash \varphi$ if and only if $\mathbf{M}^* \vDash \psi_\varphi$. Thus, since the constant symbol c was arbitrary, we may replace each constant symbol $c \in \mathscr{L}$

by its corresponding relation R_c, which shows that constant symbols are dispensable.

Let $\mathscr{L}$ be a signature, let $f \in \mathscr{L}$ be an n-ary function symbol, and let T be an $\mathscr{L}$-theory. As above, we first replace the function symbol $f \in \mathscr{L}$ by an $(n+1)$-ary relation symbol R_f and denote the corresponding signature by $\mathscr{L}^*$. Then, for each $\mathscr{L}$-formula φ we choose a variable ν not occurring in φ and, for terms $\tau_1, \ldots, \tau_n$, replace φ by the formula

$$\psi_\varphi :\equiv \begin{cases} \exists \nu \big(\varphi(f(\tau_1, \ldots, \tau_n)/\nu) \wedge R_f(\tau_1, \ldots, \tau_n, \nu) \big) & \text{or} \\ \varphi, \end{cases}$$

depending on whether $f(\tau_1, \ldots, \tau_n)$ occurs in φ or not. This way, the theory T transforms to an $\mathscr{L}^*$-theory. To this theory we add the sentence

$$\forall \nu_1 \cdots \forall \nu_n \, \exists! y \, \big(R_f(\nu_1, \ldots, \nu_n, y) \big)$$

and denote the resulting theory by T^*.
Let $\mathbf{M}$ be a an $\mathscr{L}$-structure with domain A, such that $\mathbf{M} \vDash \mathsf{T}$. Now, by induction on the complexity of terms, we extend $\mathbf{M}$ to an $\mathscr{L}^*$-structure $\mathbf{M}^*$ with the same domain by stipulating

$$\big(\tau_1^{\mathbf{M}^*}, \ldots, \tau_n^{\mathbf{M}^*}, \tau_0^{\mathbf{M}^*} \big) \in R_f^{\mathbf{M}^*} \iff f^{\mathbf{M}} \big(\tau_1^{\mathbf{M}^*}, \ldots, \tau_n^{\mathbf{M}^*} \big) = \tau_0^{\mathbf{M}^*}$$

where $f^{\mathbf{M}^*}$ is undefined in $\mathbf{M}^*$. By construction, $\mathbf{M}^* \vDash \mathsf{T}^*$ and for each $\mathscr{L}$-formula φ we have $\mathbf{M} \vDash \varphi$ if and only if $\mathbf{M}^* \vDash \psi_\varphi$. Thus, since the function symbol f was arbitrary, we may replace each function symbol $f \in \mathscr{L}$ by its corresponding relation R_f, which shows that function symbols are dispensable.

6.2 For part (a) we use the axiomatization obtained in the solution of Exercise 6.0.(a), together with what we have found in Exercise 6.1. This leads to the following:

$$\mathsf{GT}_{-1}^R : \ \forall x \forall y \exists! z R(x, y, z)$$

$$\mathsf{GT}_0^R : \ \forall x \forall y \forall z \exists v_0 \exists v_1 \exists v_2 \exists v_3 \big(R(x, y, v_0) \wedge R(y, z, v_1) \wedge R(x, v_1, v_2) \wedge R(v_0, z, v_3) \wedge v_2 = v_3 \big)$$

$$\mathsf{GT}_1^R : \ \exists x \forall y \exists z \big(R(x, y, y) \wedge R(z, y, x) \big)$$

For part (b), we define the relation

$$R_{\mathsf{e}}(x) :\iff \forall y \exists z \big(R(x, y, y) \wedge R(z, y, x) \big) .$$

By Exercise 6.1, the axiomatization above is equivalent to the one given in Exercise 6.0.(a). Thus, the same argument as in Exercise 6.0.(b) shows $\mathsf{GT}^R \vdash \exists! x\, R_e(x)$.

Similarly, for part (c) we define

$$R_{\mathrm{inv}}(x,y) :\Longleftrightarrow \exists v_0 \big(R_e(v_0) \wedge R(y,x,v_0)\big)$$

and obtain $\mathsf{GT}^R \vdash \forall x \exists! y\, R_{\mathrm{inv}}(x,y)$ by Exercises 6.1 and 6.0.(c).

6.3 Let $\mathbf{M} = (\mathbb{Z}, +, 0)$. Note that $a \mapsto -a$ is an automorphism of $\mathbf{M}$ and hence, $\mathbf{M} \vDash \varphi(a)$ if and only if $\mathbf{M} \vDash \varphi(-a)$ for any integer a. This holdsbecause due to $(-a)+(-b) = -(a+b)$ any statement about addition does not change the truth value if we replace a by $-a$. Now, assume towards a contradiction that there is an $\mathscr{L}$-formula $\psi_<$ such that $\mathsf{T}^* \vdash \varphi$, where

$$\varphi \equiv \forall x \forall y (x < y \leftrightarrow \psi_<(x,y))\,.$$

Consider $\mathbf{M}^* = (\mathbb{Z}, +, 0, <)$. Then $\mathbf{M}^* \vDash 0 < 1$ and hence, $\mathbf{M}^* \vDash \psi_<(0,1)$. Since $\psi_<$ is an $\mathscr{L}$-Formula, we get that also $\mathbf{M} \vDash \psi_<(0,1)$, and by our observation above we obtain $\mathbf{M} \vDash \psi_<(0,-1)$, which is obviously a contradiction.

Chapter 7

7.0 We proceed as in the proof of THEOREM 7.2: Let σ and τ be FINITE strings of the form $\mathbf{s}\cdots\mathbf{s}$, then

$$\sigma\mathbf{0} \cdot^{\mathbb{N}} \mathsf{s}^{\mathbb{N}}\tau\mathbf{0} \equiv \sigma\cdots\sigma\mathbf{0} \equiv \sigma\cdots\sigma\sigma\mathbf{0} \equiv \sigma\cdots\sigma\mathbf{0} +^{\mathbb{N}} \sigma\mathbf{0} \equiv (\sigma\mathbf{0} \cdot^{\mathbb{N}} \tau\mathbf{0}) +^{\mathbb{N}} \sigma\mathbf{0}.$$
$$\underbrace{\uparrow\cdots\uparrow}_{\mathsf{s}\tau}\ \underbrace{\mathbf{s}\cdots\mathbf{s}}_{\mathsf{s}\tau}\quad \underbrace{\uparrow\cdots\uparrow}_{\tau}\ \underbrace{\mathbf{s}\cdots\mathbf{s}}_{\tau}\quad \underbrace{\uparrow\cdots\uparrow}_{\tau}\ \underbrace{\mathbf{s}\cdots\mathbf{s}}_{\tau}$$

7.1 We first show that $\{\mathsf{PA}_1\text{–}\mathsf{PA}_6\} \nvdash \mathsf{PA}_0$ by constructing a model for PA_1–PA_6 in which PA_0 does not hold: Consider the structure $\mathbf{M}$ with domain $\{0\}$ and with $\mathsf{s}^{\mathbf{M}}, +^{\mathbf{M}}, \cdot^{\mathbf{M}}$ (necessarily) defined as the zero-maps. Then clearly, $\mathbf{M} \vDash \mathsf{PA}_1\text{–}\mathsf{PA}_6$ and $\mathbf{M} \nvDash \mathsf{PA}_0$.

Similarly, towards a model for PA_0 and PA_2–PA_6 let $\mathbf{N}$ be a structure with domain $\{0,1\}$ and define $\mathsf{0}^{\mathbf{N}} := 0$, $n \cdot^{\mathbf{N}} m := n \cdot m$, $n +^{\mathbf{N}} m := \min\{n+m, 1\}$, and $\mathsf{s}^{\mathbf{N}}n := 1$ for $n, m \in \{0,1\}$, where $+$ and $\cdot$ are the standard operations in $\mathbb{N}$. Then $\mathbf{N} \vDash \mathsf{PA}_0$ and $\mathbf{N} \vDash \mathsf{PA}_2$–$\mathsf{PA}_6$, but $\mathbf{N} \nvDash \mathsf{PA}_1$ since $\mathsf{s}^{\mathbf{N}}0 = \mathsf{s}^{\mathbf{N}}1 = 1$ but $1 \neq 0$.

7.2 Let $\mathbb{P}$ be the set of all standard prime numbers and for $p \in \mathbb{P}$ define $\varphi_p(x) :\equiv \underline{p} \mid x$, $\mathscr{L} := \mathscr{L}_{\mathsf{PA}} \cup \{\mathsf{c}\}$ and $\mathsf{T} := \mathsf{PA} + \{\varphi_p(\mathsf{c}) : p \in \mathbb{P}\}$. Note that divisibility can be defined in PA as will be discussed in Chapter 8. By suitably assigning c to an element of $\mathbb{N}$, we see that $\mathbb{N}$ is a model of any finite subset of T. Hence, by THEOREM 2.17 and THEOREM 5.5, T is consistent and therefore has a model $\mathbf{N}$, and by construction, every standard prime number divides $\mathsf{c}^{\mathbf{N}}$.

7.3 Following the hint, let c be a constant different from $\mathsf{0}$, let $\mathbb{P} \subseteq \mathbb{N}$ be the set of all prime numbers and for $p \neq q \in \mathbb{P}$ define $\varphi_{p,q} :\equiv p \mid \mathsf{c} \wedge q \nmid \mathsf{c}$. For each $S \subseteq \mathbb{P}$, let Φ_S be the collection of all formulae $\varphi_{p,q}$ such that $p \in S$ and $q \notin S$.

Then, for each $S \subseteq \mathbb{P}$ we can choose a $\mathsf{c}^{\mathbb{N}}$ in the domain of $\mathbb{N}$ such that $\mathbb{N}$ is a model of any finite subset of $\mathsf{PA}+\Phi_S$. With THEOREM 2.17 we obtain that $\mathsf{PA}+\Phi_S$ is consistent and by THEOREM 5.6 it has a countable model $\mathbf{N}_S$. Now, note that for every $S_0 \subseteq \mathbb{P}$ there are at most countably many $S \subseteq \mathbb{P}$ such that $\mathbf{N}_S$ and $\mathbf{N}_{S_0}$ are isomorphic: Otherwise, by construction, that model would have to contain for every such S a minimal element $a_S \in \mathbf{N}_{S_0}$ such that

$$\{p \in \mathbb{P} : p \mid a_S\} = S.$$

This implies that for any distinct $S_1, S_2 \subseteq \mathbb{P}$ we have $a_{S_1} \neq a_{S_2}$, and since the power-set of $\mathbb{P}$ is uncountable by CANTOR'S THEOREM 13.8, $\mathbf{N}_{S_0}$ would be uncountable. Hence, each model $\mathbf{N}_{S_0}$ is isomorphic to at most countably many models of the form $\mathbf{N}_S$. Moreover, since the

countable union of countable sets is countable by PROPOSITION 13.10, there are uncountably many pairwise non-isomorphic countable models $\mathbf{N}_S$ of PA.

7.4 Let $\mathbf{M}, M$ and φ be as in the statement. Assume towards a contradiction that for every non-standard $a \in M$ there is an $\tilde{a} < a$ in M such that $\mathbf{M} \nvDash \varphi(\tilde{a})$.

Define $\sigma(z) :\equiv \forall x(x < z \to \varphi(x))$. Notice that if $y \in \mathbb{N}$, then any $x < y$ is also an element of $\mathbb{N}$. Obviously, $\mathbf{M} \vDash \sigma(\mathbf{0})$.

In order to apply PA_6, we claim that if $\mathbf{M} \vDash \sigma(a_0)$ for some $a_0 \in M$, then also $\mathbf{M} \vDash \sigma(\mathsf{s}a_0)$. For this, it is enough to prove that under the assumption $\mathbf{M} \vDash \sigma(a_0)$ we have $\mathbf{M} \vDash \varphi(a_0)$: If this is not the case, then, since $\mathbf{M} \vDash \varphi(\underline{n})$ for all $n \in \mathbb{N}$, a_0 is non-standard and we can use the initial assumption to find $\tilde{a} < a_0$ such that $\mathbf{M} \nvDash \varphi(\tilde{a})$. However, this is not possible as $\mathbf{M} \vDash \forall x(x < a_0 \to \varphi(x))$ and $\tilde{a} < a_0$. Thus we have shown $\mathbf{M} \vDash \varphi(a_0)$, which proves the claim.

Now, by PA_6 and the SOUNDNESS THEOREM we obtain $\mathbf{M} \vDash \forall z(\sigma(z))$. In particular, we have $\mathbf{M} \vDash \sigma(a)$ where $a \in M$ is an arbitrary non-standard element (which exists as $\mathbf{M}$ is non-standard). This contradicts our assumption and completes the proof.

7.5 Assume that such a relation standard(x), or st(x) for short, exists, introduced by a language extension of $\mathscr{L}_{\mathsf{PA}}$. Define a new language $\mathscr{L}^* := \mathscr{L}_{\mathsf{PA}} \cup \{\mathrm{st}, \mathsf{c}\}$, where c is a new constant symbol. Now, for every $n \in \mathbb{N}$ let

$$\varphi_n(x) :\equiv x > \underline{n} \quad \text{and} \quad \psi_n :\equiv \mathrm{st}(\mathsf{c}) \wedge \varphi_n(\mathsf{c})$$

and let $\mathsf{T}^* := \mathsf{PA} \cup \{\psi_n : n \in \mathbb{N}\}$. By suitably assigning the constant symbol c, we see that $\mathbb{N}$ is a model for every finite subset of T^*. Thus, by THEOREM 2.17 and THEOREM 5.5, the $\mathscr{L}^*$-theory T^* is consistent and has a model $\mathbf{M}$.

Since $\mathbf{M} \vDash \mathrm{st}(\mathsf{c})$, by definition of st we have that $\mathsf{c}^{\mathbf{M}} = \underline{n}$ for some $n \in \mathbb{N}$, but because $\mathbf{M} \vDash \varphi_n(\mathsf{c})$, we obtain a contradiction.

7.6 We first show that the set $\{\mathbb{Z}_c : c \in M \text{ is non-standard}\}$ together with the binary relation "$\prec$" satisfies the axioms $\mathsf{DLO}_0 - -\mathsf{DLO}_4$ of dense linearly ordered sets without endpoints:

DLO_0: This is clear since $\mathbb{Z}_c = \mathbb{Z}_d$ implies $\mathbb{Z}_c \not\prec \mathbb{Z}_d$

DLO_1: By definition, we have that $\mathbb{Z}_c \prec \mathbb{Z}_d$ implies $c < d$. Hence, for all non-standard $b, c, d \in M$ we have that $\mathbb{Z}_b \prec \mathbb{Z}_c \wedge \mathbb{Z}_c \prec \mathbb{Z}_d$ implies $b < c < d$. In particular we have $b < d$ (since "$<$" is transitive), and since $\mathbb{Z}_b \neq \mathbb{Z}_d$, we obtain $\mathbb{Z}_b \prec \mathbb{Z}_d$.

DLO_2: Notice that for any non-standard $c, d \in M$, $\mathbb{Z}_c \neq \mathbb{Z}_d$ implies either $c < d$ or $d < c$, but not both. In the former case we have $\mathbb{Z}_c \prec \mathbb{Z}_d$, and in the latter case we have $\mathbb{Z}_d \prec \mathbb{Z}_c$.

DLO_3: Let $\mathbb{Z}_c \prec Z_d$, we claim that then $\mathbb{Z}_c \prec \mathbb{Z}_{\frac{c+d}{2}} \prec \mathbb{Z}_d$ where $\frac{c+d}{2}$ denotes a number $e \in M$ with $2e = c + d$ or $2e = c + d + 1$, which exists by Exercise 8.1: As $c < \frac{c+d}{2} < d$, we cannot have $\mathbb{Z}_{\frac{c+d}{2}} \prec \mathbb{Z}_c$ or $\mathbb{Z}_d \prec \mathbb{Z}_{\frac{c+d}{2}}$. Now, if $\mathbb{Z}_c = \mathbb{Z}_{\frac{c+d}{2}}$, then there is an $n \in \mathbb{N}$ such that $c + n = \frac{c+d}{2}$, but this implies that $c + 2n \geq d$, which contradicts the fact that $\mathbb{Z}_c \prec \mathbb{Z}_d$. With DLO_2, we therefore have $\mathbb{Z}_c \prec \mathbb{Z}_{\frac{c+d}{2}}$. Similarly, we can show that $\mathbb{Z}_{\frac{c+d}{2}} \prec \mathbb{Z}_d$.

DLO_4: We claim that $\mathbb{Z}_{\frac{c}{2}} \prec \mathbb{Z}_c \prec \mathbb{Z}_{2c}$ using the same notation for fractions as above. Since $\frac{c}{2} < c < 2c$, by DLO_2 it is enough to prove that $\mathbb{Z}_{\frac{c}{2}} \neq \mathbb{Z}_c \neq \mathbb{Z}_{2c}$: If, for example, $\mathbb{Z}_c = \mathbb{Z}_{2c}$, then we find $n \in \mathbb{N}$ with $c + n = 2c$, which implies $c = n$ and contradicts the fact that c is non-standard (i.e., $c \notin \mathbb{N}$). Thus, we have $\mathbb{Z}_c \prec \mathbb{Z}_{2c}$ and similarly we show that $\mathbb{Z}_{\frac{c}{2}} \prec \mathbb{Z}_c$.

For the second part of the exercise recall that the model $\mathbf{M}$ is countable and apply Exercise 3.4 to see that the set $\{\mathbb{Z}_c : c \in M \text{ is non-standard}\}$ with the binary relation "$\prec$" is isomorphic to $(\mathbb{Q}, <)$. Now, since for every non-standard $c \in M$, $\mathbb{Z}_c$ is isomorphic to $\mathbb{Z}$, we get that the ordering structure of $\mathbf{M}$ corresponds to the disjoint union of $\mathbb{N}$ and $\mathbb{Q} \times \mathbb{Z}$ (see also the figure at the end of Chapter 7).

Chapter 8

8.0 We argue by induction on x, i.e., we prove the following two statements:

$$\mathsf{PA} \vdash \forall y \forall z (\mathsf{0} + (y + z) = (\mathsf{0} + y) + z) \tag{1}$$

$$\mathsf{PA} \vdash \forall y \forall z \big((x + (y + z) = (x + y) + z) \to (\mathsf{s}x + (y + z) = (\mathsf{s}x + y) + z) \big) \tag{2}$$

Then, by PA_6, the assertion follows.

For (1) we argue as follows:

φ_0:	$\forall y (y + \mathsf{0} = y)$	instance of PA_2
φ_1:	$y + \mathsf{0} = y$	from φ_0 by L_{10} and (MP)
φ_2:	$y + \mathsf{0} = \mathsf{0} + y$	by Lemma 8.0, L_{10} and (MP)
φ_3:	$\mathsf{0} + y = y$	from φ_1 and φ_2 by L_{15} and (MP)
φ_4:	$(\mathsf{0} + y = y) \to \big(((\mathsf{0} + y) + z) = y + z\big)$	by L_{16} and L_{14}
φ_5:	$(\mathsf{0} + y) + z = y + z$	from φ_4 and φ_3 by (MP)
φ_6:	$(y + z) + \mathsf{0} = \mathsf{0} + (y + z)$	by Lemma 8.0, L_{10} and (MP)
φ_7:	$(y + z) + \mathsf{0} = y + z$	by PA_2, L_{10} and (MP)
φ_8:	$\mathsf{0} + (y + z) = y + z$	from φ_6 and φ_7 by L_{15} and (MP)
φ_9:	$\mathsf{0} + (y + z) = (\mathsf{0} + y) + z$	from φ_5 and φ_8 by L_{15} and (MP)
φ_{10}:	$\forall y \forall z \big(\mathsf{0} + (y + z) = (\mathsf{0} + y) + z\big)$	from φ_9 by $(\forall)$

For (2) we first show that

$$\mathsf{PA} + \{x + (y + z) = (x + y) + z\} \vdash \mathsf{s}x + (y + z) = (\mathsf{s}x + y) + z\,.$$

φ_0:	$\mathsf{s}x + (y + z) = (y + z) + \mathsf{s}x$	by Lemma 8.0, L_{10} and (MP)
φ_1:	$(y + z) + \mathsf{s}x = \mathsf{s}((y + z) + x)$	by PA_3, L_{10} and (MP)
φ_2:	$\mathsf{s}((y + z) + x) = \mathsf{s}(x + (y + z))$	by Lemma 8.0, L_{10}, L_{16} and (MP)
φ_3:	$x + (y + z) = (x + y) + z$	by assumption
φ_4:	$\mathsf{s}(x + (y + z)) = \mathsf{s}((x + y) + z)$	from φ_3 by L_{16} and (MP)
φ_5:	$\mathsf{s}((x + y) + z) = \mathsf{s}(z + (x + y))$	by Lemma 8.0, L_{10}, L_{16} and (MP)
φ_6:	$\mathsf{s}x + (y + z) = \mathsf{s}(z + (x + y))$	from $\varphi_0, \varphi_1, \varphi_2, \varphi_4, \varphi_5$ by Taut. (N.1)
φ_7:	$\mathsf{s}(z + (x + y)) = z + \mathsf{s}(x + y)$	by PA_3, L_{10} and (MP)
φ_8:	$z + \mathsf{s}(x + y) = \mathsf{s}(x + y) + z$	by Lemma 8.0, L_{10} and (MP)
φ_9:	$x + y = y + x$	by Lemma 8.0, L_{10} and (MP)
φ_{10}:	$\mathsf{s}(y + x) = y + \mathsf{s}x$	by PA_3
φ_{11}:	$\mathsf{s}(x + y) + z = (y + \mathsf{s}x) + z$	from φ_9 and φ_{10} by L_{15}, L_{16} and (MP)
φ_{12}:	$(y + \mathsf{s}x) + z = (\mathsf{s}x + y) + z$	by Lemma 8.0, L_{10}, L_{16} and (MP)
φ_{13}:	$\mathsf{s}x + (y + z) = (\mathsf{s}x + y) + z$	from $\varphi_6, \varphi_7, \varphi_8, \varphi_{11}, \varphi_{12}$ by Taut. (N.1)

Now, by (DT) we have

$$\mathsf{PA} \vdash (x + (y + z) = (x + y) + z) \to (\mathsf{s}x + (y + z) = (\mathsf{s}x + y) + z)$$

and hence, (2) follows by applying $(\forall)$ twice.

Remark. Alternatively, the induction step can be simplified by proving first $\mathsf{PA} \vdash \forall x \forall y (\mathsf{s}(x + y) = \mathsf{s}x + y)$.

8.1 We define

$$\begin{aligned} \mathrm{even}(x) &:\Longleftrightarrow \exists y(y + y = x) \\ \mathrm{odd}(x) &:\Longleftrightarrow \mathrm{even}(\mathtt{s}x)\,, \end{aligned}$$

which obviously proves the second statement.

For the first statement, the proof is by induction on x. Clearly, since $\mathtt{0} + \mathtt{0} = \mathtt{0}$, we have $\mathrm{even}(\mathtt{0})$. By case distinction, we may first assume $\mathrm{even}(x)$. Then, by taking $\mathtt{s}y$ instead of y, we obtain $\mathrm{even}(x) \to \mathrm{even}(\mathtt{ss}x)$, and since $\mathrm{even}(\mathtt{ss}x) \leftrightarrow \mathrm{odd}(\mathtt{s}x)$, we are done. On the other hand, if we assume $\mathrm{odd}(x)$, then we get $\mathrm{even}(\mathtt{s}x)$ by definition, and therefore we have

$$\mathsf{PA} \vdash \forall x\big(\mathrm{even}(x) \vee \mathrm{odd}(x) \to \mathrm{even}(\mathtt{s}x) \vee \mathrm{odd}(\mathtt{s}x)\big)\,.$$

Hence, by PA_6, the assertion follows.

8.2 First, we prove

$$\mathsf{PA} \vdash \big(x \neq \mathtt{0} \wedge y \neq \mathtt{0} \wedge \exists a \leq y\, \exists b \leq x\, (ax + \mathtt{1} = by)\big) \to \mathrm{coprime}(x, y)\,.$$

In other words, we need to check $\forall z(x \,|\, yz \;\to\; x \,|\, z)$. For this, suppose $x \,|\, yz$. Then, $x \,|\, byz$ and thus, $x \,|\, axz + z$. Since $x \,|\, x$, by applying Lemma 8.10(b) twice we get $x \,|\, axz$ and thus by Lemma 8.10(a) we obtain $x \,|\, z$. The result then follows by (DT) and $(\forall)$.

It remains to show that

$$\mathsf{PA} \vdash \mathrm{coprime}(x, y) \to (x \neq \mathtt{0} \wedge y \neq \mathtt{0} \wedge \exists a \leq y\, \exists b \leq x\, (ax + \mathtt{1} = by))\,.$$

The proof is by strong induction on $x + y$. So, suppose the implication holds for any pair (x', y') with $x' + y' < x + y$ and assume $\mathrm{coprime}(x, y)$. Without loss of generality, let $y \leq x$. By the Principle of Division with Remainder we find $r < y$ and q such that $x = qy + r$ and $\mathrm{coprime}(y, r)$ (by Lemma 8.17). By Lemma 8.11 we have $\mathrm{coprime}(r, y)$. Since $r + y < x + y$, by our induction hypothesis there are $a' \leq y$ and $b' \leq r$ such that $a'r + \mathtt{1} = b'y$. Hence, for $a = a'$ and $b = a'q + b'$ we get

$$ax + \mathtt{1} = a'(qy + r) + \mathtt{1} = a'qy + (a'r + \mathtt{1}) = a'qy + b'y = by.$$

It remains to check that $a \leq y$ and $b \leq x$. By assumption, $a = a' \leq y$, and $b = a'q + b' = qa' + b' \leq qy + r = x$. This completes the proof.

8.3 For part (a), let φ be an $\mathscr{L}_{\mathsf{PA}}$-formula with $\mathrm{free}(\varphi) = \{x\}$ and suppose, by contraposition, that $\exists x\, \neg\varphi(x)$. Let $\psi(x) :\equiv \neg\varphi(x)$. Then, by the Least Number Principle, $\exists x\big(\psi(x) \wedge \forall y < x\, (\neg\psi(y))\big)$, which means that $\exists x\big(\neg\varphi(x) \wedge \forall y < x\, (\varphi(y))\big)$. If $\mathtt{0}$ is a witness for x in this formula, then

$\neg\varphi(0)$ holds and we are done. Otherwise, by LEMMA 8.3, we find y such that $\mathtt{s}y = x$. Since $\varphi(y)$ holds, we have $\neg(\varphi(y) \to \varphi(\mathtt{s}y))$, which proves PA_6 by contraposition.

For part (b), we construct a model $\mathbf{M}$ with domain $A := \mathbb{N} \cup \overline{\mathbb{N}} = \{\mathbf{0}, \mathbf{1}, \mathbf{2}, \ldots\} \cup \{\overline{\mathbf{0}}, \overline{\mathbf{1}}, \overline{\mathbf{2}}, \ldots\}$ consisting of two copies of the standard natural numbers $\mathbb{N}$ which are disjoint, except for one common element $\overline{\mathbf{0}} = \mathbf{0}$. Let $0^{\mathbf{M}} := \mathbf{0}$, define $\mathtt{s}$ in $\mathbf{M}$ as in the standard model and define $\mathtt{s}\overline{n} := \overline{\mathtt{s}n}$ for every $n \in \mathbb{N}$. Similarly, take the standard addition on $\mathbb{N}$ and use the following rules if summands in $\overline{\mathbb{N}}$ are involved:

$$\overline{n} + \overline{m} := \overline{n + m}$$
$$n + \overline{m} := \overline{n + m}$$
$$\overline{n} + m := \overline{n + m}$$

The resulting partial order $<$ in $\mathbf{M}$ then has the property that $n < m$ is equivalent both to $\overline{n} < \overline{m}$ and $n < \overline{m}$.
Regarding multiplication in $\mathbf{M}$ we proceed similarly, by taking standard multiplication on $\mathbb{N}$ together with the following rules:

$$\overline{n} \cdot \overline{m} := \overline{n \cdot m}$$
$$n \cdot \overline{m} := \overline{n \cdot m}$$
$$\overline{n} \cdot m := \overline{n \cdot m}$$

One can check that this gives a model $\mathbf{M}$ satisfying PA_0–PA_5. It remains to prove that the LEAST NUMBER PRINCIPLE holds in $\mathbf{M}$ and that PA_6 fails in $\mathbf{M}$. For the former statement, suppose $a \in A$ satisfies $\mathbf{M} \vDash \varphi(a)$ for some formula φ. If there is such an a which is a standard natural number we may use the LEAST NUMBER PRINCIPLE in the standard model and find an appropriate $n \in \mathbb{N}$. Since $\mathbf{M} \vDash \neg(\overline{m} < n)$ for every $\overline{m} \in \overline{\mathbb{N}}$, the LEAST NUMBER PRINCIPLE holds in this case. Otherwise, $\mathbf{M} \vDash \neg\varphi(m)$ for every $m \in \mathbb{N}$. Similar to the first case, we find an $\overline{n} \in \overline{\mathbb{N}}$ with $\varphi(\overline{n})$ and $\neg\varphi(\overline{m})$ for every $\overline{m} < \overline{n}$. Thus, $\neg\varphi(x)$ holds for every $x < \overline{n}$, which proves the LEAST NUMBER PRINCIPLE also in this case.

To show that $\mathbf{M} \nvDash \mathsf{PA}_6$, by part (a) it suffices to prove that $\mathbf{M} \nvDash \forall x(x = 0 \vee \exists y(x = \mathtt{s}y))$ (i.e., LEMMA 8.3 does not hold in $\mathbf{M}$). But $\overline{\mathbf{1}}$ clearly is neither $\mathbf{0}$ nor has a predecessor, so the claim follows.

8.4 Assume $\varphi(1) \wedge \forall x\big(\varphi(x) \to \varphi(2x) \wedge \varphi(x-1)\big)$. This implies $\varphi(1) \to \varphi(1-1)$ and hence, $\varphi(0)$. We now argue by strong induction on x: Let $x > 1$ and assume $\varphi(y)$ holds for every $y < x$. By EXERCISE 8.1, we may distinguish the cases $\mathrm{even}(x)$ and $\mathrm{odd}(x)$. So, suppose first $\mathrm{even}(x)$, i.e., $\exists z(2z = x)$. Since $x > 1$, we have $z < x$ and the assertion follows from our assumption $\varphi(z) \to \varphi(2z)$.

For the other case, assuming $\mathrm{odd}(x)$, we know that $\mathtt{s}x$ is even and $\mathtt{s}x > \mathtt{3}$. Thus, there is a z with $\mathtt{2}z = \mathtt{s}x$, and $\varphi(\mathtt{s}x)$ follows as described in the previous case. Hence, with $\varphi(x') \to \varphi(x' - \mathtt{1})$ for all $x' < x$ we obtain $\varphi(x)$, and by the Strong Induction Principle 8.14 we obtain $\forall x \varphi(x)$.

Chapter 9

9.0 We use metainduction on n. The case when $n \equiv \mathbf{0}$ follows from PA_4. Assume $\mathsf{PA} \vdash \underline{m} \cdot \underline{n} = \underline{m \cdot^{\mathbb{N}} n}$. Then, using N_0, PA_5, N_1 and Theorem 7.2, the following equations are provable in PA:

$$\underline{m} \cdot \mathbf{s}\underline{n} = \underline{m} \cdot \mathtt{s}\underline{n} = \underline{m} \cdot \underline{n} + \underline{m} = \underline{m \cdot^{\mathbb{N}} n} + \underline{m} = \underline{m \cdot^{\mathbb{N}} n +^{\mathbb{N}} m} = \underline{m \cdot^{\mathbb{N}} \mathbf{s}n}$$

9.1 (a) We first claim that

$$\mathsf{PA} \vdash \forall x \geq 2\big(\mathrm{prime}(x) \vee \exists y \geq 2\, \exists z \geq 2\big(yz = x\big)\big)\,.$$

Let $x \geq 1$ be such that $\neg\,\mathrm{prime}(x)$. Thus, we find a y such that $y \mid x$, $y \neq x$, and $y \neq 1$. Then there is a z with $yz = x$ and $z > 1$ because $y \neq x$, which proves the claim.

Assume now that there is an $x \geq 2$ which does not satisfy the statement. By Least Number Principle we assume further that x is minimal with this property, i.e., $\forall y\big(\neg\,\mathrm{prime}(y) \vee y \nmid x\big)$. As $x \mid x$ we see that $\neg\,\mathrm{prime}(x)$ must be true. We apply the initial claim and write $x = yz$ with some $1 < y < x$. By minimality of x, there is y' such that $\mathrm{prime}(y')$ and $y' \mid y$. By Lemma 8.9 we obtain that $y' \mid x_0$, a contradiction.

(b) Let $x \geq 2$ be such that $\mathrm{prime}(x)$ and let y, z be arbitrary with $x \mid yz$ and assume $x \nmid z$. We will show that this implies $x \mid y$. We first prove that x and z are coprime using Proposition 8.18. Let u be such that$u \mid x$ and $u \mid z$. Since x is prime, we have either $u = x$ or $u = 1$. The first case contradicts the assumption $x \nmid z$. Hence, we obtain $u = 1$ and therefore $\mathrm{coprime}(u, z)$. By Bézout's Lemma (see Exercise 8.2) there are a and b such that $ax + 1 = bz$. Hence $axy + y = byz$, and, since $x \mid axy$ and by assumption $x \mid byz$, we get $x \mid y$ by Lemma 8.10. This shows the first direction.

Conversely, let $x \geq 2$ be such that $\neg\,\mathrm{prime}(x)$. Then we find y, z such that $1 \leq y, z < x$ and $yz = x$, and we obviously have neither $x \mid y$ nor $x \mid z$.

9.2 We define

$$n! = y :\Longleftrightarrow \exists t\big(\mathrm{seq}(t) \wedge \mathrm{lh}(t) = \mathtt{s}n \wedge t_0 = 1 \wedge \forall i < n\big(t_{\mathtt{s}i} = \mathtt{s}i \cdot t_i\big) \wedge t_n = y\big)\,.$$

We have to show that for each n there exists a value for $n!$ and that this value is unique. By definition, such a value is clearly unique if it exists.

For the existence we use induction on n: In the base case $n = \mathtt{0}$, the sequence $\langle \mathtt{1} \rangle$ with one element suffices. Assume there is a sequence s of length $\mathtt{s}n$ encoding $n!$. Then the sequence $s^\frown\langle n! \cdot (n+1)\rangle$ encodes $(n+1)!$. Finally, notice that by COROLLARY 9.3.(b) the function is $\mathbb{N}$-conform.

9.3 We start by proving that the least common multiple of two values x and y exists.

CLAIM. $\mathsf{PA} \vdash \forall x \forall y \exists! m \big(x \mid m \wedge y \mid m \wedge \forall z < m \big(x \nmid z \vee y \nmid z\big)\big)$

Proof of Claim. As $x \mid xy$ and $y \mid xy$ holds, we know that a common multiple must exist. By the LEAST NUMBER PRINCIPLE we can choose the smallest non-zero multiple m — which is unique by minimality. $\dashv_{\text{Claim}}$

Therefore, we can define $\operatorname{lcm}(x, y)$ for any two values x and y. Furthermore, any common multiple of x and y is divisible by $\operatorname{lcm}(x, y)$. To see this, let z be a common multiple of x and y. Now, with the PRINCIPLE OF DIVISION WITH REMAINDER we can write $z = qm + r$ with $\mathtt{0} \leq r < m$ and $q \neq \mathtt{0}$ as $m \leq z$. Then, as $r = z - qm$, by LEMMA 8.10 we have $x \mid r$ and $y \mid r$ which contradicts the minimality of m, unless $r = \mathtt{0}$ and $m \mid z$.

Let F be a function which is definable in PA. We define now

$$\operatorname{lcm}_{i<k} F(i) = m :\Longleftrightarrow \forall i < k \big(F(i) \mid m\big) \wedge \forall y < m \, \exists i < k \big(F(i) \nmid y\big)$$

and claim that this definition is functional:

$$\mathsf{PA} \vdash \forall k > \mathtt{0} \, \exists! m \big(\forall i < k \big(F(i) \mid m\big) \wedge \forall y < m \, \exists i < k \big(F(i) \nmid y\big)\big)$$

Using induction on k, we show the above claim and the statement that every common multiple of $F(i)$ for $i < k$ is also a multiple of m. For $k = \mathtt{1}$ we can simply take $m = F(\mathtt{0})$ and for $k = \mathtt{2}$ we use $\operatorname{lcm}\big(F(\mathtt{0}), F(\mathtt{1})\big)$ defined as above. If the statement holds for $k \geq \mathtt{2}$, we use our induction hypothesis to find $m' := \operatorname{lcm}_{i<k} F(i)$ and then apply the CLAIM again to obtain $m := \operatorname{lcm}\big(m', F(k)\big)$. We claim that this m satisfies the desired properties. Indeed, since for all $i < k$ we have $F(i) \mid m'$ and $m' \mid m$, by LEMMA 8.9 we obtain $F(i) \mid m$. Hence, $F(i) \mid m$ for all $i \leq k$. Now suppose towards a contradiction that $y < m$ such that $F(i) \mid y$ for all $i \leq k$. Then, by construction of m, we either have $m' \nmid y$ or $F(k) \nmid y$. Since the latter case is clearly contradictory, we may assume that $m' \nmid y$. But since y is a common multiple of $F(i)$ for all $i < k$, by our induction hypothesis we get $m' \mid y$, a contradiction. That every common multiple of $F(i)$ for $i \leq k$ is a multiple of m can be derived as in the case $k = \mathtt{2}$. This proves that such an m exists. It there was another n with the same properties, then both would be common multiples of all $F(i)$ for $i \leq k$. But then $m \mid n$ and $n \mid m$ which implies $m = n$.

As the defining formula only uses bounded universal quantification, it is an $\exists$-formula and thus it is $\mathbb{N}$-conform.

9.4 (a) By Euclid's theorem on the infinitude of primes, we have

$$\mathbb{N} \vDash \forall n \exists p(\mathrm{prime}(p) \wedge p > n).$$

Hence for $\varphi(x) :\equiv \exists p(\mathrm{prime}(p) \wedge p > x)$ we have $\mathbb{N} \vDash \varphi(\underline{n})$ for every natural number n. By Corollary 9.3, φ is $\mathbb{N}$-conform, since $\mathrm{prime}(x)$ is equivalent to a Δ-formula. Now, by $\mathbb{N}$-conformity we obtain $\mathsf{PA} \vdash \varphi(\underline{n})$ for all $n \in \mathbb{N}$. By Exercise 7.4 there is a non-standard number $a \in M$, where M is the domain of $\mathbf{M}$, such that $\mathbf{M} \vDash \forall x(x < a \to \varphi(x))$. Now let $b \in M$ such that $b = \mathbf{s}^{\mathbf{M}}a$. Then we have $\mathbf{M} \vDash \varphi(b)$, hence there is $p > b$ such that $\mathrm{prime}(p)$. Then p is a non-standard prime number.

(b) The second part can be shown in the same way by considering

$$\varphi(x) :\equiv \exists y \forall p < x(\mathrm{prime}(p) \to p \mid y).$$

Note that $\mathbb{N} \vDash \varphi(\underline{n})$ for every $n \in \mathbb{N}$, since we can set $y = n!$.

9.5 Firstly, we define finite products as follows:

$$\mathrm{prod}(s,x) :\Longleftrightarrow \mathrm{seq}(s) \wedge \mathrm{lh}(s) > \mathsf{0} \wedge \exists t\Big(\mathrm{seq}(t) \wedge \mathrm{lh}(t) = \mathrm{lh}(s) \wedge s_0 = t_0 \wedge \forall k < \mathrm{lh}(s)\big(t_{\mathsf{s}k} = t_k \cdot s_{\mathsf{s}k}\big) \wedge t_{\mathrm{lh}(s)} = x\Big)$$

The idea is that $x = s_0 \cdot \ldots \cdot s_{\mathrm{lh}(s)-1}$ and t codes the construction process of the product. That this is functional in case $\mathrm{seq}(s)$ follows from similar arguments as in Exercise 9.2. One can further prove inductively by making use of Exercise 9.1 that

$$\mathsf{PA} \vdash \forall x \forall s \forall p\big(\mathrm{prod}(s,x) \wedge \mathrm{prime}(p) \wedge p \mid x \to \exists k < \mathrm{lh}(s)(p \mid s_k)\big).$$

Now, we can use this to easily encode prime decompositions:

$$\mathrm{p\_seq}(s,x) :\Longleftrightarrow \mathrm{prod}(s,x) \forall k < \mathrm{lh}(s)\big(\mathrm{prime}(s_k)\big)$$

We need to prove that every number $x \geq 2$ has such a prime decomposition and that it is unique. For the existence we prove

$$\mathsf{PA} \vdash \forall x\big(x \leq \mathsf{1} \vee \exists s\big(\mathrm{p\_seq}(s,x)\big)\big).$$

Towards a contradiction suppose by the Least Number Principle that $x > 1$ is a minimal counterexample. If $\mathrm{prime}(x)$, then the sequence $\langle x \rangle$ contradicts this assumption. Hence, there are y, z such that $\mathsf{1} < y$, $z < x$,

and $x = yz$. By assumption, there are sequences s and s' encoding the prime decompositions of y and z, i.e., $\mathrm{p\_seq}(s, y)$ and $\mathrm{p\_seq}(s', z)$. Then we have $\mathrm{p\_seq}(s^\frown s', x)$.

It remains to check that the prime decomposition is unique up to the order of the factors. Note that the sequence encoding the prime decomposition is not necessarily ordered (this significantly simplified the existence proof). Hence we introduce the following definition:

$$\mathrm{p\_seq\_ord}(s, x) :\Longleftrightarrow \mathrm{p\_seq}(s, x) \wedge \forall k < \mathrm{lh}(s) \forall j < k (s_j \leq s_k)$$

By induction, one can prove the following:

$$\mathsf{PA} \vdash \forall s \forall x \big(\mathrm{p\_seq}(s, x) \to \exists! s' (\mathrm{p\_seq\_ord}(s', x))\big)$$

In the induction step one removes the last number in the sequence, orders inductively the rest of the sequence and then places the removed number in the right place to obtain again an ordered sequence. We leave the details to the reader.

For uniqueness we prove

$$\mathsf{PA} \vdash \forall x \big(\forall s, s' \big((\mathrm{p\_seq\_ord}(s, x) \wedge \mathrm{p\_seq\_ord}(s', x)) \to s = s'\big)\big).$$

Suppose again that x is a minimal counterexample with s, s' encoding different ordered prime decompositions and let $\mathrm{lh}(s) = \mathtt{s}n$ and $\mathrm{lh}(s') = \mathtt{s}n'$ (they cannot be $\mathtt{0}$ since $x > \mathtt{1}$). As $s_0 \mid x$ and $\mathrm{prime}(s_n)$ and $\mathrm{prod}(s', x)$ there is $k < n'$ such that $s_n \mid s'_k$. Now, since s'_k is also prime, this implies $s_0 = s'_k$. Since s is ordered, s_n must be the maximal prime divisor of x: If $p > s_n$ was a prime divisor of x, then p would divide s_i for some $i < \mathrm{lh}(s)$ and hence $p = s_i \leq s_n$, a contradiction. With similar arguments one can show that $k = m$, i.e., $s_n = s'_m$. Now, we construct t and t' such that $s = t^\frown \langle s_n \rangle$ and $s' = t'^\frown \langle s'_m \rangle$ in the obvious way. Let x' be such that $\mathrm{prod}(t, x)$. Then by construction we also have $\mathrm{prod}(t', x)$ and further $\mathrm{p\_seq\_ord}(t', x) \wedge \mathrm{p\_seq\_ord}(t', x')$. By induction, we can conclude that $t = t'$ and hence $s = s'$.

9.6 (a) Notice first that no valid formula or term has a Gödel number in whose prime factorisation $\mathbf{2}$ has an odd power. Furthermore, the only such number without $\mathbf{2}$ as a factor represents $\mathtt{0}$ or a variable. Thus, any Gödel number x is

(i) odd, in which case it represents a variable,

(ii) in $\{\mathbf{2}k \mid \mathbf{0} \leq k \leq \mathbf{10}\}$, or

(iii) of the form $\mathbf{2}^{2a} \cdot \mathbf{3}^b \cdot \mathbf{5}^c$, where $a \geq \mathbf{1}$ and $\mathbf{2}a, b, c$ are Gödel numbers.

We use induction to show that any Gödel number x encodes at most one term or formula.

We first carry out the induction step and assume that the above is true if $x \leq \mathbf{20}$. The statement is clear if x is odd, so assume we are in the case (iii). Since $\mathbf{2}a, b, c$ are smaller than x, we can conclude that b and c and hence also x encode a unique term or formula by induction.

To show the base case when $x \leq \mathbf{20}$, we look at each case separately. By the previous remark on the power of the factor $\mathbf{2}$, we only have to consider factorisations of even numbers up to $\mathbf{20}$, where $\mathbf{2}$ has a non-zero even power:

Gödel number	possible symbols
$\mathbf{4} = \mathbf{2^2} \cdot \mathbf{3^0} \cdot \mathbf{5^0}$	$+$, $\mathtt{s0}$
$\mathbf{12} = \mathbf{2^2} \cdot \mathbf{3^1} \cdot \mathbf{5^0}$	$\wedge$, $\mathtt{s}v_0$
$\mathbf{16} = \mathbf{2^4} \cdot \mathbf{3^0} \cdot \mathbf{5^0}$	$\rightarrow$, $\mathtt{0} + \mathtt{0}$
$\mathbf{20} = \mathbf{2^2} \cdot \mathbf{3^0} \cdot \mathbf{5^1}$	$\forall$

As $+$, $\wedge$, and $\rightarrow$ are not valid terms or formulae on their own, there is no ambiguity with these Gödel numbers.

(b) Let $\varphi \equiv \forall v_0 \neg = \mathtt{0}\, v_0$. Using the Gödel numbers $\mathbf{2^8} \cdot \mathbf{3^0} \cdot \mathbf{5^1} = \mathbf{1280}$ for "$= \mathtt{0} v_0$" and $\mathbf{2^{10}} \cdot \mathbf{3^{1280}}$ for "$\neg = \mathtt{0} v_0$", we find the Gödel number

$$\mathbf{2^{20}} \cdot \mathbf{3^1} \cdot \mathbf{5^{2^{10}}} \cdot \mathbf{3^{1280}}$$

for the formula φ, and a sequence encoding the formula φ is then given by

$$\left\langle \mathbf{1280},\, \mathbf{2^{10}} \cdot \mathbf{3^{1280}},\, \mathbf{2^{20}} \cdot \mathbf{3^1} \cdot \mathbf{5^{2^{10}}} \cdot \mathbf{3^{1280}} \right\rangle .$$

Remark: The code for this sequence, which would be a Gödel code encoding φ, is much too large to be written down explicitly.

9.7 (a) Instead of encoding an infinite number of variables, we only encode the 13 logical symbols (including brackets) and reserve two numbers for the variables, i.e., we set $b = 15$. All logical symbols ζ with Gödel number $\#\zeta$ get the alternative Gödel number $\#_{\text{alt}}\,\zeta := \frac{\#\zeta}{\mathbf{2}}$. In addition, we use the numbers $\mathbf{11}$ and $\mathbf{12}$ for brackets as well as $\mathbf{13}$ and $\mathbf{14}$ to encode the variables.

Now, any formula represented by a string of symbols can be translated to a sequence of numbers in $\{\mathbf{0}, \ldots, b\}$ by replacing each logical symbol ζ by its alternative Gödel number $\#_{\text{alt}}\,\zeta$ and each variable v_k by $k+1$ repetitions of $\mathbf{13}$ followed by $\mathbf{14}$. This translation into sequences of numbers is injective.

For example, the formula $\forall v_0 \exists v_1 \big(v_0 + \mathtt{0} = v_1\big)$, which corresponds to $\forall v_0 \exists v_1 = + v_0\, \mathtt{0}\, v_1$ in Polish notation, translates to the sequence

$$\langle \mathbf{10}, \mathbf{13}, \mathbf{14}, \mathbf{9}, \mathbf{13}, \mathbf{13}, \mathbf{14}, \mathbf{4}, \mathbf{2}, \mathbf{13}, \mathbf{14}, \mathbf{0}, \mathbf{13}, \mathbf{13}, \mathbf{14} \rangle .$$

(b) As in the case of finite products (see EXERCISE 9.5) we can define in PA a relation $\mathrm{sum}(s,x)$ which states that s encodes the construction process of the finite sum x.
Let $b \geq 2$. Now, we formalise the statement that

$$x = n = n_k b^k + \ldots + n_1 b + n_0 \equiv: (n)_b \equiv: (n_k, \ldots, n_0)_b,$$

where $\mathrm{seq}(n)$ and $k + 1 = \mathrm{lh}(n)$.

$$(n)_b = x :\Longleftrightarrow \mathrm{seq}(n) \ \wedge \ \exists s\big(\mathrm{sum}(s,x) \ \wedge \ \mathrm{lh}(s) = \mathrm{lh}(n) \\ \wedge \forall i < \mathrm{lh}(n)(s_i = n_i \cdot b^i)\big)$$

We need to check the following:

$$\mathsf{PA} \vdash \forall b \forall x \big(b \geq 2 \wedge x \geq 2 \to \exists n : x = (n)_b\big)$$

We use induction on x. The base case is easy. For the induction step, using PRINCIPLE OF DIVISION WITH REMAINDER we can write $x = mb + n_0$. By induction, we have $m = (n_k, \ldots, n_1)_b$ und hence

$$x = (n_k b^{k-1} + \ldots + n_1)b + n_0 = n_k b^k + \ldots n_1 b + n_0 = (n_k, \ldots, n_0)_b.$$

(c) We define $(n_k, \ldots, n_0)_b * (m_l, \ldots, m_0)_b = y$ by stipulating

$$\begin{aligned} y &= (n_k, \ldots, n_0, m_l, \ldots, m_0)_b \\ &= n_k b^{l+k+1} + \ldots + n_0 b^{l+1} + m_l b^l + \ldots + m_0\,. \end{aligned}$$

The defining formula is functional and thus the function $*$ is definable.

(d) We can translate the above encoding of all symbols as sequences $\langle n_0, \ldots, n_k \rangle$ to numbers $(n_0, \ldots, n_k)_b$ in base b notation. The operation $*$ then allows us to recursively encode all formulas as for the standard Gödel encoding. For example, if $v = (\mathsf{13}, \ldots, \mathsf{13}, \mathsf{14})_b$ is the code for some variable v_k and $f = (n_l \ldots n_0)_b$ encodes any term or formula, then

$$\mathrm{all}(v, f) :\equiv (\mathsf{10})_b * v * f.$$

In order for this to work instead of Gödel coding, we also need to check that the base b notation is unique, which can easily be achieved by induction.

9.8 Define

$$s \restriction k = t :\Longleftrightarrow \mathrm{seq}(s) \wedge \mathrm{seq}(t) \wedge \big(k \geq \mathrm{lh}(s) \to s = t\big) \\ \wedge \big(k < \mathrm{lh}(s) \to \big(\mathrm{lh}(t) = k \wedge \forall i < k\big(t_i = s_i\big)\big)\big)$$

It is clear that such a t must exist for any code s of a sequence and any number k. As a sequence is entirely defined by all its elements, we also have uniqueness. Therefore, the defining formula is functional and $\restriction$ is well-defined.

9.9 Let τ, τ_0, φ, ν be as in LEMMA 9.14.

(a) We show $\mathsf{PA} \vdash \mathrm{sb\_term}(\ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\tau\urcorner, t) \leftrightarrow t = \ulcorner\tau(\nu/\tau_0)\urcorner$ using induction on term construction of τ. By definition we have

$$\mathsf{PA} \vdash \mathrm{sb\_term}(\ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\tau\urcorner, t) \leftrightarrow \exists c \exists c' \exists c'' \big(\mathrm{c\_sb\_term}(c, c', c'', v, t_0, t, t')\big)\,.$$

First assume τ is a constant or a variable different from ν. If c encodes the term τ and c'' the term τ_0 then $c''{}^\frown c$ satisfies

$$\mathrm{c\_sb\_term}(c, c''{}^\frown c, c'', \ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\tau\urcorner, \ulcorner\tau\urcorner)\,,$$

which gives the equivalence since $\tau(\nu/t_0) = \tau$ and because the required t is uniquely defined in c_sb_term. Now, if τ is equal to ν, then $\tau(\nu/\tau_0) = \tau_0$, and therefore, if c encodes the sequence $\langle\ulcorner\nu\urcorner\rangle$ then

$$\mathrm{c\_sb\_term}(c, c''{}^\frown c'', c'', \ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\tau\urcorner, \ulcorner\tau_0\urcorner)$$

and the equivalence follows as above. It remains to show the equivalence in the case when $\tau \equiv \mathsf{s}\tau_1$ or when $\tau \equiv F\tau_1\tau_2$ for $F \in \{+, \cdot\}$ and terms τ_1 and τ_2. We know by induction that for $i \in \{1, 2\}$:

$$\mathsf{PA} \vdash \mathrm{sb\_term}(\ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\tau_i\urcorner, t_i) \leftrightarrow t_i = \ulcorner\tau_i(\nu/\tau_0)\urcorner$$

If $\tau \equiv F\tau_1\tau_2$, then from the definition of c_sb_term we obtain

$$\begin{aligned}
\ulcorner\tau(\nu/\tau_0)\urcorner = \ulcorner F\tau_1(\nu/\tau_0)\tau_2(\nu/\tau_0)\urcorner &= 2^{\ulcorner F\urcorner} \cdot 3^{t_1} \cdot 5^{t_2} \\
&\Leftrightarrow_{\mathsf{PA}} \bigwedge_{i=1}^{2} t_i = \ulcorner\tau_i(\nu/t_0)\urcorner \\
&\Leftrightarrow_{\mathsf{PA}} \bigwedge_{i=1}^{2} \mathrm{sb_term}(\ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\tau_i\urcorner, t_i) \\
&\Leftrightarrow_{\mathsf{PA}} \bigwedge_{i=1}^{2} \exists c_i \exists c_i' \exists c'' \big(\mathrm{c_sb_term}(c_i, c_i', c'', \ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\tau_i\urcorner, t_i)\big) \\
&\Leftrightarrow_{\mathsf{PA}} \exists c \exists c' \exists c'' \big(\mathrm{c_sb_term}(c, c', c'', \ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\tau\urcorner, t)\big) \\
&\Leftrightarrow_{\mathsf{PA}} \mathrm{sb_term}(\ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\tau\urcorner, t)
\end{aligned}$$

where the penultimate equivalence uses that we can go from c_1 and c_2 to c by appending $\ulcorner F\tau_1\tau_2\urcorner$ to the sequence $c_1{}^\frown c_2$. Similarly, for c' by appending $2^{\ulcorner F\urcorner} \cdot 3^{t_1} \cdot 5^{t_2}$ to $c_1'{}^\frown \tilde{c}_2'$, where $\tilde{c}_2'$ is the sequence c_2' without the initial subsequence c''. Therefore, c' encodes a term construction of the term $2^{\ulcorner F\urcorner} \cdot 3^{t_1} \cdot 5^{t_2} = t$ and hence, proving the desired equivalence.

The case when $\tau \equiv \mathtt{s}\tau_0$ can be handled similarly.

(b) We show $\mathsf{PA} \vdash \mathrm{sb\_fml}(\ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\varphi\urcorner, f) \leftrightarrow f = \ulcorner\varphi(\nu/\tau_0)\urcorner$ using induction on formula construction of φ. In the case when $\varphi \equiv \tau_1 = \tau_2$ is atomic note that

$$\begin{aligned}
\ulcorner\varphi(\nu/\tau_0)\urcorner = \ulcorner\tau_1(\nu/\tau_0) = \tau_2(\nu/\tau_0)\urcorner &= 2^8 \cdot 3^{f_1} \cdot 5^{f_2} \\
&\Leftrightarrow_{\mathsf{PA}} \bigwedge_{i=1}^{2} f_i = \ulcorner\tau_i(\nu/t_0)\urcorner \\
&\Leftrightarrow_{\mathsf{PA}} \bigwedge_{i=1}^{2} \mathrm{sb_term}(\ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\tau_i\urcorner, f_i) \\
&\Leftrightarrow_{\mathsf{PA}} \bigwedge_{i=1}^{2} \exists c_i \exists c_i' \exists c'' \big(\mathrm{c_sb_term}(c_i, c_i', c'', \ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\tau_i\urcorner, f_i)\big) \\
&\Leftrightarrow_{\mathsf{PA}} \exists c \exists c' \big(\mathrm{c_sb_fml}(c, c', \ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\varphi\urcorner, f)\big) \\
&\Leftrightarrow_{\mathsf{PA}} \mathrm{sb_fml}(\ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\varphi\urcorner, f)
\end{aligned}$$

where we used (a) for the second equivalence and the penultimate step can be seen by picking c a code for $\langle \ulcorner\tau_1\urcorner, \ulcorner\tau_2\urcorner, \mathrm{eq}(\ulcorner\tau_1\urcorner, \ulcorner\tau_2\urcorner)\rangle$ and c' a code for $\langle f_1, f_2, \mathrm{eq}(f_1, f_2)\rangle$ and $f = 2^8 \cdot 3^{f_1} \cdot 5^{f_2}$.

If φ is of the form $\neg\psi$, $\psi_1 \wedge \psi_2$, $\psi_1 \vee \psi_2$, $\psi_1 \to \psi_2$, $\forall x\psi$, or $\exists x\psi$ we proceed similarly. We will only show $\psi_1 \to \psi_2$:

$$\begin{aligned}
\ulcorner\varphi(\nu/\tau_0)\urcorner = \ulcorner\psi_1(\nu/\tau_0) \to \psi_2(\nu/\tau_0)\urcorner &= 2^{16} \cdot 3^{f_1} \cdot 5^{f_2} \\
&\Leftrightarrow_{\mathsf{PA}} \bigwedge_{i=1}^{2} f_i = \ulcorner\psi_i(\nu/t_0)\urcorner \\
&\Leftrightarrow_{\mathsf{PA}} \bigwedge_{i=1}^{2} \mathrm{sb_fml}(\ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\psi_i\urcorner, f_i) \\
&\Leftrightarrow_{\mathsf{PA}} \bigwedge_{i=1}^{2} \exists c_i \exists c_i' \big(\mathrm{c_sb_fml}(c_i, c_i', \ulcorner\nu\urcorner, \ulcorner\tau_0\urcorner, \ulcorner\psi_i\urcorner, f_i)\big)
\end{aligned}$$

$$\Leftrightarrow_{\mathsf{PA}} \exists c \exists c' \big(\mathrm{c\_sb\_fml}(c, c', \ulcorner \nu \urcorner, \ulcorner \tau_0 \urcorner, \ulcorner \varphi \urcorner, f)\big)$$

$$\Leftrightarrow_{\mathsf{PA}} \mathrm{sb\_fml}(\ulcorner \nu \urcorner, \ulcorner \tau_0 \urcorner, \ulcorner \varphi \urcorner, f)$$

where the second equivalence uses induction and for the penultimate one note that c can be chosen as the code of the sequence encoded by $c_1 \frown c_2$ with $\mathrm{imp}(\ulcorner \psi_1 \urcorner, \ulcorner \psi_2 \urcorner)$ appended to it and similarly for c' but with $\mathrm{imp}(f_1, f_2) = f$.

Chapter 10

10.0 For part (a), we argue with a proof by cases. Suppose $\mathrm{prv}(\ulcorner\varphi\urcorner)$ and let c be a code such that $\mathrm{c\_prv}(c, \ulcorner\varphi\urcorner)$ holds. We construct a new code c' by setting $c'_i := c_i$ for every $i < \mathrm{lh}(c)$, $c'_{\mathrm{lh}(c)} := \mathrm{imp}(\ulcorner\varphi\urcorner, \mathrm{or}(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner))$ (instance of L_6), and $c'_{\mathtt{s}\,\mathrm{lh}(c)} := \ulcorner\varphi \vee \psi\urcorner$. Since $\mathrm{or}(\ulcorner\varphi\urcorner, \ulcorner\psi\urcorner) = \ulcorner\varphi \vee \psi\urcorner$, the sequence c' with $\mathrm{lh}(c') = \mathtt{ss}\,\mathrm{lh}(c)$ satisfies $\mathrm{mp}(c'_{\mathrm{lh}(c)}, c'_{\mathrm{lh}(c)-1}, c'_{\mathtt{s}\,\mathrm{lh}(c)})$, and hence, $\mathrm{c\_prv}(c', \ulcorner\varphi \vee \psi\urcorner)$, which shows that $\mathrm{prv}(\ulcorner\varphi \vee \psi\urcorner)$.
The other case is very similar. Starting with the assumption $\mathrm{prv}(\ulcorner\psi\urcorner)$, we find a code for c_prv and can add an instance of L_7 instead of L_6 to obtain the same conclusion as above.

Regarding part (b), with the help of the First Incompleteness Theorem we will show that the converse does not hold. If $\mathrm{prv}(\ulcorner\varphi \vee \psi\urcorner) \to \big(\mathrm{prv}(\ulcorner\varphi\urcorner) \vee \mathrm{prv}(\ulcorner\psi\urcorner)\big)$ was provable in PA, then, by the Soundness Theorem 3.7, we have

$$\mathbb{N} \vDash \mathrm{prv}(\#(\varphi \vee \psi)) \to \big(\mathrm{prv}(\#\varphi) \vee \mathrm{prv}(\#\psi)\big) \qquad (*)$$

for every $\mathscr{L}_{\mathsf{PA}}$-formulae φ and ψ. Now, let σ be an $\mathscr{L}_{\mathsf{PA}}$-sentence with $\mathsf{PA} \nvdash \sigma$ and $\mathsf{PA} \nvdash \neg\sigma$. By Lemma 9.16, we obtain $\mathbb{N} \nvDash \mathrm{prv}(\#\sigma)$ and $\mathbb{N} \nvDash \mathrm{prv}(\#(\neg\sigma))$. Furthermore, as an instance of L_0, we have $\mathsf{PA} \vdash \sigma \vee \neg\sigma$, which gives $\mathbb{N} \vDash \mathrm{prv}(\#(\sigma \vee \neg\sigma))$ by the same lemma. Thus, we have

$$\mathbb{N} \vDash \mathrm{prv}(\#(\sigma \vee \neg\sigma)) \wedge \neg\big(\mathrm{prv}(\#\sigma) \vee \mathrm{prv}(\#\neg\sigma)\big)$$

which is a contradiction to $(*)$.

10.1 Pick an $\mathscr{L}_{\mathsf{PA}}$-sentence σ satisfying $\sigma \Leftrightarrow_{\mathsf{PA}} \neg\,\mathrm{prv}(\ulcorner\sigma\urcorner)$, as in the proof of the First Incompleteness Theorem 10.5, and let

$$\varphi(x) :\equiv \big(\neg\,\mathrm{c\_prv}(x, \ulcorner\sigma\urcorner)\big)\,.$$

Then, since $\mathbb{N} \vDash \sigma$, for each $n \in \mathbb{N}$ we have $\mathbb{N} \vDash \big(\neg\,\mathrm{c\_prv}(n, \#\sigma)\big)$, and therefore we have $\mathsf{PA} \vdash \varphi(\underline{n})$ for each $n \in \mathbb{N}$ (cf. Proposition 9.2). On the other hand, since $\mathsf{PA} + \neg\sigma$ is consistent, there exists a model $\mathbf{M}$ of $\mathsf{PA} + \neg\sigma$, in particular, $\mathbf{M} \vDash \exists y(\mathrm{c\_prv}(y, \ulcorner\sigma\urcorner))$. Let a be a witness for y in the domain of $\mathbf{M}$. Then, $\mathbf{M} \nvDash \varphi(a)$, and thus, $\mathbf{M} \nvDash \forall x\varphi(x)$, which shows that $\mathsf{PA} \nvdash \forall x\varphi(x)$.

10.2 For $i = 1$ or $i = 2$, let

$$\begin{aligned}\xi_i(v_0, v_1) :\equiv \forall v_2 \forall v_3 \forall v_4 \forall v_5 \big(&(\mathrm{sb\_fml}(\ulcorner v_0 \urcorner, \mathrm{gn}(v_0), v_0, v_2) \\ &\wedge \mathrm{sb\_fml}(\ulcorner v_1 \urcorner, \mathrm{gn}(v_1), v_2, v_3) \wedge \mathrm{sb\_fml}(\ulcorner v_0 \urcorner, \mathrm{gn}(v_0), v_1, v_4) \\ &\wedge \mathrm{sb\_fml}(\ulcorner v_1 \urcorner, \mathrm{gn}(v_1), v_4, v_5)) \to \varphi_i(v_3, v_5)\big)\end{aligned}$$

and define $\sigma_i :\equiv \xi_i(\ulcorner\xi_1\urcorner, \ulcorner\xi_2\urcorner)$. Then we have:

$$\begin{aligned}
\sigma_i \equiv\ & \forall v_2 \forall v_3 \forall v_4 \forall v_5 \big((\mathrm{sb_fml}(\mathsf{1}, \mathrm{gn}(\ulcorner\xi_1\urcorner), \ulcorner\xi(v_0, v_1)\urcorner, v_2) \\
& \quad \wedge \mathrm{sb_fml}(\mathsf{3}, \mathrm{gn}(\ulcorner\xi_2\urcorner), v_2, v_3) \wedge \mathrm{sb_fml}(\mathsf{1}, \mathrm{gn}(\ulcorner\xi_1\urcorner), \ulcorner\xi_2(v_0, v_1)\urcorner, v_4) \\
& \quad \wedge \mathrm{sb_fml}(\mathsf{3}, \mathrm{gn}(\ulcorner\xi_2\urcorner), v_4, v_5)) \to \varphi_i(v_3, v_5) \big) \\
\Leftrightarrow_{\mathsf{PA}}\ & \forall v_2 \forall v_3 \forall v_4 \forall v_5 \big((\mathrm{sb_fml}(\mathsf{1}, \ulcorner\ulcorner\xi_1\urcorner\urcorner, \ulcorner\xi_1(v_0, v_1)\urcorner, v_2) \\
& \quad \wedge \mathrm{sb_fml}(\mathsf{3}, \ulcorner\ulcorner\xi_2\urcorner\urcorner, v_2, v_3) \wedge \mathrm{sb_fml}(\mathsf{1}, \ulcorner\ulcorner\xi_1\urcorner\urcorner, \ulcorner\xi_2(v_0, v_1)\urcorner, v_4) \\
& \quad \wedge \mathrm{sb_fml}(\mathsf{3}, \ulcorner\ulcorner\xi_2\urcorner\urcorner, v_4, v_5)) \to \varphi_i(v_3, v_5) \big) \\
\Leftrightarrow_{\mathsf{PA}}\ & \forall v_2 \forall v_3 \forall v_4 \forall v_5 \big((v_2 = \ulcorner\xi_1(\ulcorner\xi_1\urcorner, v_1)\urcorner \wedge \mathrm{sb_fml}(\mathsf{3}, \ulcorner\ulcorner\xi_2\urcorner\urcorner, v_2, v_3) \\
& \quad \wedge v_4 = \ulcorner\xi_2(\ulcorner\xi_1\urcorner, v_1)\urcorner \wedge \mathrm{sb_fml}(\mathsf{3}, \ulcorner\ulcorner\xi_2\urcorner\urcorner, v_4, v_5)) \to \varphi_i(v_3, v_5) \big) \\
\Leftrightarrow_{\mathsf{PA}}\ & \forall v_3 \forall v_5 \big((v_3 = \ulcorner\xi_1(\ulcorner\xi_1\urcorner, \ulcorner\xi_2\urcorner)\urcorner \wedge v_5 = \ulcorner\xi_2(\ulcorner\xi_1\urcorner, \ulcorner\xi_2\urcorner\urcorner) \\
& \quad \to \varphi_i(v_3, v_5) \big) \\
\Leftrightarrow_{\mathsf{PA}}\ & \varphi(\ulcorner\xi_1(\ulcorner\xi_1\urcorner, \ulcorner\xi_2\urcorner)\urcorner, \ulcorner\xi_2(\ulcorner\xi_1\urcorner, \ulcorner\xi_2\urcorner)\urcorner) \\
\equiv\ & \varphi_i(\ulcorner\sigma_1\urcorner, \ulcorner\sigma_2\urcorner)
\end{aligned}$$

10.3 If there is a positive even integer $n_0 \in \mathbb{N}$ which cannot be written as a sum of two primes, then, since n_0 is finite, this integer n_0 exists in every model of PA, which implies that $\mathsf{PA} \vdash \neg\mathsf{GC}$. Hence, $\mathsf{PA} \nvdash \neg\mathsf{GC}$ implies that every positive even integer $n \in \mathbb{N}$ can be written as a sum of two primes, and therefore, $\mathbb{N} \models \mathsf{GC}$.

10.4 Let $\mathsf{T} = \mathsf{PA} + \neg\sigma$, where σ is the $\mathscr{L}_{\mathsf{PA}}$-sentence such that $\mathsf{PA} \nvdash \sigma$ and $\mathsf{PA} \nvdash \neg\sigma$ considered in the proof of First Incompleteness Theorem. By construction, T is consistent. It remains to check that T is not ω-consistent. Consider the formula

$$\varphi(x) :\equiv \mathrm{c\_prv}(x, \ulcorner\sigma\urcorner).$$

Since $\mathsf{PA} \vdash \sigma \leftrightarrow \neg\,\mathrm{prv}(\ulcorner\sigma\urcorner)$ and $\mathsf{T} \vdash \neg\sigma$, we have $\mathsf{T} \vdash \mathrm{prv}(\ulcorner\sigma\urcorner)$, hence $\mathsf{T} \vdash \exists x\varphi(x)$. In order to check that T is not ω-consistent, we need to prove that for $n \in \mathbb{N}$ we have $\mathsf{T} \vdash \neg\varphi(\underline{n})$. Assume towards a contradiction that this is not the case. Hence, there is $n \in \mathbb{N}$ such that $\mathsf{T} \nvdash \neg\varphi(\underline{n})$. Then there is a model $\mathbf{M}$ of $\mathsf{T} + \varphi(\underline{n})$, i.e., $\mathbf{M} \models \mathrm{c\_prv}(\underline{n}, \ulcorner\sigma\urcorner)$. But since c_prv is equivalent to a formula which contains only bounded quantifiers, n codes a standard proof of σ and thus $\mathbf{M} \models \sigma$, a contradiction.

Chapter 11

11.0 We show that $\mathsf{PA} \nvdash \mathrm{con}_{\mathsf{PA}}$, i.e., we show that $\mathsf{PA} \nvdash \neg\,\mathrm{prv}(\ulcorner \mathtt{0} = \mathtt{1} \urcorner)$: Assume towards a contradiction that $\mathsf{PA} \vdash \neg\,\mathrm{prv}(\ulcorner \mathtt{0} = \mathtt{1} \urcorner)$. With L_9 and (MP) we deduce:

$$\begin{aligned}
&\mathsf{PA} \vdash \neg\,\mathrm{prv}(\ulcorner \mathtt{0} = \mathtt{1} \urcorner)\\
&\mathsf{PA} \vdash \neg\,\mathrm{prv}(\ulcorner \mathtt{0} = \mathtt{1} \urcorner) \to \big(\mathrm{prv}(\ulcorner \mathtt{0} = \mathtt{1} \urcorner) \to \mathtt{0} = \mathtt{1}\big)\\
&\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \mathtt{0} = \mathtt{1} \urcorner) \to \mathtt{0} = \mathtt{1}
\end{aligned}$$

Thus, by Löb's Theorem we obtain $\mathsf{PA} \vdash \mathtt{0} = \mathtt{1}$, which contradicts PA_0.

11.1 We will first prove the following more general statements for $V \subseteq \{x_1, \dots, x_n\}$:

$$\mathsf{PA} \vdash \ulcorner \tau(x_1/\underline{m_1}, \dots, x_n/\underline{m_n}) \urcorner = \lceil \tau \rceil_V^{\mathrm{gn}}(x_1/\underline{m_1}, \dots, x_n/\underline{m_n}) \qquad (1)$$
$$\mathsf{PA} \vdash \ulcorner \varphi(x_1/\underline{m_1}, \dots, x_n/\underline{m_n}) \urcorner = \lceil \varphi \rceil_V^{\mathrm{gn}}(x_1/\underline{m_1}, \dots, x_n/\underline{m_n}) \qquad (2)$$

Under the assumption that (1) and (2) hold, the claim follows for $V = \mathrm{free}(\varphi) \subseteq \{x_1, \dots, x_n\}$. For the sake of simplicity, we will write $(x_i/\underline{m_i})$ instead of $(x_1/\underline{m_1}, \dots, x_n/\underline{m_n})$ and $(x_i/\,\mathrm{gn}(x_i))$ instead of $(x_1/\,\mathrm{gn}(x_1), \dots, x_n/\,\mathrm{gn}(x_n))$. We show (1) by induction on term construction:

- If $\tau \equiv \mathtt{0}$ then there are no free variables, so we can ignore the substitution. Moreover, we have $\lceil \tau \rceil_V^{\mathrm{gn}} = \lceil \tau \rceil = \ulcorner \mathtt{0} \urcorner = \ulcorner \tau \urcorner$.
- If τ is a variable ν which is not in V, then ν is not a free variable of $\lceil \tau \rceil_V^{\mathrm{gn}}$ and hence we obtain:

$$\lceil \tau \rceil_V^{\mathrm{gn}}(x_i/\underline{m_i}) = \ulcorner \nu \urcorner(x_i/\,\mathrm{gn}(x_i))(x_i/\underline{m_i}) = \ulcorner \nu \urcorner = \ulcorner \nu(x_i/\underline{m_i}) \urcorner$$

- If τ is any variable which occurs in V (without loss of generality we may assume $\tau \equiv x_1$), then by Lemma 10.3:

$$\begin{aligned}
\lceil \tau \rceil_V^{\mathrm{gn}}(x_i/\underline{m_i}) &= \lceil \tau \rceil_V(x_i/\,\mathrm{gn}(x_i))(x_i/\underline{m_i}) = x_1(x_i/\,\mathrm{gn}(x_i))(x_i/\underline{m_i})\\
&= \mathrm{gn}(x_1)(x_i/\underline{m_i}) = \mathrm{gn}(\underline{m_1}) = \ulcorner \underline{m_1} \urcorner = \ulcorner \tau(x_i/\underline{m_i}) \urcorner
\end{aligned}$$

- Assume now that $\tau \equiv \mathtt{s}\sigma$ where the statement holds for σ, then we obtain:

$$\begin{aligned}
\lceil \tau \rceil_V^{\mathrm{gn}}(x_1/\underline{m_1}, \dots, x_n/\underline{m_n}) &= \lceil \mathtt{s}\sigma \rceil_V(x_i/\,\mathrm{gn}(x_i))(x_i/\underline{m_i})\\
&= \mathrm{succ}(\lceil \sigma \rceil_V)(x_i/\,\mathrm{gn}(x_i))(x_i/\underline{m_i})
\end{aligned}$$

$$
\begin{aligned}
&= \mathrm{succ}(\lceil\sigma\rceil_V^{\mathrm{gn}}(x_i/\underline{m_i})) \\
&= \mathrm{succ}(\ulcorner\sigma(x_i/\underline{m_i})\urcorner) \\
&= \ulcorner \mathtt{s}\sigma(x_i/\underline{m_i})\urcorner = \ulcorner\tau(x_i/\underline{m_i})\urcorner
\end{aligned}
$$

- In the case when $\tau \equiv \sigma_1 + \sigma_2$ where the statement holds for σ_1 and σ_2, then we obtain:

$$
\begin{aligned}
\lceil\tau\rceil_V^{\mathrm{gn}}(x_i/\underline{m_i}) &= \lceil\sigma_1 + \sigma_2\rceil_V(x_i/\,\mathrm{gn}(x_i))(x_i/\underline{m_i}) \\
&= \mathrm{add}(\lceil\sigma_1\rceil_V, \lceil\sigma_2\rceil_V)(x_i/\,\mathrm{gn}(x_i))(x_i/\underline{m_i}) \\
&= \mathrm{add}(\lceil\sigma_1\rceil_V^{\mathrm{gn}}(x_i/\underline{m_i}), \lceil\sigma_2\rceil_V^{\mathrm{gn}}(x_i/\underline{m_i})) \\
&= \mathrm{add}(\ulcorner\sigma_1(x_i/\underline{m_i})\urcorner, \ulcorner\sigma_2(x_i/\underline{m_i})\urcorner) \\
&= \ulcorner(\sigma_1 + \sigma_2)(x_i/\underline{m_i})\urcorner
\end{aligned}
$$

- The case when $\tau = \sigma_1 \cdot \sigma_2$ follows analogously.

We now prove (2) by induction on formula construction:

- If $\varphi \equiv \tau_1 = \tau_2$, then with the previous result we obtain:

$$
\begin{aligned}
\lceil\varphi\rceil_V^{\mathrm{gn}}(x_i/\underline{m_i}) &= \lceil\tau_1 = \tau_2\rceil(x_i/\,\mathrm{gn}(x_i))(x_i/\underline{m_i}) \\
&= \mathrm{eq}(\lceil\tau_1\rceil_V, \lceil\tau_2\rceil_V)(x_i/\,\mathrm{gn}(x_i))(x_i/\underline{m_i}) \\
&= \mathrm{eq}(\lceil\tau_1\rceil_V^{\mathrm{gn}}(x_i/\underline{m_i}), \lceil\tau_2\rceil_V^{\mathrm{gn}}(x_i/\underline{m_i})) \\
&= \mathrm{eq}(\ulcorner\tau_1(x_i/\underline{m_i})\urcorner, \ulcorner\tau_2(x_i/\underline{m_i})\urcorner) \\
&= \ulcorner(\tau_1 = \tau_2)(x_i/\underline{m_i})\urcorner
\end{aligned}
$$

- If $\varphi \equiv \varphi_1 \wedge \varphi_2$, then by induction we obtain:

$$
\begin{aligned}
\lceil\varphi\rceil_V^{\mathrm{gn}}(x_i/\underline{m_i}) &= \lceil\varphi_1 \wedge \varphi_2\rceil_V(x_i/\,\mathrm{gn}(x_i))(x_i/\underline{m_i}) \\
&= \mathrm{and}(\lceil\varphi_1\rceil_V, \lceil\varphi_2\rceil_V)(x_i/\,\mathrm{gn}(x_i))(x_i/\underline{m_i}) \\
&= \mathrm{and}(\lceil\varphi_1\rceil_V^{\mathrm{gn}}(x_i/\underline{m_i}), \lceil\varphi_2\rceil_V^{\mathrm{gn}}(x_i/\underline{m_i})) \\
&= \mathrm{and}(\ulcorner\varphi_1(x_i/\underline{m_i})\urcorner, \ulcorner\varphi_2(x_i/\underline{m_i})\urcorner) \\
&= \ulcorner(\varphi_1 \wedge \varphi_2)(x_i/\underline{m_i})\urcorner
\end{aligned}
$$

- If $\neg\varphi, \varphi_1 \vee \varphi_2, \varphi_1 \to \varphi_2$ we proceed similarly.
- If $\varphi \equiv \exists\nu\psi$ where $\nu \notin V$, then $V \setminus \{\nu\} = V$ and we have:

$$
\begin{aligned}
\lceil\varphi\rceil_V^{\mathrm{gn}}(x_i/\underline{m_i}) &= \lceil\exists\nu\psi\rceil_V(x_i/\,\mathrm{gn}(x_i))(x_i/\underline{m_i}) \\
&= \mathrm{ex}(\ulcorner\nu\urcorner, \lceil\psi\rceil_V)(x_i/\,\mathrm{gn}(x_i))(x_i/\underline{m_i}) \\
&= \mathrm{ex}(\ulcorner\nu\urcorner, \lceil\psi\rceil_V^{\mathrm{gn}}(x_i/\underline{m_i})) \\
&= \mathrm{ex}(\ulcorner\nu\urcorner, \ulcorner\psi(x_i/\underline{m_i})\urcorner) \\
&= \ulcorner\exists\nu\psi(x_i/\underline{m_i})\urcorner
\end{aligned}
$$

- If $\varphi \equiv \exists \nu \psi$ such that $\nu \in V$ (without loss of generality we may assume $\nu \equiv x_1$) and $V' = V \setminus \{x_1\}$, then we have:

$$\begin{aligned}
\ulcorner\varphi\urcorner_V^{\mathrm{gn}}(x_i/\underline{m_i}) &= \ulcorner\exists x_1\psi\urcorner_V(x_i/\operatorname{gn}(x_i))(x_i/\underline{m_i}) \\
&= \operatorname{ex}(\ulcorner x_1\urcorner, \ulcorner\psi\urcorner_{V'})(x_i/\operatorname{gn}(x_i))(x_i/\underline{m_i}) \\
&= \operatorname{ex}(\ulcorner x_1\urcorner, \ulcorner\psi\urcorner_{V'}^{\mathrm{gn}}(x_2/\underline{m_2}, \ldots, x_n/\underline{m_n}) \\
&= \operatorname{ex}(\ulcorner x_1\urcorner, \ulcorner\psi(x_2/\underline{m_2}, \ldots, x_n/\underline{m_n})\urcorner) \\
&= \ulcorner\exists x_1\psi(x_2/\underline{m_2}, \ldots, x_n/\underline{m_n})\urcorner \\
&= \ulcorner\varphi(x_i/\underline{m_i})\urcorner
\end{aligned}$$

  In the third and last we have used that x_1 is neither a free variable of $\ulcorner\psi\urcorner_{V'}$ nor of φ.

- The last case, when $\varphi \equiv \forall \nu \psi$, is treated similarly.

11.2 From the proof of Theorem 11.4 we know that (∗) holds for formulas of the form $v_l = v_k$. Thus, we obtain

$$\begin{aligned}
&\mathsf{PA} \vdash v_l = v_k \to \operatorname{prv}(\ulcorner v_l = v_k\urcorner^{\mathrm{gn}}) \\
&\mathsf{PA} \vdash \forall v_l\big(v_l = v_k \to \operatorname{prv}(\ulcorner v_l = v_k\urcorner^{\mathrm{gn}})\big) \\
&\mathsf{PA} \vdash \forall v_l\big(v_l = v_k \to \operatorname{prv}(\ulcorner v_l = v_k\urcorner^{\mathrm{gn}})\big) \\
&\qquad\qquad \to \big(v_i + v_j = v_k \to \operatorname{prv}(\ulcorner v_i + v_j = v_k\urcorner^{\mathrm{gn}})\big)
\end{aligned}$$

$$\mathsf{PA} \vdash v_i + v_j = v_k \to \operatorname{prv}(\ulcorner v_i + v_j = v_k\urcorner^{\mathrm{gn}})$$

where we used L_{10} with the admissible substitution $(v_l = v_k)(v_l/(v_i+v_j))$. With similar arguments, we can show that (∗) holds also for formulae of the form $v_i \cdot v_j = v_k$.

11.3 We have to show that $\mathsf{PA} \vdash \varphi \vee \psi \to \operatorname{prv}(\ulcorner\varphi \vee \psi\urcorner^{\mathrm{gn}})$: By Corollary 11.9, L_6 and L_7 we obtain

$$\mathsf{PA} \vdash \operatorname{prv}(\ulcorner\varphi\urcorner^{\mathrm{gn}}) \to \operatorname{prv}(\ulcorner\varphi \vee \psi\urcorner^{\mathrm{gn}})$$

as well as

$$\mathsf{PA} \vdash \operatorname{prv}(\ulcorner\psi\urcorner^{\mathrm{gn}}) \to \operatorname{prv}(\ulcorner\varphi \vee \psi\urcorner^{\mathrm{gn}}),$$

Hence, with L_8 and (MP) we deduce

$$\mathsf{PA} \vdash \big(\operatorname{prv}(\ulcorner\varphi\urcorner^{\mathrm{gn}}) \vee \operatorname{prv}(\ulcorner\psi\urcorner^{\mathrm{gn}})\big) \to \operatorname{prv}(\ulcorner\varphi \vee \psi\urcorner^{\mathrm{gn}}).$$

From $\mathsf{PA} \vdash \varphi \to \operatorname{prv}(\ulcorner\varphi\urcorner^{\mathrm{gn}})$ and $\mathsf{PA} \vdash \operatorname{prv}(\ulcorner\varphi\urcorner^{\mathrm{gn}}) \to \operatorname{prv}(\ulcorner\varphi \vee \psi\urcorner^{\mathrm{gn}})$ we obtain $\mathsf{PA} \vdash \varphi \to \operatorname{prv}(\ulcorner\varphi \vee \psi\urcorner^{\mathrm{gn}})$ with Tautology (D.0), and similarly

for ψ. We can thus conclude with the following instance of L_8 by applying twice (MP):

$$\mathsf{PA} \vdash \big(\varphi \to \mathrm{prv}(\lceil \varphi \vee \psi \rceil^{\mathrm{gn}})\big) \to \Big(\big(\psi \to \mathrm{prv}(\lceil \varphi \vee \psi \rceil^{\mathrm{gn}})\big) \to \big(\varphi \vee \psi \to \mathrm{prv}(\lceil \varphi \vee \psi \rceil^{\mathrm{gn}})\big)\Big)\,.$$

11.4 (a) We define $\psi :\equiv \mathrm{prv}(\ulcorner \mathrm{prv}(\ulcorner \varphi \urcorner) \to \varphi \urcorner) \to \mathrm{prv}(\ulcorner \varphi \urcorner)$ and show that $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \psi \urcorner) \to \psi$. By Löb's Theorem, this allows us to conclude that $\mathsf{PA} \vdash \psi$ — notice that ψ is an $\mathscr{L}_{\mathsf{PA}}$-sentence and not just a formula: By (DT) it is enough to show that

$$\mathsf{PA} + \Big\{ \mathrm{prv}(\ulcorner \psi \urcorner), \underbrace{\mathrm{prv}(\ulcorner \mathrm{prv}(\ulcorner \varphi \urcorner) \to \varphi \urcorner)}_{\equiv:\,\vartheta} \Big\} \vdash \mathrm{prv}(\ulcorner \varphi \urcorner)\,.$$

From $\mathrm{prv}(\ulcorner \psi \urcorner)$, (MP), and D_1 applied to ψ we obtain:

$$\mathrm{prv}(\ulcorner \vartheta \urcorner) \to \mathrm{prv}(\ulcorner \mathrm{prv}(\ulcorner \varphi \urcorner) \urcorner)$$

With $\vartheta \equiv \mathrm{prv}(\ulcorner \mathrm{prv}(\ulcorner \varphi \urcorner) \to \varphi \urcorner)$, D_2 applied to $\mathrm{prv}(\ulcorner \varphi \urcorner) \to \varphi$, and (MP), we obtain $\mathrm{prv}(\ulcorner \vartheta \urcorner)$, and therefore also $\mathrm{prv}(\ulcorner \mathrm{prv}(\ulcorner \varphi \urcorner) \urcorner)$. Finally, D_1 applied to $\mathrm{prv}(\ulcorner \varphi \urcorner) \to \varphi$ gives us

$$\vartheta \to (\mathrm{prv}(\ulcorner \mathrm{prv}(\ulcorner \varphi \urcorner) \urcorner) \to \mathrm{prv}(\ulcorner \varphi \urcorner))$$

from which we obtain the desired result using (MP).

(b) Set $\psi :\equiv \mathrm{prv}(\ulcorner \neg\, \mathrm{prv}(\ulcorner \varphi \urcorner) \urcorner) \to \mathrm{prv}(\ulcorner \varphi \urcorner)$. Then, by Tautology (G), we have to show that $\mathsf{PA} \vdash \psi$, and by Löb's Theorem, it is enough to show that $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \psi \urcorner) \to \psi$, and by applying (DT) twice, it is enough to show that

$$\mathsf{PA} + \Big\{ \mathrm{prv}(\ulcorner \psi \urcorner), \mathrm{prv}(\ulcorner \neg\, \mathrm{prv}(\ulcorner \varphi \urcorner) \urcorner) \Big\} \vdash \mathrm{prv}(\ulcorner \varphi \urcorner)\,.$$

From $\mathrm{prv}(\ulcorner \psi \urcorner)$, D_1 and (MP) we deduce

$$\mathrm{prv}(\ulcorner \mathrm{prv}(\ulcorner \neg\, \mathrm{prv}(\ulcorner \varphi \urcorner) \urcorner) \urcorner) \to \mathrm{prv}(\ulcorner \mathrm{prv}(\ulcorner \varphi \urcorner) \urcorner)\,.$$

With $\mathrm{prv}(\ulcorner \neg\, \mathrm{prv}(\ulcorner \varphi \urcorner) \urcorner)$, D_2 and (MP) we obtain

$$\mathrm{prv}(\ulcorner \mathrm{prv}(\ulcorner \neg\, \mathrm{prv}(\ulcorner \varphi \urcorner) \urcorner) \urcorner)$$

and therefore also $\mathrm{prv}(\ulcorner \mathrm{prv}(\ulcorner \varphi \urcorner) \urcorner)$. Now, by L_9 we have

$$\mathsf{PA} \vdash \neg\, \mathrm{prv}(\ulcorner \varphi \urcorner) \to \big(\mathrm{prv}(\ulcorner \varphi \urcorner) \to \varphi\big)\,,$$

hence, by D_0, we also have

$$\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \neg\, \mathrm{prv}(\ulcorner \varphi \urcorner) \to \big(\mathrm{prv}(\ulcorner \varphi \urcorner) \to \varphi\big) \urcorner)\,.$$

Applying D_1 and (MP), we obtain

$$\mathrm{prv}(\ulcorner \neg \mathrm{prv}(\ulcorner \varphi \urcorner) \urcorner) \to \mathrm{prv}(\ulcorner \mathrm{prv}(\ulcorner \varphi \urcorner) \to \varphi \urcorner)$$

and by (MP) we have $\mathrm{prv}(\ulcorner \mathrm{prv}(\ulcorner \varphi \urcorner) \to \varphi \urcorner)$. Now, we can deduce $\mathrm{prv}(\ulcorner \mathrm{prv}(\ulcorner \varphi \urcorner) \urcorner) \to \mathrm{prv}(\ulcorner \varphi \urcorner)$ using D_1 and finally get $\mathrm{prv}(\varphi)$.

(c) If $\mathsf{PA} \vdash \neg\, \mathrm{prv}(\ulcorner \neg\, \mathrm{prv}(\ulcorner \mathsf{0} = \mathsf{1} \urcorner) \urcorner)$, then by L_1 we have

$$\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \mathsf{0} = \mathsf{1} \urcorner) \to \neg\, \mathrm{prv}(\ulcorner \neg\, \mathrm{prv}(\ulcorner \mathsf{0} = \mathsf{1} \urcorner) \urcorner)\,,$$

or equivalently, by Tautology (G):

$$\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \neg\, \mathrm{prv}(\ulcorner \mathsf{0} = \mathsf{1} \urcorner) \urcorner) \to \neg\, \mathrm{prv}(\ulcorner \mathsf{0} = \mathsf{1} \urcorner)$$

Now, by Löb's Theorem we then obtain $\mathsf{PA} \vdash \neg\, \mathrm{prv}(\ulcorner \mathsf{0} = \mathsf{1} \urcorner)$ which contradicts the Second Incompleteness Theorem 11.0. Therefore, $\mathsf{PA} \nvdash \neg\, \mathrm{prv}(\ulcorner \neg\, \mathrm{prv}(\ulcorner \mathsf{0} = \mathsf{1} \urcorner) \urcorner)$, i.e., the Second Incompleteness Theorem is *not* provable within PA.

11.5 We assume that $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \varphi \urcorner) \wedge \big(\mathrm{prv}(\ulcorner \psi \urcorner) \to \psi\big)$. Notice that if ψ is an $\mathscr{L}_{\mathsf{PA}}$-sentence, then we can use L_4 and Löb's Theorem to obtain $\mathsf{PA} \vdash \psi$. We can then conclude with L_1 and (MP).
In general this does not work if ψ is an arbitrary $\mathscr{L}_{\mathsf{PA}}$-formula. In this case, we first use D_0 to find $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \mathrm{prv}(\ulcorner \psi \urcorner) \to \psi \urcorner)$ and then obtain $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \psi \urcorner)$ with Exercise 11.4.(a). This implies again $\mathsf{PA} \vdash \psi$ and hence we can conclude as above.

11.6 We first prove $\mathsf{PA} + \big\{\, \mathrm{prv}(\ulcorner \varphi \leftrightarrow \mathrm{con}_{\mathsf{PA}} \urcorner), \mathrm{prv}(\ulcorner \varphi \urcorner) \big\} \vdash \neg\, \mathrm{con}_{\mathsf{PA}}$. Note that by definition of $\leftrightarrow$ and by Corollary 10.2 we have

$$\mathrm{prv}(\ulcorner \varphi \leftrightarrow \mathrm{con}_{\mathsf{PA}} \urcorner) \Leftrightarrow_{\mathsf{PA}} \mathrm{prv}(\ulcorner \varphi \to \mathrm{con}_{\mathsf{PA}} \urcorner) \wedge \mathrm{prv}(\ulcorner \mathrm{con}_{\mathsf{PA}} \to \varphi \urcorner)\,.$$

With L_3, D_1 and (MP) we obtain $\mathrm{prv}(\ulcorner \varphi \urcorner) \to \mathrm{prv}(\ulcorner \mathrm{con}_{\mathsf{PA}} \urcorner)$ and thus also $\mathrm{prv}(\ulcorner \mathrm{con}_{\mathsf{PA}} \urcorner)$. From Exercise 11.4.(b) we get

$$\mathsf{PA} \vdash \neg\, \mathrm{prv}(\ulcorner \mathsf{0} = \mathsf{1} \urcorner) \to \neg\, \mathrm{prv}(\ulcorner \neg\, \mathrm{prv}(\ulcorner \mathsf{0} = \mathsf{1} \urcorner) \urcorner)$$

or equivalently, $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \mathrm{con}_{\mathsf{PA}} \urcorner) \to \neg\, \mathrm{con}_{\mathsf{PA}}$. Now, with (MP) and by (DT) we obtain

$$\mathsf{PA} + \big\{\, \mathrm{prv}(\ulcorner \varphi \leftrightarrow \mathrm{con}_{\mathsf{PA}} \urcorner) \big\} \vdash \mathrm{prv}(\ulcorner \varphi \urcorner) \to \neg\, \mathrm{con}_{\mathsf{PA}}\,.$$

Next we show that

$$\mathsf{T} := \mathsf{PA} + \big\{\, \mathrm{prv}(\ulcorner \varphi \leftrightarrow \mathrm{con}_{\mathsf{PA}} \urcorner), \neg\, \mathrm{prv}(\ulcorner \varphi \urcorner), \neg\, \mathrm{con}_{\mathsf{PA}} \,\big\} \vdash \lightning.$$

Just as above we can obtain $\mathrm{prv}(\ulcorner \mathrm{con}_{\mathsf{PA}} \urcorner) \to \mathrm{prv}(\ulcorner \varphi \urcorner)$. Towards a contradiction we will show that $\mathsf{T} \vdash \mathrm{prv}(\ulcorner \mathrm{con}_{\mathsf{PA}} \urcorner)$, this will in turn imply $\mathsf{T} \vdash \mathrm{prv}(\ulcorner \varphi \urcorner)$ and therefore $\mathsf{T} \vdash \lightning$. Since $\mathsf{PA} \vdash \neg(\mathtt{0} = \mathtt{1})$, it follows from D_0 that $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \neg(\mathtt{0} = \mathtt{1}) \urcorner)$ and therefore the same is provable in T. We now apply D_1 to an instance of L_9 to obtain $\mathrm{prv}(\ulcorner \neg(\mathtt{0} = \mathtt{1}) \to \big(\mathtt{0} = \mathtt{1} \to \mathrm{con}_{\mathsf{PA}}\big) \urcorner)$. Together with multiple applications of D_1 and (MP), we first deduce $\mathsf{T} \vdash \mathrm{prv}(\ulcorner \mathtt{0} = \mathtt{1} \to \mathrm{con}_{\mathsf{PA}} \urcorner)$ and then $\mathsf{T} \vdash \mathrm{prv}(\ulcorner \mathrm{con}_{\mathsf{PA}} \urcorner)$, using that $\neg\, \mathrm{con}_{\mathsf{PA}} \in \mathsf{T}$. Thus, with Corollary 2.8 and (DT) we finally obtain

$$\mathsf{PA} + \big\{ \mathrm{prv}(\ulcorner \varphi \leftrightarrow \mathrm{con}_{\mathsf{PA}} \urcorner) \big\} \vdash \neg\, \mathrm{con}_{\mathsf{PA}} \to \mathrm{prv}(\ulcorner \varphi \urcorner)$$

and together with the first part, this gives us the desired result.

In the standard model this result says that if $\mathbb{N} \vDash \mathrm{prv}(\#\varphi \leftrightarrow \mathrm{con}_{\mathsf{PA}})$ then we have $\mathbb{N} \vDash \mathrm{prv}(\#\varphi)$ if and only if $\mathbb{N} \vDash \neg\, \mathrm{con}_{\mathsf{PA}}$. Notice that by Lemma 9.16 $\mathbb{N} \nvDash \neg\, \mathrm{con}_{\mathsf{PA}}$ because $\mathsf{PA} \nvdash \mathtt{0} = \mathtt{1}$. Therefore, if some $\mathscr{L}_{\mathsf{PA}}$-formula φ is provably equivalent to the consistency of PA, then we have $\mathbb{N} \nvDash \mathrm{prv}(\#\varphi)$. Intuitively this makes sense, because otherwise we have $\mathsf{PA} \vdash \varphi$ and therefore $\mathsf{PA} \vdash \mathrm{con}_{\mathsf{PA}}$.

11.7 First note:

$$\begin{aligned}
\mathrm{con}^R_{\mathsf{PA}} &\Leftrightarrow_{\mathsf{PA}} \neg\, \mathrm{prv}^R(\ulcorner \mathtt{0} = \mathtt{1} \urcorner) \Leftrightarrow_{\mathsf{PA}} \neg \exists c \big(\mathrm{c_prv}^R(c, \ulcorner \mathtt{0} = \mathtt{1} \urcorner) \big) \\
&\Leftrightarrow_{\mathsf{PA}} \neg \exists c \big(\mathrm{c_prv}(c, \ulcorner \mathtt{0} = \mathtt{1} \urcorner) \wedge \neg \exists c' < c \big(\mathrm{c_prv}(c', \ulcorner \mathtt{0} \neq \mathtt{1} \urcorner) \big) \big) \\
&\Leftrightarrow_{\mathsf{PA}} \forall c \big(\big(\neg\, \mathrm{c_prv}(c, \ulcorner \mathtt{0} = \mathtt{1} \urcorner) \big) \vee \exists c' < c \big(\mathrm{c_prv}(c', \ulcorner \mathtt{0} \neq \mathtt{1} \urcorner) \big) \big)
\end{aligned}$$

where we used Tautologies (Q.0), (L.0) and (F). We will show that $\mathsf{PA} \vdash \forall x \big(\psi(x) \big)$ where

$$\psi(x) :\equiv \big(\neg\, \mathrm{c\_prv}(x, \ulcorner \mathtt{0} = \mathtt{1} \urcorner) \big) \vee \exists c' < x \big(\mathrm{c\_prv}(c', \ulcorner \mathtt{0} \neq \mathtt{1} \urcorner) \big).$$

It is clear that $\mathsf{PA} \vdash \mathrm{prv}(\ulcorner \mathtt{0} \neq \mathtt{1} \urcorner)$. In particular, we find $\tilde{c}$ such that $\mathsf{PA} \vdash \mathrm{c\_prv}(\tilde{c}, \ulcorner \mathtt{0} \neq \mathtt{1} \urcorner)$ and clearly also

$$\mathsf{PA} \vdash \forall x \big(x > \tilde{c} \to \exists c' < x \big(\mathrm{c\_prv}(c', \ulcorner \mathtt{0} \neq \mathtt{1} \urcorner) \big) \big) .$$

Therefore, we obtain $\mathsf{PA} \vdash \forall x > \tilde{c} \, \big(\psi(x) \big)$. Since c_prv is a Δ-formula (see Chapter 10), so is ψ and also $\forall x \leq \tilde{c} \, \big(\psi(x) \big)$. Furthermore, we know that $\mathbb{N} \vDash \forall x \leq \tilde{c} \, \big(\psi(x) \big)$ (because $\mathsf{PA} \nvdash \mathtt{0} = \mathtt{1}$), and thus, $\mathbb{N} \nvDash \mathrm{prv}(\#(\mathtt{0} = \mathtt{1}))$. Hence, by $\mathbb{N}$-conformity we get $\mathsf{PA} \vdash \forall x \leq \tilde{c} \, \big(\psi(x) \big)$, and with the above we finally obtain $\mathsf{PA} \vdash \forall x \big(\psi(x) \big)$.

Chapter 12

12.0 We claim that if $\varphi(x)$ is a quantifier-free formula with $\mathrm{free}(\varphi) = \{x\}$ such that infinitely many x satisfy $\varphi(x)$ (we will show later how to formalise this statement), then $\neg\varphi(x)$ cannot be satisfied by any x. Now, if there is a quantifier-free formula with $\mathrm{free}(\varphi) = \{x\}$ such that $\mathsf{PrA} \vdash \varphi(x) \leftrightarrow \exists y(x = \mathsf{2}y)$, then there are infinitely many witnesses for $\varphi(x)$ as well as for $\neg\varphi(x)$, which contradicts our claim.

In order to prove the claim, suppose $\varphi(x)$ is a quantifier-free formula with $\mathrm{free}(\varphi) = \{x\}$ such that infinitely many x satisfy $\varphi(x)$, i.e., for every $n \in \mathbb{N}$ we have $\mathsf{PrA} \vdash \exists v_0 \cdots \exists v_n \bigwedge_{i \neq j}(v_i \neq v_j \wedge \varphi(v_i))$. Then, every atomic term in φ is either $\mathsf{0}$ or x and every term in φ can be written in the form $\underline{n} + \underline{m}x$ for some $n, m \in \mathbb{N}$. Consequently, every atomic formula in φ is equivalent to an atomic formula of the form $\underline{n} + \underline{m}x = \underline{k} + \underline{l}x$ for some $n, m, k, l \in \mathbb{N}$. If $q := (n-k)/(l-m)$ is a rational number which belongs to $\mathbb{N}$, then

$$\underline{n} + \underline{m}x = \underline{k} + \underline{l}x \Leftrightarrow_{\mathsf{PrA}} x = \underline{q}\,.$$

If $n = k$ and $l = m$, then

$$\underline{n} + \underline{m}x = \underline{k} + \underline{l}x \Leftrightarrow_{\mathsf{PrA}} \mathsf{0} = \mathsf{0}\,.$$

If none of the above conditions are satisfied, then

$$\underline{n} + \underline{m}x = \underline{k} + \underline{l}x \Leftrightarrow_{\mathsf{PrA}} \mathsf{0} = \mathsf{1}\,.$$

Since $\varphi(x)$ is satisfied by infinitely many x, we may replace any occurrence of an atomic formula equivalent to $x = \underline{q}$ by the formula $\mathsf{0} = \mathsf{1}$ and get a formula $\varphi'(x)$ with the same properties. Now, every atomic formula ψ in φ' satisfies $\mathsf{PrA} \vdash \psi$ or $\mathsf{PrA} \vdash \neg\psi$. Since $\mathsf{PrA} \vdash \exists x\,\varphi'(x)$, it follows $\mathsf{PrA} \vdash \varphi'$ and thus, $\mathsf{PrA} \vdash \forall x\,\varphi'(x)$. So, $\neg\varphi'$ cannot be satisfied by any x, which proves the claim.

12.1 To show $\mathsf{PrA} \vdash \tau \equiv_n \tau' \rightarrow \underline{m}\tau \equiv_{mn} \underline{m}\tau'$, we first assume

$$\exists z(\underline{n}z + \tau = \tau')\,.$$

By multiplying the equation with $\underline{m}$, we obtain

$$\exists z(\underline{m} \cdot \underline{n}z + \underline{m}\tau = \underline{m}\tau')\,.$$

Since $\underline{m} \cdot \underline{n} = \underline{mn}$ and multiplication is associative, this implies

$$\underline{m}\tau \equiv_{mn} \underline{m}\tau'\,.$$

Notice that this argument also works if we assume $\exists z(\underline{n}z+\tau' = \tau)$ instead of $\exists z(\underline{n}z + \tau = \tau')$.

For the converse, suppose $\exists z(\underline{mn}z + \underline{m}\tau = \underline{m}\tau')$. By LEMMA 12.0, this leads to $\exists z(\underline{m}(\underline{n}z+\tau) = \underline{m}\tau')$ and $\exists z(\underline{n}z+\tau = \tau')$, which implies $\tau \equiv_n \tau'$. As above, the same argument works for $\exists z(\underline{mn}z + \underline{m}\tau' = \underline{m}\tau)$.

12.2 Let us first assume there is some z such that

$$\underline{m}v + \underline{n} + \underline{k}z = \underline{l}\,,$$

for some $m, n, k, l \in \mathbb{N}$ with $\gcd(m, k) = \mathbf{1}$. By BÉZOUT'S LEMMA there are a and b in $\mathbb{N}$ such that $\underline{am} = \underline{bk} + \mathtt{1}$. Since $a \neq \mathbf{0}$, our assumption is equivalent to

$$\begin{aligned}\underline{al} &= \underline{am}v + \underline{an} + \underline{ak}z\\ &= (\underline{bk} + \mathtt{1})v + \underline{an} + \underline{ak}z\\ &= \underline{k}(\underline{b}v) + \underline{k}(\underline{a}z) + v + \underline{an}\\ &= \underline{k}(\underline{b}v + \underline{a}z) + v + \underline{an}\,.\end{aligned}$$

Equivalently, we obtain

$$v \equiv_k \underline{a(l-n) + Nk}$$

for a sufficiently large number $N \in \mathbb{N}$ so that $al + Nk \geq an$.

If, on the other hand, we start with an equation of the form

$$\underline{m}v + \underline{n} = \underline{l} + \underline{k}z\,,$$

we can argue similarly to obtain

$$\underline{al} + \underline{ak}z = \underline{am}v + \underline{an}\,.$$

Depending whether $\underline{b}v \leq \underline{a}z$ or not, we can rewrite the above equation equivalently to

$$v + \underline{an} = \underline{al} + (\underline{a}z - \underline{b}v)\underline{k} \qquad \text{or} \qquad (\underline{b}v - \underline{a}z)\underline{k} + v + \underline{an} = \underline{al}\,.$$

In both cases this is equivalent to

$$v \equiv_k \underline{a(l-n) + Nk}$$

for a sufficiently large number $N \in \mathbb{N}$ so that $al + Nk \geq an$.

12.3 Let us first introduce the relation symbol $\neq$ by stipulating

$$x \neq y :\Longleftrightarrow \neg(x = y)\,.$$

We will first check that T over the signature $\mathscr{L}' := \mathscr{L} \cup \{\neq\}$ admits quantifier elimination, using THEOREM 12.2.
Observe that part (a) of the theorem is satisfied, since

$$\neg(\tau = \tau') \Leftrightarrow \tau \neq \tau' \quad \text{and} \quad \neg(\tau \neq \tau') \Leftrightarrow \tau = \tau' .$$

For part (b), let $\varphi \equiv \exists\nu(\varphi_1 \wedge \cdots \wedge \varphi_n)$, for some atomic formulae φ_i. Without loss of generality, we may assume that ν is free in every φ_i.

If there exists some φ_i of the form $\nu = \tau$ or $\tau = \nu$ for some term $\tau \not\equiv \nu$, then

$$\varphi \Leftrightarrow_{\mathsf{T}} \varphi_1(\nu/\tau) \wedge \cdots \wedge \varphi_n(\nu/\tau) ,$$

where φ is quantifier-free and has the same free variables as $\varphi_1 \wedge \cdots \wedge \varphi_n$ except ν.

If there is some φ_i of the form $\nu \neq \nu$, then we may replace φ_i by $c_0 = c_1$ and afterwards substitute every occurrence of ν by c_0 and obtain a desired formula equivalent to φ.

Otherwise, all φ_i's are of the form $\nu = \nu$, $\nu \neq \tau$, or $\tau \neq \nu$ for some term $\tau \not\equiv \nu$. So, we may replace $\nu = \nu$ by $c_0 = c_0$, $\nu \neq \tau$ by $\tau = \tau$ and $\tau \neq \nu$ by $\tau = \tau$ and obtain an equivalent formula ψ with the desired properties — in fact, we even have $\mathsf{T} \vdash \psi$.

Thus, THEOREM 12.2 is applicable and yields quantifier elimination for T over $\mathscr{L}'$. Finally, since every quantifier-free $\mathscr{L}'$-formula is equivalent to a quantifier-free $\mathscr{L}$-formula by replacing every occurrence of $\tau \neq \tau'$ by $\neg(\tau = \tau')$, we conclude that T admits quantifier elimination.

12.4 In order to simplify the notation we write just T instead of $\mathbf{Th}(\mathbb{N}, <, \mathsf{s}, \mathbf{0})$ and set $\mathscr{L} := \{<, \mathsf{s}, 0\}$. The proof is based on the QUANTIFIER ELIMINATION THEOREM 12.2: Let φ be an atomic formula. Then φ is either of the form $\tau = \tau'$ or of the form $\tau < \tau'$ for some terms τ and τ'. Since $\mathbb{N}$ is totally ordered under $<$, we have

$$\neg(\tau = \tau') \Leftrightarrow_{\mathsf{T}} \tau < \tau' \vee \tau' < \tau \qquad \text{and} \qquad \neg(\tau < \tau') \Leftrightarrow_{\mathsf{T}} \tau = \tau' \vee \tau' < \tau .$$

This takes care of part (a) in THEOREM 12.2. Now, assume $\varphi \equiv \exists\nu(\varphi_1 \wedge \cdots \wedge \varphi_k)$ for some variable ν and atomic formulae φ_i. Without loss of generality, ν occurs free in every φ_i and by replacing

$$\underbrace{\mathsf{s} \cdots \mathsf{s}}_{n\text{-times}}\tau \quad \text{with} \quad \tau + \underline{n}$$

we may also assume that every term in φ is of the form $\underline{n}$ or $\mu + \underline{n}$ for some $n \in \mathbb{N}$ and some variable μ.

We claim that φ is equivalent to formula to a formula $\exists\nu(\varphi'_1 \wedge \cdots \wedge \varphi'_l)$, where each φ'_i in one of the following forms, where $\mu \not\equiv \nu$ are variables

and $n, m \in \mathbb{N}$:

$$\nu < \mu + \underline{n} \tag{1}$$
$$\mu + \underline{n} < \nu \tag{2}$$

$$\mu < \nu + \underline{n} \tag{3}$$
$$\nu + \underline{n} < \mu \tag{4}$$

Pick any φ_i. If φ_i is of the form

$$\nu + \underline{n} = \nu + \underline{m}\,,$$

we may replace it equivalently by $\underline{n} = \underline{m}$. Similarly, if φ_i is of the form

$$\nu + \underline{n} < \nu + \underline{m}\,,$$

we may replace it by $\underline{n} < \underline{m}$. If φ_i is of the form

$$\nu + \underline{n} = \mu + \underline{m}\,,$$

for some variable $\mu \not\equiv \nu$ and $n \leq m$, we have

$$\varphi \;\Leftrightarrow_{\mathsf{T}}\; \varphi_1(\nu/\mu + \underline{m-n}) \wedge \cdots \wedge \varphi_k(\nu/\mu + \underline{m-n})\,.$$

If, on the other hand, $m < n$, then

$$\nu + \underline{n} = \mu + \underline{m} \;\Leftrightarrow_{\mathsf{T}}\; (\nu + \underline{n-m-\mathbf{1}} < \mu) \wedge (\mu < \nu + \underline{n-m+\mathbf{1}})\,.$$

If φ_i is of the form $\nu + \underline{n} = \underline{m}$ and $n \leq m$, we can replace every instance of ν in every φ_j by $\underline{m-n}$ and we are done. Otherwise, if $m < n$, φ_i can never be satisfied and we can replace it, for example, by $\mathtt{0} < \mathtt{0}$.

If φ_i is of the form $\nu + \underline{n} < \underline{m}$ and $n \leq m$, we have

$$\varphi \;\Leftrightarrow_{\mathsf{T}}\; \bigvee_{l=\mathbf{0}}^{m-n-\mathbf{1}} \big(\varphi_1(\nu/\underline{l}) \wedge \cdots \wedge \varphi_k(\nu/\underline{l})\big)\,.$$

If instead $m < n$, φ_i can never be satisfied and we can replace it by $\mathtt{0} < \mathtt{0}$.

If φ_i is of the form $\underline{m} < \nu + \underline{n}$ and $n \leq m$, we can replace every instance of ν by $\nu + \underline{m-n+\mathbf{1}}$ and replace φ_i by $\mathtt{0} = \mathtt{0}$.

If, on the other hand, $m < n$, then $\mathsf{T} \vdash \varphi_i$. So, we can replace it by $\mathtt{0} = \mathtt{0}$. This completes the proof of the claim.

For the remaining four types of atomic formulae, let us first assume that φ is a conjunction of formulae of type (1) and (2). Without loss of generality, φ is of the form

$$(\nu < \mu_1 + \underline{n_1}) \wedge \cdots \wedge (\nu < \mu_i + \underline{n_i}) \wedge (\lambda_1 + \underline{m_1} < \nu) \wedge \cdots \wedge (\lambda_j + \underline{m_j} < \nu)$$

for $n_p, m_q \in \mathbb{N}$ and variables $\mu_p \not\equiv \nu$ and $\lambda_q \not\equiv \nu$. Then, we have

$$\varphi \Leftrightarrow_{\mathsf{T}} \bigwedge_{p,q} (\lambda_q + \underline{m_q} + 1 < \mu_p + \underline{n_p}) \,. \tag{$*$}$$

To complete the proof, we also have to allow formulas of type (3) and (4). This works out similarly, we just have to put the terms $\underline{m_q}$ and $\underline{n_p}$ on the other side of the inequality for the respective formulas in ($*$). Hence, T admits quantifier elimination.

We still need to prove that addition $+$ is not definable in T. So, let us assume addition is definable in T, i.e., let φ_+ be a formula such that

$$\mathsf{T} \vdash \forall x \forall y \exists! z \, \varphi_+(x, y, z) \quad \text{and} \quad \mathbb{N} \vDash \varphi_+(\underline{m}, \underline{n}, \underline{m+n}) \,.$$

for all $m, n \in \mathbb{N}$. Let $\psi(x) \equiv \exists y (x = y + y) \equiv \exists y \varphi_+(y, y, x)$. On the one hand, since T admits quantifier elimination, there is a quantifier-free $\mathscr{L}$-formula $\psi'(x)$ such that $\psi \Leftrightarrow_{\mathsf{T}} \psi'$. Furthermore, $<$ is definable in PrA, PrA is complete, and $\mathbb{N} \vDash \mathsf{PrA}$. Thus, $\psi'(x) \Leftrightarrow_{\mathsf{PrA}} \psi(x)$. On the other hand, a very similar argument to that given for Exercise 12.0 shows that one cannot find a quantifier-free formula equivalent to $\exists y (x = y + y)$ in PrA over the signature $\{<, \mathsf{s}, 0, +\}$. This is a contradiction and thus, addition $+$ is not definable in T.

12.5 Let φ be an arbitrary $\mathscr{L}$-formula. We have to show that there exists a quantifier-free $\mathscr{L}$-formula ψ, such that $\varphi \Leftrightarrow_{\mathsf{T}} \psi$. By Theorem 2.14, we may assume φ is in PNF. Since we can replace every occurrence of $\forall v_i$ by $\neg \exists v_i \neg$, it suffices to prove the case when $\varphi \equiv \exists \nu \varphi'$ for some quantifier-free formula φ'. Furthermore, by the Disjunctive Normal Form Theorem 2.12 we may assume

$$\begin{aligned} \varphi &\equiv \exists \nu \big((\varphi_{1,1} \wedge \cdots \wedge \varphi_{1,n_1}) \vee \cdots \vee (\varphi_{k,1} \wedge \cdots \wedge \varphi_{k,n_k}) \big) \\ &\Leftrightarrow \big(\exists \nu (\varphi_{1,1} \wedge \cdots \wedge \varphi_{1,n_1}) \big) \vee \cdots \vee \big(\exists \nu (\varphi_{k,1} \wedge \cdots \wedge \varphi_{k,n_k}) \big) \end{aligned}$$

for some formulae $\varphi_{i,j}$ which are atomic or negations of atomic formulae. Thus, without loss of generality we may assume that $k = 1$ and that $\nu \in \mathrm{free}(\varphi_{1,i})$ for every i. Since $\sim$ is reflexive, i.e., $\forall x (x \sim x)$, by L_{14} we may replace any occurrence of $\nu \sim \nu$ and $\nu = \nu$ by $c_0 = c_0$ to obtain an equivalent formula.

Now, if there is a formula $\varphi_{1,i}$ of the form $\nu = \tau$ or $\tau = \nu$ for some term $\tau \not\equiv \nu$, we may replace any instance of ν by τ. Since $\tau \not\equiv \nu$, we are done. Otherwise, note that any $\varphi_{1,i}$ of the form $\neg(\nu = \tau)$ or $\neg(\tau = \nu)$ for some term $\tau \not\equiv \nu$ can be replaced by $c_0 = c_0$, because of the fourth axiom in T.

Thus, we may assume that every $\varphi_{1,i}$ is of one of the following forms (where $\mu \not\equiv \nu$ are variables and $n \in \mathbb{N}$):

$$\nu \sim \mu \tag{1}$$
$$\nu \sim c_n \tag{2}$$
$$\neg(\nu \sim \mu) \tag{3}$$
$$\neg(\nu \sim c_n) \tag{4}$$

If there are no formulae of the form (1) or (2), then $\varphi \Leftrightarrow_{\mathsf{T}} c_0 = c_0$, by the last axiom of T. Otherwise, we construct a formula $\psi \Leftrightarrow_{\mathsf{T}} \varphi$ by taking the conjunction over the following formulae:

$\mu_1 \sim \mu_2$	for every μ_1 and μ_2 in (1)
$c_n \sim c_m$	for every c_n and c_m in (2)
$\mu \sim c_n$	for every μ in (1) and c_n in (2)
$\neg(\mu_1 \sim \mu_2)$	for every μ_1 in (1) and μ_2 in (3)
$\neg(c_n \sim \mu)$	for every c_n in (2) and μ in (3)
$\neg(\mu \sim c_m)$	for every μ in (1) and c_m in (4)
$\neg(c_n \sim c_m)$	for every c_n in (2) and c_m in (4)

This shows that T admits quantifier elimination.

12.6 Suppose we have a relation $\psi_{\mathrm{div}}(x, y)$ in PrA, which coincides with the relation $x \mid y$ in the standard model $\mathbb{N}$. Let

$$\varphi(x) :\equiv \forall y\big(\psi_{\mathrm{div}}(y, x) \to (y = \mathsf{1} \vee y = x)\big)\,,$$

i.e., $\varphi(x)$ holds if and only if $x = \mathsf{1}$ or x is a prime. By Lemma 12.10 we find natural numbers $p > \mathsf{0}$ and n_0 such that:

$$\mathbb{N}^* \vDash \forall n \geq n_0 \left(\varphi(n) \leftrightarrow \varphi(n + p)\right) \tag{$*$}$$

If we now chose a prime $q \geq n_0$, then we have $\mathbb{N}^* \vDash \varphi(q)$ and by successively applying ($*$) we obtain $\mathbb{N}^* \vDash \varphi(q+p), \ldots, \mathbb{N}^* \vDash \varphi(q+qp)$, but since $q + qp = q(\mathsf{1} + p)$ is neither a prime nor $\mathsf{1}$, we have a contradiction.

Chapter 13

13.0 By EXAMPLE 1.0 we know that there exists a set x and that $\forall z(z = z)$. We now apply the Axiom Schema of Separation with respect to the set x and the formula $\varphi(z) :\equiv z \neq z$, and obtain the set $y := \{z \in x : z \neq z\}$, which obviously satisfies $\neg\exists z(z \in x)$, or equivalently, $\forall z(z \notin x)$.

13.1 (a) We define F_α as in the hint. It is clear that F_α is a class function. By the TRANSFINITE RECURSION THEOREM there is a unique class function G_α on Ω such that for all $\beta \in \Omega$:

$$G_\alpha(\beta) = F_\alpha(G|_\beta) = F_\alpha(\{\langle \delta, G(\delta)\rangle : \delta \in \beta\}).$$

In particular we have:

- If $\beta = 0 = \emptyset$, then $G_\alpha(0) = F_\alpha(\emptyset) = \alpha$.
- If $\beta = \gamma + 1$, then

$$G_\alpha(\gamma+1) = F_\alpha(\{\langle \delta, G(\delta)\rangle : \delta \in \gamma+1\}) = G_\alpha(\beta) \cup \{G_\alpha(\beta)\} = G_\alpha(\beta) + 1\,.$$

- If $\beta \in \Omega \setminus \{\emptyset\}$ is a limit ordinal, then

$$G_\alpha(\beta) = F_\alpha(\{\langle \delta, G(\delta)\rangle : \delta \in \beta\}) = \bigcup_{\delta\in\beta} G_\alpha(\delta)\,.$$

Thus, addition of ordinals defined by stipulating $\alpha + \beta := G_\alpha(\beta)$ has the required properties.

(b) We proceed as above but define F_α as follows:

$$F_\alpha(x) = \begin{cases} \emptyset & \text{if } x = \emptyset \\ x(\beta) + \alpha & \text{if } \mathrm{dom}(x) = \beta + 1 \text{ and } \beta \in \Omega \\ \bigcup_{\delta\in\beta} x(\delta) & \text{if } \mathrm{dom}(x) = \beta \text{ and } \beta \in \Omega \setminus \{\emptyset\} \text{ is a limit ordinal} \\ \emptyset & \text{otherwise} \end{cases}$$

(c) We proceed as above but define F_α as follows:

$$F_\alpha(x) = \begin{cases} \{\emptyset\} & \text{if } x = \emptyset \\ x(\beta) \cdot \alpha & \text{if } \mathrm{dom}(x) = \beta + 1 \text{ and } \beta \in \Omega \\ \bigcup_{\delta\in\beta} x(\delta) & \text{if } \mathrm{dom}(x) = \beta \text{ and } \beta \in \Omega \setminus \{\emptyset\} \text{ is a limit ordinal} \\ \emptyset & \text{otherwise} \end{cases}$$

13.2 For part (a), assume towards a contradiction that for some limit ordinal $\alpha \in \Omega$, $\lambda = \bigcup_{\delta\in\alpha} \aleph_\delta$ is such that $|\lambda| < \lambda$ (i.e., λ is not a cardinal). Then, since cardinals are ordinals, $|\lambda| \in \lambda$, this implies that there exists a $\delta \in \beta$ such that $|\lambda| \leq \aleph_\delta < \aleph_{\delta+1} \in \lambda$ which is obviously a contradiction. Notice that λ is the smallest cardinal which is bigger than $\aleph_\delta$ for each $\delta \in \alpha$.

For part (b), assume towards a contradiction that there exists a cardinal κ such that for all ordinals $\alpha \in \Omega$ we have $\aleph_\alpha \neq \kappa$. Let

$$I := \{\alpha \in \Omega : \aleph_\alpha < \kappa\}\,.$$

Then I is a *set* of ordinals, and therefore, the class $\Omega \setminus I$ is non-empty and has a least element, say β_0. If $\beta_0 = \alpha+1$, then $\aleph_\alpha < \kappa$, which implies that $\aleph_\alpha^+ \leq \kappa$, and since $\aleph_\alpha^+ = \aleph_{\alpha+1} = \aleph_{\beta_0}$ and $\beta_0 \notin I$, we have $\kappa = \aleph_{\beta_0}$. If β_0 is a limit ordinal, then $\aleph_\delta < \kappa$ for all $\delta \in \beta_0$, and by the remark above we have $\kappa = \bigcup_{\delta\in\beta_0} \aleph_\delta$, i.e., $\kappa = \aleph_{\beta_0}$.

13.3 Firstly, observe that if $|A| = |A'|$ and $|B| = |B'|$, then we have $|A \cup B| = |A' \cup B'|$, $|A \times B| = |A' \times B'|$, and $|^B A| = |^{B'} A'|$, where in the first case we additionally assume that A and B as well as A' and B' are disjoint.

We first show that addition is commutative and associative: For this, let $A := \kappa \times \{0\}$ and $B := \lambda \times \{1\}$, and notice that $\kappa + \lambda = |A \dot\cup B|$, where $\dot\cup$ is the disjoint union. Therefore, we have

$$\kappa + \lambda = |A \dot\cup B| = |B \dot\cup A| = \lambda + \kappa\,,$$

which shows that addition is commutative, and for $C := \mu \times \{2\}$ we have

$$\kappa + (\lambda + \mu) = |A \dot\cup (B \dot\cup C)| = |(A \dot\cup B) \dot\cup C| = (\kappa + \lambda) + \mu\,,$$

which shows that addition is associative. Note that we have used our observation above, e.g. in order to derive $\mu = |C|$.

Similarly, we can show that multiplication is commutative and associative: We have

$$\kappa \cdot \lambda = |\kappa \times \lambda| = |\lambda \times \kappa| = \kappa \cdot \lambda\,,$$

which shows that multiplication is commutative, and

$$\kappa \cdot (\lambda \cdot \mu) = |\kappa \times (\lambda \times \mu)| = |(\kappa \times \lambda) \times \mu| = (\kappa \cdot \lambda) \cdot \mu\,,$$

which shows that multiplication is associative.

For distributivity, let again $A := \kappa \times \{0\}$, $B := \lambda \times \{1\}$, and $C := \mu \times \{2\}$. Then we have

$$\kappa \cdot (\lambda + \mu) = |A \times (B \dot\cup C)| = |(A \times B) \dot\cup (A \times C)| = \kappa \cdot \lambda + \kappa \cdot \mu\,.$$

For the remaining equalities let again $B := \lambda \times \{1\}$ and $C := \mu \times \{2\}$. Then we have:

$$\kappa^{\lambda+\mu} = |{}^{\lambda+\mu}\kappa| = |{}^{B \dot\cup C}\kappa| = |{}^{B}\kappa \times {}^{C}\kappa| = |{}^{B}\kappa| \cdot |{}^{C}\kappa| = \kappa^{\lambda} \cdot \kappa^{\mu}$$

$$\kappa^{\lambda\cdot\mu} = |{}^{\lambda\cdot\mu}\kappa| = |{}^{\lambda\times\mu}\kappa| = |{}^{\mu}({}^{\lambda}\kappa)| = |{}^{\mu}(\kappa^{\lambda})| = (\kappa^{\lambda})^{\mu}$$

$$(\kappa \cdot \lambda)^{\mu} = |{}^{\mu}(\kappa \times \lambda)| = |{}^{\mu}\kappa \times {}^{\mu}\lambda| = |\kappa^{\mu} \times \lambda^{\mu}| = \kappa^{\mu} \cdot \lambda^{\mu}$$

Note that we have $|{}^{B \dot\cup C}\kappa| = |{}^{B}\kappa \times {}^{C}\kappa|$ because we can bijectively map a pair $\langle f, g\rangle \in {}^{B}\kappa \times {}^{C}\kappa$ to $h \in {}^{B \dot\cup C}\kappa$ by stipulating

$$h(x) = \begin{cases} f(x), & x \in B \\ g(x), & x \in C \end{cases}$$

With similar bijections one can also show $|{}^{\lambda\times\mu}\kappa| = |{}^{\mu}({}^{\lambda}\kappa)|$ and $|{}^{\mu}\kappa \times {}^{\mu}\lambda| = |\kappa^{\mu} \times \lambda^{\mu}|$.

13.4 Let κ be an infinite cardinal, i.e., $\kappa \geq \aleph_0$.

(a) By definition, $\operatorname{seq}(\kappa)$ is the set of all finite sequences which can be built with elements of κ. Note that we have $|\kappa^n| = |\kappa|$ for all $n \in \omega$ and hence there is a bijection $f_n : \kappa^n \to \kappa$. This implies

$$|\operatorname{seq}(\kappa)| = \left|\left\{s \in {}^{n}\kappa : n \in \omega\right\}\right| = \left|\bigcup_{n\in\omega} {}^{n}\kappa\right| = \left|\bigcup_{n\in\omega} \kappa^{n}\right| = \aleph_0 \cdot \kappa = \kappa,$$

where for the penultimate equality we used the bijection

$$s \mapsto \langle n, f_n(s)\rangle$$

and for the last equality we used that there are injections

$$\kappa \hookrightarrow (\aleph_0 \times \kappa) \hookrightarrow (\kappa \times \kappa),$$

which implies

$$|\kappa| \leq |\aleph_0 \times \kappa| \leq |\kappa^2| = |\kappa| = \kappa.$$

(b) On the one hand, we obviously have $\kappa \leq |\operatorname{fin}(\kappa)|$. On the other hand, every finite subset $s = \{s_0, \ldots, s_{n-1}\} \subseteq \kappa$ can be ordered such that $s_0 < \cdots < s_{n-1}$, and therefore, s corresponds to a unique finite sequence $\langle s_0, \ldots, s_{n-1}\rangle$. This gives us an injection $\operatorname{fin}(\kappa) \hookrightarrow \operatorname{seq}(\kappa)$ and by (a) we have

$$\kappa \leq |\operatorname{fin}(\kappa)| \leq |\operatorname{seq}(\kappa)| = \kappa,$$

which shows that $|\operatorname{fin}(\kappa)| = \kappa$.

13.5 The domain of the model $\mathbb{N} \vDash \mathsf{PA}$ we construct within Z is the set ω. Notice that for the construction of ω we need neither the Axiom Schema of Replacement nor the Axiom of Foundation. Recall that $\mathscr{L}_{\mathsf{PA}} = \{\mathbf{0}, \mathbf{s}, +, \cdot\}$. First, we define $\mathbf{0}^{\mathbb{N}}$, $\mathbf{s}^{\mathbb{N}}$, $+^{\mathbb{N}}$, and $\cdot^{\mathbb{N}}$ as follows:

$$\mathbf{0}^{\mathbb{N}} := \emptyset$$

$$\mathbf{s}^{\mathbb{N}} := \big\{\langle n, m\rangle \in \omega \times \omega : m = n \cup \{n\}\big\}$$

$$+^{\mathbb{N}} := \bigcap \Big\{ f \in \mathscr{P}\big((\omega \times \omega) \times \omega\big) : \forall x \in \omega \left(\big\langle\langle x, \emptyset\rangle, x\big\rangle \in f\right) \wedge$$
$$\forall x \forall y \forall z \forall z' \big(\big\langle\langle x, y\rangle, z\big\rangle \in f \wedge \big\langle\langle x, y\rangle, z'\big\rangle \in f \to z = z'\big) \wedge$$
$$\forall x \forall y \forall z \big(\big\langle\langle x, y\rangle, z\big\rangle \in f \to \big\langle\langle x, \mathbf{s}^{\mathbb{N}}(y)\rangle, \mathbf{s}^{\mathbb{N}}(z)\big\rangle \in f\big)\Big\}$$

$$\cdot^{\mathbb{N}} := \bigcap \Big\{ g \in \mathscr{P}\big((\omega \times \omega) \times \omega\big) : \forall x \in \omega \left(\big\langle\langle x, \emptyset\rangle, \emptyset\big\rangle \in g\right) \wedge$$
$$\forall x \forall y \forall z \forall z' \big(\big\langle\langle x, y\rangle, z\big\rangle \in g \wedge \big\langle\langle x, y\rangle, z'\big\rangle \in g \to z = z'\big) \wedge$$
$$\forall x \forall y \forall z \big(\big\langle\langle x, y\rangle, z\big\rangle \in g \to \big\langle\langle x, \mathbf{s}^{\mathbb{N}}(y)\rangle, z +^{\mathbb{N}} x\big\rangle \in g\big)\Big\}$$

We have to check that $\mathbf{s}^{\mathbb{N}}$, $+^{\mathbb{N}}$, and $\cdot^{\mathbb{N}}$ are functions and that the structure $\mathbb{N} = \big(\omega, \mathbf{0}^{\mathbb{N}}, \mathbf{s}^{\mathbb{N}}, +^{\mathbb{N}}, \cdot^{\mathbb{N}}\big)$ is a model of PA.

$\mathbf{s}^{\mathbb{N}}$ *is a unary function with domain* ω: By definition of $\mathbf{s}^{\mathbb{N}}$, for every $n \in \omega$ there is a unique $m \in \omega$ such that $\langle n, m\rangle \in \mathbf{s}^{\mathbb{N}}$.

$+^{\mathbb{N}}$ *is a binary function with domain* $\omega \times \omega$: First notice that by definition of the set $+^{\mathbb{N}}$, for all $x, y \in \omega$ there exists at most one $z \in \omega$ such that $\big\langle\langle x, y\rangle, z\big\rangle \in +^{\mathbb{N}}$. Thus, if $+^{\mathbb{N}} \neq \emptyset$, then $+^{\mathbb{N}}$ is a function. Now, we show that for all $\langle x, y\rangle \in \omega \times \omega$ there exists a unique $z \in \omega$ such that $\big\langle\langle x, y\rangle, z\big\rangle \in +^{\mathbb{N}}$. For this, let

$$\varphi(y) :\equiv \forall x \in \omega \, \exists! z \in \omega \big(\big\langle\langle x, y\rangle, z\big\rangle \in +^{\mathbb{N}}\big)\,.$$

By definition of $+^{\mathbb{N}}$, we have $\varphi(\emptyset)$ and for each $y \in \omega$ we have $\varphi(y) \to \varphi(\mathbf{s}^{\mathbb{N}}(y))$, where $\mathbf{s}^{\mathbb{N}}(y) = y \cup \{y\}$. If there exists a $y \in \omega$ such that $\neg\varphi(y)$, then, since ω is well-ordered, there exists a $y_0 \in \omega$ such that $\neg\varphi(y_0)$ and $\forall y < y_0 \, \varphi(y)$. Since $y_0 \neq \emptyset$, we have $y_0 = y_1 \cup \{y_1\}$ for some $y_1 \in \omega$. Now, since $y_1 < y_0$, we have $\varphi(y_1)$ and by $\varphi(y_1) \to \varphi(y_0)$ we get $\varphi(y_0)$, which is a contradiction to the choice of y_0. Thus, we have $\forall y \in \omega \left(\varphi(y)\right)$, which shows that

$$\forall x \in \omega \, \forall y \in \omega \, \exists! z \in \omega \big(\big\langle\langle x, y\rangle, z\big\rangle \in +^{\mathbb{N}}\big)\,,$$

and therefore, $+^{\mathbb{N}}$ is a function.

$\cdot^{\mathbb{N}}$ *is a binary function with domain* $\omega \times \omega$: The proof is essentially the same as the proof for $+^{\mathbb{N}}$.

$\mathbb{N} \vDash \mathsf{PA}_0$: Since for all $x \in \omega$ we have $\mathbf{s}^{\mathbb{N}}(x) = x \cup \{x\}$ and $x \cup \{x\} \neq \emptyset$, there is no $x \in \omega$ with $\mathbf{s}^{\mathbb{N}}(x) = \emptyset$.

$\mathbb{N} \vDash \mathsf{PA}_1$: If $\mathbf{s}^{\mathbb{N}}(x) = \mathbf{s}^{\mathbb{N}}(y)$, then $x \cup \{x\} = y \cup \{y\}$, which implies that either $x = y$ or $x \in y \in x$. In the former case, we are done, and in the latter case, we obtain that $x \in x$ and therefore $\{x\} \subseteq x$ (since x is transitive), which contradicts the fact that x is well-ordered by $\in$ (since $\{x\} \subseteq x$ does not have an $\in$-minimal element).

$\mathbb{N} \vDash \mathsf{PA}_2$: By definition we have

$$\forall x \in \omega \big(x +^{\mathbb{N}} \emptyset = x \big) .$$

$\mathbb{N} \vDash \mathsf{PA}_3$: Notice that if $x +^{\mathbb{N}} y = z$, then $\mathbf{s}^{\mathbb{N}}(x +^{\mathbb{N}} y) = \mathbf{s}^{\mathbb{N}} z$. Now, since $+^{\mathbb{N}}$ is a function and

$$x +^{\mathbb{N}} y = z \;\rightarrow\; x +^{\mathbb{N}} \mathbf{s}^{\mathbb{N}}(y) = \mathbf{s}^{\mathbb{N}} z ,$$

we obtain $\forall x \in \omega \forall y \in \omega \big(\mathbf{s}^{\mathbb{N}}(x +^{\mathbb{N}} y) = x +^{\mathbb{N}} \mathbf{s}^{\mathbb{N}}(y) \big)$.

$\mathbb{N} \vDash \mathsf{PA}_4$: By definition we have

$$\forall x \in \omega \big(x \cdot^{\mathbb{N}} \emptyset = \emptyset \big) .$$

$\mathbb{N} \vDash \mathsf{PA}_5$: This is similar to $\mathbb{N} \vDash \mathsf{PA}_3$.

$\mathbb{N} \vDash \mathsf{PA}_6$: Let φ be an $\mathscr{L}_{\mathsf{PA}}$-formula with $\mathrm{free}(\varphi) = \{x\}$ and assume

$$\varphi(\emptyset) \wedge \forall x \in \omega \, \big(\varphi(x) \rightarrow \varphi(\mathbf{s}^{\mathbb{N}}(x)) \big) .$$

If the set $A := \big\{ x \in \omega : \neg\varphi(x) \big\} \subseteq \omega$ is non-empty, then, since ω is well-ordered, it contains a least element, say x_0. Now, since $\varphi(\emptyset)$ and $\neg\varphi(x_0)$, we have $x_0 \neq \emptyset$, and therefore $x_0 = \mathbf{s}^{\mathbb{N}}(x_1)$ for some $x_1 \in \omega$ with $\varphi(x_1)$, which contradicts the assumption $\varphi(x_1) \rightarrow \varphi(\mathbf{s}^{\mathbb{N}}(x_1))$. Hence, the set A is empty, i.e., $\forall x \in \omega \, \big(\varphi(x) \big)$, which shows that $\mathbb{N} \vDash \mathsf{PA}_6$.

13.6 The proof is taken from Weiss [57, Ch. 4]. Assume towards a contradiction that there is a set x_0 with $\mathrm{trans}^*(x_0)$ such that the set

$$S := \big\{ y_2 \in x_0 : \exists y_1 \in x_0 \, (y_1 \subseteq y_2 \wedge y_1 \neq y_2 \wedge y_1 \notin y_2) \big\}$$

is non-empty.

Using the Axiom of Foundation we can find $y_2 \in x_0$ be such that $y_2 \in S$ and $y_2 \cap S = \emptyset$. Now, let $y_1 \in x_0$ be such that $y_1 \subseteq y_2$, $y_1 \neq y_2$, and $y_1 \notin y_2$. By the Axiom of Foundation there is a $z \in y_2 \setminus y_1$ with $z \cap (y_2 \setminus y_1) = \emptyset$. Thus, by $\mathrm{trans}(y_2)$ we have $z \subseteq y_2$, and since $z \cap (y_2 \setminus y_1) = \emptyset$ we obtain $z \subseteq y_1$. Furthermore, since $z \in y_2$ and $y_1 \notin y_2$, we have $z \neq y_1$, and from

$z \subseteq y_1$ we obtain $y_1 \setminus z \neq \emptyset$. Applying the Axiom of Foundation again, we find $u \in y_1 \setminus z$ with $u \cap (y_1 \setminus z) = \emptyset$. Since trans$(y_1)$, we have $u \subseteq y_1$ and with $u \cap (y_1 \setminus z) = \emptyset$ we obtain $u \subseteq z$. Now, since $z \in y_2$ and $y_2 \cap S = \emptyset$, we have $z \notin S$, and since $z \in x_0$ by trans(x_0) and $u \subseteq z$, we therefore have $u = z$ or $u \in z$. However, $u = z$ contradicts $z \in y_2 \setminus y_1$ (i.e., $u \notin y_1$) and $u \in y_1 \setminus z$ (i.e., $u \in y_1$), and $u \in z$ contradicts $u \in y_1 \setminus z$.

Thus, the set S is empty which shows that for all sets x with trans$^*(x)$ we have:

$$\forall y_1, y_2 \in x \left(y_1 \subseteq y_2 \to (y_1 = y_2 \vee y_1 \in y_2)\right)$$

13.7 We first prove $(a) \Rightarrow (b)$: If ordinal(x), then by definition we have trans(x). Let $y \in x$, and let $z_1 \in z_2 \in y$. We have to show that $z_1 \in y$. Since trans(x), we have that $z_2 \in x$ and therefore also $z_1 \in x$. Now, since ordinal(x), we have ord$_\in(x)$, which implies that either $z_1 \in y$, or $z_1 = y$, or $z_1 \ni y$. If $z_1 \in y$, we are done, and if $z_1 = y$ or $y \in z_1$, then the set $\{y, z_2, z_1\} \subseteq x$ does not have an $\in$-minimal element, which contradicts the fact that x is well-ordered by $\in$.

For $(b) \Rightarrow (c)$, assume trans$^*(x)$ and let $y_1, y_2 \in x$. Notice that since trans(x) we have $y_1 \subsetneq x$ and $y_2 \subsetneq x$. We have to show that either $y_1 \in y_2$, or $y_1 = y_2$, or $y_1 \ni y_2$. We consider the following two cases:

$y_1 \subseteq y_2$: By Exercise 13.6 we have either $y_1 = y_2$ or $y_1 \in y_2$.

$y_1 \not\subseteq y_2$: With Axiom of Foundation we can find $z \in y_1 \setminus y_2$ such that $z \cap (y_1 \setminus y_2) = \emptyset$. Since trans$^*(x)$ and $z \in y_1 \in x$, we have $z \subseteq x$. Furthermore, since trans(y_1) and $z \in y_1$, we have $z \subseteq y_1$, and since $z \cap (y_1 \setminus y_2) = \emptyset$, we have $z \subseteq y_2$. Thus, we have $z, y_2 \in x$ and $z \subseteq y_2$, and by Exercise 13.6 we obtain either $z = y_2$ or $z \in y_2$. Now, $z \in y_2$ is not possible since $z \in y_1 \setminus y_2$, and therefore we have $z = y_2$, i.e., $y_2 \in y_1$.

Thus, in both cases we have either $y_1 \in y_2$, or $y_1 = y_2$, or $y_1 \ni y_2$.

For $(c) \Rightarrow (a)$ we suppose towards a contradiction

$$\exists y \subseteq x \big(y \neq \emptyset \wedge \neg \exists z \min_\in(z, y)\big) .$$

We thus pick such a non-empty set $y \subseteq x$. Note that by the choice of y we have $\forall z \in y \exists z' \big(z' \in z \wedge z' \in y\big)$, or equivalently,

$$\forall z \in y \exists z' \big(z' \in (z \cap y)\big) .$$

Since y is non-empty, we find $z_0 \in y$, and by the choice of y, we then find $z_1 \in z_0 \cap y$. By the Axiom of Choice, we can then repeat this process indefinitely by finding for each $i \in \omega$ a $z_{i+1} \in z_i \cap y$. This gives an infinite decreasing sequence $z_0 \ni z_1 \ni \ldots$ which contradicts the Axiom of Foundation.

13.8 First we show $|\mathbb{R}| = |(0,1)|$, where $(0,1) = \{r \in \mathbb{R} : 0 < r < 1\}$: For this, consider the function

$$\begin{aligned} f : \mathbb{R} &\longrightarrow (0,1) \\ x &\longmapsto \tfrac{1}{2^x+1} \end{aligned}$$

which is obviously a bijection between $\mathbb{R}$ and the interval $(0,1)$.

Now, we show $|\mathscr{P}(\omega)| \leq |(0,1)|$: For this, consider the function

$$\begin{aligned} g : \mathscr{P}(\omega) &\longrightarrow \quad (0,1) \\ x \quad &\longmapsto \sum_{n \in x} 3^{-(n+1)} \end{aligned}$$

where $g(\emptyset) := \frac{3}{4}$, which is obviously an injection from $\mathscr{P}(\omega)$ into the interval $(0,1)$.

Finally, we show $|(0,1)| \leq |\mathscr{P}(\omega)|$: For this, we use the fact that every positive real number $r \in \mathbb{R}$ has a unique representation as a finite or infinite continued fraction, i.e.,

$$r = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \cfrac{1}{a_5 + \cdots}}}}}$$

where $a_n \in \omega$ and $a_n > 0$ for all $n \geq 1$. For simplicity, we write $r = [a_0, a_1, a_2, \ldots]$. Notice that every $r \in (0,1)$ has a continued fraction which starts with $a_0 = 0$. Now, for each $r \in (0,1)$ with finite or infinite continued fraction $[0, a_1, a_2, \ldots]$, we assign a finite or infinite set $x_r \subseteq \omega$ by stipulating

$$x_r := \left\{0,\, a_1,\, a_1 + a_2,\, a_1 + a_2 + a_3,\, \ldots, a_1 + a_2 + \ldots + a_n,\, \ldots\right\}.$$

Then, by the uniqueness of the representation of real numbers by continued fraction, the mapping which maps real numbers $r \in (0,1)$ to subsets $x_r \subseteq \omega$ is injective.

Now, combining these results we have $|\mathbb{R}| = |(0,1)| = |\mathscr{P}(\omega)|$, and by Cantor's Theorem 13.8 we get that $\mathbb{R}$ is uncountable.

Chapter 14

14.0 (a) First, notice that V_0 is transitive. Now, if α is a non-empty limit ordinal and for each $\beta \in \alpha$, V_β is transitive, then $\bigcup_{\beta\in\alpha} \mathrm{V}_\beta$ is obviously transitive, too. Finally, if $\alpha = \beta + 1$ and V_β is transitive, then $\mathrm{V}_\alpha = \mathscr{P}(\mathrm{V}_\beta)$, and for all sets x and y, such that $y \in x \in \mathrm{V}_\alpha$ we have $y \in x \subseteq \mathrm{V}_\beta$. Hence, $y \in \mathrm{V}_\beta$, and since V_β is transitive, we get $y \subseteq \mathrm{V}_\beta$ which shows that $y \in \mathrm{V}_\alpha$. Therefore, by transfinite induction we get that for each $\alpha \in \Omega$, V_α is transitive.

(b) Since arbitrary unions of transitive sets are transitive, this follows immediately from (a).

(c) If $\beta = \alpha+1$, then, since V_β is transitive, $\mathrm{V}_\alpha \subseteq \mathrm{V}_\beta$, and since $\mathrm{V}_\alpha \in \mathrm{V}_\beta$ but $\mathrm{V}_\alpha \notin \mathrm{V}_\alpha$, we get $\mathrm{V}_\alpha \subsetneq \mathrm{V}_\beta$. Now, if β is a limit ordinal and $\alpha \in \beta$, then $\mathrm{V}_\alpha \subsetneq \mathrm{V}_{\alpha+1} \subsetneq \mathrm{V}_\beta$.

(d) If $\alpha = 0$, then we obviously have $\alpha \subseteq \mathrm{V}_\alpha$ and $\alpha \in \mathrm{V}_{\alpha+1}$. Now, let α_0 be an ordinal and assume that for all $\beta \in \alpha_0$ we have $\beta \subseteq \mathrm{V}_\beta$ and $\beta \in \mathrm{V}_{\beta+1}$. If α_0 is a limit ordinal, then the assumption implies $\alpha_0 \subseteq \mathrm{V}_{\alpha_0}$, and consequently we get $\alpha_0 \in \mathrm{V}_{\alpha_0+1}$. Finally, if $\alpha_0 = \beta + 1$, then, by the assumption, $\beta \in \mathrm{V}_{\alpha_0}$, and since V_{α_0} is transitive, $\alpha_0 \subseteq \mathrm{V}_{\alpha_0}$, and consequently $\alpha_0 \in \mathrm{V}_{\alpha_0+1}$. Therefore, by transfinite induction we get that for each $\alpha \in \Omega$, $\alpha \subseteq \mathrm{V}_\alpha$ and $\alpha \in \mathrm{V}_{\alpha+1}$.

14.1 For (a) assume that $x \in H_\kappa$ and let $\alpha = \mathrm{rk}(x)$, where $\mathrm{rk}(x)$ is defined as in the hint. Now, consider the set

$$A := \{\mathrm{rk}(y) : y \in \mathrm{TC}(x)\}.$$

We will show that $A = \alpha$. Then $|\mathrm{TC}(x)| < \kappa$ implies $|\alpha| = |A| < \kappa$ and hence $\alpha < \kappa$, which proves the claim.

Obviously, we have $A \subseteq \alpha$. Hence, it remains to check $\alpha \subseteq A$. Suppose towards a contradiction that there is a $\beta < \alpha$ such that $\beta \notin A$. Then the set

$$B := \{y \in \mathrm{TC}(\{x\}) : \mathrm{rk}(y) > \beta \wedge \neg\exists z \in y(\mathrm{rk}(z) = \beta)\}$$

is non-empty (since $\alpha \in B$). By the Axiom of Foundation, there is a $y \in B$ such that $y \cap B = \emptyset$. Since $\mathrm{rk}(y) > \beta$, there is a $z \in y$ with $\mathrm{rk}(z) \geq \beta$. Now there are two cases:

$\mathrm{rk}(z) = \beta$: This clearly contradicts our assumption that $y \in B$.

$\mathrm{rk}(z) > \beta$: Then $z \notin B$, so there is $w \in z$ such that $\mathrm{rk}(w) = \beta$. Since $\mathrm{TC}(\{x\})$ is transitive, we have $w \neq x$, we have $w \in \mathrm{TC}(x)$ and therefore $\beta = \mathrm{rk}(w) \in A$, a contradiction.

Now, we turn to (b). We set $\kappa_1 = \aleph_0 = \omega$ and $\kappa_2 = \aleph_1$. Firstly, we have $H_\omega \subseteq \mathrm{V}_\omega$ by (a). For the other direction, one can easily prove by induction that $\mathrm{V}_n \subseteq H_\omega$ for each $n \in \omega$, which shows $\mathrm{V}_\omega \subseteq H_\omega$.

For the other statement, observe that $\mathrm{V}_{\omega+1} \in \mathrm{V}_{\aleph_1}$ by Fact 14.0. However, since V_ω is countably infinite, by Cantor's Theorem we have $|\mathrm{V}_{\omega+1}| = |\mathscr{P}(\mathrm{V}_\omega)| \geq \aleph_1$, and hence, $\mathrm{V}_{\omega+1} \notin H_{\aleph_1}$, which implies that $H_{\aleph_1} \neq \mathrm{V}_{\aleph_1}$.

14.2 Since $H_\omega = H_{\aleph_0}$ is a transitive subset of $\mathbf{V}$ and subsets, finite unions and power sets of hereditarily finite sets are hereditarily finite sets, all the axioms of ZFC, except the Axiom of Infinity, hold in the structure $(H_\omega, \in)$.

On the other hand, $(H_{\aleph_1}, \in)$ satisfies the Axiom of Infinity because $\omega \in H_{\aleph_1}$. However, the Axiom of Power Set is not fulfilled because $\omega \in H_{\aleph_1}$ but $\mathscr{P}(\omega) \notin H_{\aleph_1}$ by Cantor's Theorem. The other axioms hold due to similar arguments as for $(H_\omega, \in)$.

14.3 Since $\mathrm{V}_{\omega+\omega}$ is transitive and $\emptyset$ and ω belong to $\mathrm{V}_{\omega+\omega}$, the set $\mathrm{V}_{\omega+\omega}$ is obviously a model for the Axiom of Extensionality, the Axiom of Empty Set, the Axiom of Pairing, the Axiom Schema of Separation, the Axiom of Union, the Axiom of Choice, and the Axiom of Infinity. To see that $\mathrm{V}_{\omega+\omega}$ is also a model for the Axiom of Power Set, notice that for each $x \in \mathrm{V}_{\omega+\omega}$ there is an $\alpha \in \omega + \omega$ such that $x \in \mathrm{V}_\alpha$. Now, since V_α is transitive, for all $y \in x$ we have $y \in \mathrm{V}_\alpha$, and since $\mathrm{V}_{\alpha+1} = \mathscr{P}(\mathrm{V}_\alpha)$, we obtain $\mathscr{P}(x) \subseteq \mathrm{V}_{\alpha+1}$, and therefore $\mathscr{P}(x) \in \mathrm{V}_{\alpha+2}$, and since $\mathrm{V}_{\alpha+2} \subseteq \mathrm{V}_{\omega+\omega}$, we have $\mathscr{P}(x) \in \mathrm{V}_{\omega+\omega}$.

Assume towards a contradiction that the Axiom Schema of Replacement holds in $\mathrm{V}_{\omega+\omega}$. Let F be the class function defined by stipulating

$$F(n) := \begin{cases} \mathrm{V}_{\omega+n} & \text{if } n \in \omega, \\ \emptyset & \text{otherwise.} \end{cases}$$

Since $\omega \in \mathrm{V}_{\omega+\omega}$, by the Axiom Schema of Replacement we get that the set $A := \{F(n) : n \in \omega\}$ is a set in $\mathrm{V}_{\omega+\omega}$. Now, since the Axiom of Union holds in $\mathrm{V}_{\omega+\omega}$, we get that also $\bigcup A = \mathrm{V}_{\omega+\omega}$ is in $\mathrm{V}_{\omega+\omega}$, which is obviously a contradiction.

14.4 Notice first that if ZF is equivalent to a finite set of axioms, then ZF is equivalent to one single axiom, say σ_{ZF}. By replacing L_α by V_α in Lévy's Reflection Theorem 14.8, we obtain that there is a $\alpha_0 \in \Omega$ such that σ_{ZF} is absolute between V_{α_0} and $\mathbf{V}$. In particular, we have $\mathrm{V}_{\alpha_0} \vDash \sigma_{\mathsf{ZF}}$, or equivalently,

$$\mathrm{V}_{\alpha_0} \vDash \mathsf{ZF}\,.$$

Since $\mathrm{V}_{\alpha_0} \vDash \mathsf{ZF}$, by Lévy's Reflection Theorem 14.8, there is an ordinal $\alpha_1 \in \mathrm{V}_{\alpha_0}$ (i.e., $\alpha_1 \in \alpha_0$), such that σ_{ZF} is absolute between V_{α_1} and V_{α_1}. Thus, we have $\mathrm{V}_{\alpha_1} \vDash \mathsf{ZF}$, where $\alpha_0 \ni \alpha_1$. Proceeding this way, we obtain an infinite decreasing sequence $\alpha_0 \ni \alpha_1 \ni \ldots \ni \alpha_n \ni \ldots$ of ordinals, which is a contradiction to the Axiom of Foundation.

14.5 (a) We first show by transfinite induction that for all ordinals $\alpha \in \kappa$, $|\mathrm{V}_\alpha| < \kappa$: If $\alpha = 0$, then $|\mathrm{V}_0| = 0 < \kappa$. Now, let $\alpha \in \kappa$ and assume that $|\mathrm{V}_\alpha| = \lambda$ for some $\lambda < \kappa$. Then $|\mathrm{V}_{\alpha+1}| = |\mathscr{P}(\mathrm{V}_\alpha)| = 2^\lambda$ and by (1) we have $2^\lambda < \kappa$, thus $|\mathrm{V}_{\alpha+1}| < \kappa$. Finally, assume that $\delta \in \kappa$ is a limit ordinal and for all $\alpha \in \delta$ we have $|\mathrm{V}_\alpha| = \lambda_\alpha < \kappa$, then, since $\delta \in \kappa$, for the set

$$A := \big\{\lambda_\alpha : \alpha \in \delta\big\} \subseteq \kappa$$

we have $|A| \leq |\delta| < \kappa$. Thus, by (2) we have $\lambda_\delta := \bigcup A \in \kappa$. In particular, by the Solution to Exercise 13.2, λ_δ is a cardinal with $\lambda_\delta < \kappa$, and by the definition of λ_δ we obtain

$$\Big|\bigcup_{\alpha\in\delta} \mathrm{V}_\alpha\Big| \leq |\delta| \cdot \lambda_\delta < \kappa\,.$$

We are now ready to show that V_κ is a model for the Axiom Schema of Replacement, the other axioms can be shown similarly to the Solution to Exercise 14.3. Let $F : \mathrm{V}_\kappa \to \mathrm{V}_\kappa$ be a function and let $X \in \mathrm{V}_\kappa$ be a set. Then there is an ordinal $\alpha \in \kappa$ such that $X \in \mathrm{V}_\alpha$ and for each $y \in X$ there is a $\beta \in \kappa$ such that $F(y) \in \mathrm{V}_\beta$. Let $G : X \to \kappa$ be defined by stipulating

$$G(y) := \min\big\{\beta \in \kappa : F(y) \in \mathrm{V}_\beta\big\} \quad \text{and let} \quad A := \big\{G(y) : y \in X\big\}\,.$$

Recall that by Theorem 13.1.(f), the class of ordinals is well-ordered by $\in$, which shows that the function G is well-defined. Now, since $X \in \mathrm{V}_\alpha$, $|\mathrm{V}_\alpha| < \kappa$ and $|A| \leq |\mathrm{V}_\alpha|$, we have $|A| < \kappa$, and since κ is inaccessible and $A \subseteq \kappa$, by (2) we obtain $|\bigcup A| < \kappa$. Thus, for the ordinal $\lambda := \bigcup A$ we have $\lambda \in \kappa$ and $F[X] \in \mathrm{V}_\lambda$, which shows that $F[X]$ is a set in V_κ.

(b) If ZFC is consistent, then there is a model $\mathbf{V} \vDash \mathsf{ZFC}$. Assume towards a contradiction that

$$\mathsf{ZFC} \vdash \text{there exists an inaccessible cardinal}\,.$$

Then there exists an inaccessible cardinal $\kappa_0 \in \mathbf{V}$ and by (a) we have $\mathrm{V}_{\kappa_0} \vDash \mathsf{ZFC}$. Hence, by our assumption we find an inaccessible cardinal $\kappa_1 \in \mathrm{V}_{\kappa_0}$, and by (a) we have $\mathrm{V}_{\kappa_1} \vDash \mathsf{ZFC}$. Proceeding this way, as in the Solution to Exercise 14.4 we obtain an infinite decreasing sequence $\kappa_0 > \kappa_1 > \ldots$ of cardinals, which is a contradiction to the Axiom of Foundation.

Chapter 15

15.0 (a) If $\mathscr{F}$ is a filter over S which contains F, then, since $\emptyset \notin \mathscr{F}$, no finite intersection of elements of $\mathscr{F}$ is empty, in particular, no finite intersection of elements of F is empty.

For the other direction assume that no finite intersection of elements of F is empty and define

$$\mathscr{F} := \left\{ y \subseteq S : \bigcap_{i=0}^{n} x_i \subseteq y \text{ for some } \{x_0, \ldots, x_n\} \subseteq F \right\},$$

i.e., $\mathscr{F}$ consists of all supersets of finite intersections of elements of F. By the properties of F, $\mathscr{F}$ does not contain the empty set and is closed under finite intersections and supersets. Thus, $\mathscr{F}$ is a filter over S.

(b) If $\mathscr{U}$ is an ultrafilter over S, then, since $\mathscr{U}$ is a filter, every intersection of finitely many elements of $\mathscr{U}$ is non-empty, and by definition of ultrafilter, for all $x \subseteq S$ we have either $x \in \mathscr{U}$ or $S \setminus x \in \mathscr{U}$.

For the other direction assume towards a contradiction that for some $x, y \in \mathscr{U}$ we have $x \cap y \notin \mathscr{U}$. Then $z := S \setminus (x \cap y) \in \mathscr{U}$ and $x \cap y \cap z = \emptyset$, which contradicts the fact that finite intersections of elements of $\mathscr{U}$ are non-empty. Furthermore, let $x \in \mathscr{U}$ and let $y \supseteq x$. Then we have $y \in \mathscr{U}$, since otherwise, $S \setminus y \in \mathscr{U}$ and $x \cap (S \setminus y) = \emptyset$. Thus, $\mathscr{U}$ is a filter, and since for all $x \subseteq S$ we have either $x \in \mathscr{U}$ or $S \setminus x \in \mathscr{U}$, $\mathscr{U}$ is an ultrafilter.

15.1 Let $\mathscr{F}$ be a filter over a non-empty set S. We want to extend $\mathscr{F}$ to an ultrafilter over S. By the proof of Theorem 13.3, there exists an ordinal $\alpha \in \Omega$ and a bijection $f : \alpha \to \mathscr{P}(S)$. For the sake of simplicity, for every $\beta \in \alpha$ let $x_\beta := f(\beta)$. First, we construct by transfinite induction for every $\beta \in \alpha + 1$ a filter $\mathscr{F}_\beta$ such that for all $\gamma \in \beta \in \alpha$ we have $\mathscr{F} \subseteq \mathscr{F}_\gamma \subseteq \mathscr{F}_\beta \subseteq \mathscr{F}_\alpha$, and then we show that $\mathscr{F}_\alpha$ is an ultrafilter.

Let $\mathscr{F}_0 := \mathscr{F}$, for non-empty limit ordinals α let

$$\mathscr{F}_\alpha := \bigcup_{\beta \in \alpha} \mathscr{F}_\beta ,$$

and for successor ordinals let

$$\mathscr{F}_{\beta+1} := \begin{cases} \mathscr{F}_\beta \cup \{x_\beta\} & \text{if } x_\beta \cap \bigcap X \neq \emptyset \text{ for every finite set } X \subseteq \mathscr{F}_\beta, \\ \mathscr{F}_\beta & \text{otherwise.} \end{cases}$$

Notice that by Exercise 15.0.(a), for every $\beta \in \alpha + 1$, $\mathscr{F}_\beta$ can be extended to a filter.

Now we show that $\mathscr{F}_\alpha$ is an ultrafilter. By Exercise 15.0.(b), we have to show that every intersection of finitely many elements of $\mathscr{F}_\alpha$ is non-empty and for all $x \subseteq S$ we have either $x \in \mathscr{F}_\alpha$ or $S \setminus x \in \mathscr{F}_\alpha$. The former is an immediate consequence of the construction. For the latter, assume towards a contradiction that there is an $x \subseteq S$ such that neither x nor $S \setminus x$ belongs to $\mathscr{F}_\alpha$. Let $\gamma, \beta \in \alpha$ be such that $x = x_\gamma$ and $x_\beta = S \setminus x$. Without loss of generality, we may assume that $\gamma \in \beta$. Since $x_\gamma, x_\beta \notin \mathscr{F}_\alpha$, there are finite sets $X_\gamma \subseteq \mathscr{F}_\gamma$ and $X_\beta \subseteq \mathscr{F}_\beta$ such that

$$x_\gamma \cap \bigcap X_\gamma = \emptyset \quad \text{and} \quad x_\beta \cap \bigcap X_\beta = \emptyset\,.$$

Now, since $X_\gamma \cup X_\beta$ is a finite subset of $\mathscr{F}_\beta$, by the definition of x_γ and x_β we get that

$$\bigcap (X_\gamma \cup X_\beta) = \emptyset\,,$$

which contradicts the fact that $\mathscr{F}_\beta$ can be extended to a filter.

15.2 Since definition of $c^{\mathbf{M}^*}$ does not depend on a choice of representatives, it is well-defined.

Let $R \in \mathscr{L}$ be an n-ary relation symbol and let f_i and g_i as in the proof for $F^{\mathbf{M}^*}$. We have to show that

$$\left\{\iota \in I : \langle f_0(\iota), \dots, f_{n-1}(\iota)\rangle \in R^{\mathbf{M}_\iota}\right\} \in \mathscr{U} \iff \left\{\iota \in I : \langle g_0(\iota), \dots, g_{n-1}(\iota)\rangle \in R^{\mathbf{M}_\iota}\right\} \in \mathscr{U}\,.$$

For $0 \leq i < n$ let

$$x_i := \left\{\iota \in I : f_i(\iota) = g_i(\iota)\right\},$$

and let $X := x_0 \cap \dots \cap x_{n-1}$. Then, since $x_i \in \mathscr{U}$ for each $0 \leq i < n$ and since $\mathscr{U}$ is a filter, we have $X \in \mathscr{U}$. Thus, we have:

$$\begin{aligned}
&\left\{\iota \in I : \langle f_0(\iota), \dots, f_{n-1}(\iota)\rangle \in R^{\mathbf{M}_\iota}\right\} \in \mathscr{U} \\
\iff &\left\{\iota \in I : \langle f_0(\iota), \dots, f_{n-1}(\iota)\rangle \in R^{\mathbf{M}_\iota}\right\} \cap X \in \mathscr{U} \\
\iff &\left\{\iota \in I : \langle g_0(\iota), \dots, g_{n-1}(\iota)\rangle \in R^{\mathbf{M}_\iota}\right\} \cap X \in \mathscr{U} \\
\iff &\left\{\iota \in I : \langle g_0(\iota), \dots, g_{n-1}(\iota)\rangle \in R^{\mathbf{M}_\iota}\right\} \in \mathscr{U}
\end{aligned}$$

15.3 As in the proof of Łoś's Theorem 15.2 we proceed by induction on the number of symbols $\neg$, $\vee$ and $\forall$ which appear in σ', where the argument for atomic sentences and for $\sigma' \equiv \neg\sigma_0$ remains unchanged.

The case when $\sigma' \equiv \sigma_1 \vee \sigma_2$ is justified by

$$\begin{aligned}
\mathbf{M}^* \vDash \sigma_1 \vee \sigma_2 &\iff \mathbf{M}^* \vDash \sigma_1 \quad \text{OR} \quad \mathbf{M}^* \vDash \sigma_2 \\
&\iff \underbrace{\left\{\iota \in I : \mathbf{M}_\iota \vDash \sigma_1\right\}}_{=:x_1} \in \mathscr{U} \text{ OR } \underbrace{\left\{\iota \in I : \mathbf{M}_\iota \vDash \sigma_2\right\}}_{=:x_2} \in \mathscr{U} \\
&\iff x_1 \cup x_2 \in \mathscr{U} \\
&\iff \left\{\iota \in I : \mathbf{M}_\iota \vDash \sigma_1 \vee \sigma_2\right\} \in \mathscr{U},
\end{aligned}$$

where the penultimate equivalence uses the fact that if $x_1 \notin \mathscr{U}$ and $x_2 \notin \mathscr{U}$, then $I \setminus (x_1 \cup x_2) = (I \setminus x_1) \cap (I \setminus x_2) \in \mathscr{U}$ and hence $x_1 \cup x_2 \notin \mathscr{U}$. Finally, suppose $\sigma' \equiv \forall \nu \sigma_0$ and that for any $[g] \in A^*$ we have

$$\mathbf{M}^* \tfrac{[g]}{\nu} \vDash \sigma_0(\nu) \iff \left\{\iota \in I : \mathbf{M}_\iota \tfrac{g(\iota)}{\nu} \vDash \sigma_0(\nu)\right\} \in \mathscr{U}.$$

Just as for $\exists$ we then find:

$$\begin{aligned}
\mathbf{M}^* \vDash \forall \nu \sigma_0 &\iff \text{FOR ALL } [g] \text{ IN } A^* : \ \mathbf{M}^* \tfrac{[g]}{\nu} \vDash \sigma_0(\nu) \\
&\iff \text{FOR ALL } [g] \text{ IN } A^* : \underbrace{\left\{\iota \in I : \mathbf{M}_\iota \tfrac{g(\iota)}{\nu} \vDash \sigma_0(\nu)\right\}}_{=:x} \in \mathscr{U}
\end{aligned}$$

Now, since $\left\{\iota \in I : \mathbf{M}_\iota \vDash \forall \nu \sigma_0\right\} \subseteq x$, we obtain

$$\left\{\iota \in I : \mathbf{M}_\iota \vDash \forall \nu \sigma_0\right\} \in \mathscr{U} \implies \mathbf{M}^* \vDash \forall \nu \sigma_0 \,.$$

For the converse implication, notice that $\left\{\iota \in I : \mathbf{M}_\iota \vDash \forall \nu \sigma_0\right\} \in \mathscr{U}$ if and only if $I \setminus \left\{\iota \in I : \mathbf{M}_\iota \vDash \forall \nu \sigma_0\right\} \notin \mathscr{U}$, where

$$I \setminus \left\{\iota \in I : \mathbf{M}_\iota \vDash \forall \nu \sigma_0\right\} = \left\{\iota \in I : \mathbf{M}_\iota \vDash \exists \nu \neg \sigma_0\right\}.$$

With the help of the Axiom of Choice, we define a function

$$\begin{array}{rcl} g_0 : I & \to & \bigcup_{\iota \in I} A_\iota \\ \iota & \mapsto & a_\iota \end{array}$$

by stipulating that a_ι is a witness for $\mathbf{M}_\iota \vDash \exists \nu \neg \sigma_0$ if such a witness exists, or in the case when $\mathbf{M}_\iota \vDash \forall \nu \sigma_0$, a_ι is an arbitrary element of A_ι. Then we have

$$\left\{\iota \in I : \mathbf{M}_\iota \vDash \exists \nu \neg \sigma_0\right\} = \left\{\iota \in I : \mathbf{M}_\iota \tfrac{g_0(\iota)}{\nu} \vDash \neg \sigma_0(\nu)\right\}.$$

Let us assume $\mathbf{M}^* \vDash \forall \nu \sigma_0$. Then by the implication

$$\mathbf{M}^* \vDash \forall \nu \sigma_0 \implies \text{FOR ALL } [g] \text{ IN } A^* : \left\{\iota \in I : \mathbf{M}_\iota \tfrac{g(\iota)}{\nu} \vDash \sigma_0(\nu)\right\} \in \mathscr{U}$$

we have

$$\left\{\iota \in I : \mathbf{M}_\iota \tfrac{g_0(\iota)}{\nu} \vDash \neg\sigma_0(\nu)\right\} \notin \mathscr{U} ,$$

and we conclude as follows:

$$\begin{aligned} & \left\{\iota \in I : \mathbf{M}_\iota \tfrac{g_0(\iota)}{\nu} \vDash \neg\sigma_0(\nu)\right\} \notin \mathscr{U} \\ \iff & \left\{\iota \in I : \mathbf{M}_\iota \vDash \exists\nu\neg\sigma_0\right\} \notin \mathscr{U} \\ \iff & I \setminus \left\{\iota \in I : \mathbf{M}_\iota \vDash \forall\nu\sigma_0\right\} \notin \mathscr{U} \\ \iff & \left\{\iota \in I : \mathbf{M}_\iota \vDash \forall\nu\sigma_0\right\} \in \mathscr{U} \end{aligned}$$

Thus, we have

$$\mathbf{M}^* \vDash \forall\nu\sigma_0 \implies \left\{\iota \in I : \mathbf{M}_\iota \vDash \forall\nu\sigma_0\right\} \in \mathscr{U} .$$

15.4 Let $\mathscr{L} = \emptyset$ (i.e., $\mathscr{L}$ is the empty language), and let $I = \mathbb{N}$ and for each $i \in I$ let $\mathbf{M}_i$ be an $\mathscr{L}$-structure with domain $A_\iota = \{0, \ldots, \iota\}$ for $\iota \in \mathbb{N}$. Hence, each A_ι is a finite $\mathscr{L}$-structure. Now let $\mathscr{U}$ be a non-trivial ultrafilter on $\mathbb{N}$. Let $\mathbf{M}^*$ be the corresponding ultraproduct. Note that in all but finitely many structures the sentence

$$\sigma_n \equiv \exists x_1 \ldots \exists x_n \Big(\bigwedge_{i \neq j} x_i \neq x_j \Big)$$

is fulfilled in all but finitely many $\mathbf{M}_i$, hence,

$$\{\iota \in \mathbb{N} : \mathbf{M}_\iota \vDash \sigma_n\} \in \mathscr{U}$$

for each $n \in \mathbb{N}$, which implies that the domain of $\mathbf{M}$ is infinite.

Chapter 16

16.0 Let $\mathscr{U} = \{x \subseteq \omega : n_0 \in x\}$ be a principal ultrafilter, where $n_0 \in \omega$. Then we have:

$$\begin{aligned} f \sim g &\iff \{n \in \omega : f(n) = g(n)\} \in \mathscr{U} \\ &\iff f(n_0) = g(n_0) \end{aligned}$$

The second equivalence can be verified as follows: If $f(n_0) = g(n_0)$ then the set $x = \{n \in \omega : f(n) = g(n)\}$ contains $\{n_0\}$, which implies that $x \in \mathscr{U}$. Define the function $h : \mathbb{N}^*_\omega \to \mathbb{N}_\omega$ by stipulating $h(f) := fn_0$. Then h is an isomorphism between $\mathbb{N}^*_\omega$ and $\mathbb{N}_\omega$, i.e., $\mathbb{N}^*_\omega$ is isomorphic to the model $\mathbb{N}_\omega$.

16.1 Let $\mathscr{U}$ be the ultrafilter with which we have constructed the ultrapower $\mathbb{N}^*_\omega$, and let σ be an $\mathscr{L}_{\mathsf{PA}}$-sentence. We have to show that

$$\mathbb{N}^*_\omega \vDash \sigma \iff \mathbb{N}_\omega \vDash \sigma\,.$$

By Łoś's Theorem we have

$$\mathbb{N}^*_\omega \vDash \sigma \iff \{\iota \in \omega : (\mathbb{N}_\omega)_\iota \vDash \sigma\} \in \mathscr{U}\,,$$

and since $(\mathbb{N}_\omega)_\iota = \mathbb{N}_\omega$ for all $\iota \in \omega$, we have

$$\mathbb{N}^*_\omega \vDash \sigma \iff \mathbb{N}_\omega \vDash \sigma\,.$$

16.2 Let $G := \{[g_k] : k \in \omega\}$ be a countable set of elements of ω^*. We use a diagonal argument to construct a function $f : \omega \to \omega$, such that $[f] \notin G$. The function f is defined as follows:

$$\begin{aligned} f(0) &:= g_0(0) + 1 \\ f(1) &:= \max\{g_0(1),\, g_1(1)\} + 1 \\ &\vdots \qquad\qquad\quad \vdots \\ f(n) &:= \max\{g_0(n),\, g_1(n), \ldots,\, g_n(n)\} + 1 \\ &\vdots \qquad\qquad\quad \vdots \end{aligned}$$

By construction, for all $n \geq k$ we have that $f(n) > g_k(n)$, and since $\mathscr{U}$ is a non-trivial ultrafilter (i.e., $\mathscr{U}$ does not contain finite sets), for all $k \in \omega$ we have

$$\{n \in \omega : f(n) > g_k(n)\} \in \mathscr{U}\,.$$

Hence, for every $k \in \omega$ we have $\mathbb{N}^*_\omega \vDash [f] > [g_k]$, and therefore $[f] \notin G$.

16.3 Let $[g] < [g']$ and let $h : \omega \to \omega$ be such that $[g] + [h] = [g']$. Since $\mathbb{N}^*_\omega \vDash \mathsf{PA}$, such a function h exists and $[h]$ is unique. Furthermore, for each $k \in \omega$ let $c_k : \omega \to \omega$ be such that for all $n \in \omega$, $c_k(n) = k$. We consider the following two cases.

Case 1: $\exists M \in \omega\, \exists x \in \mathscr{U}\, \forall i \in x \big(h(i) \leq M\big)$.

We claim that in this case, there exists a $k \in \omega$ with $0 \leq k \leq M$ such that $[h] = [c_k]$, i.e., there is a $y \in \mathscr{U}$ such that for all $i \in y$, $h(i) = k$. In order to prove the claim, for each $0 \leq l \leq M$ let

$$y_l := \big\{ i \in x : h(i) = l \big\}.$$

Then $x = y_0 \,\dot\cup \cdots \dot\cup\, y_M$ (i.e., x is the disjoint union of the sets $y_0, \ldots y_M$), and since $x \in \mathscr{U}$ and $\mathscr{U}$ is an ultrafilter, either $y_0 \in \mathscr{U}$ or $y_1 \,\dot\cup \cdots \dot\cup\, y_M \in \mathscr{U}$. If $y_0 \in \mathscr{U}$, then for all $i \in y_0$ we have $h(i) = 0$, which implies $[h] = [c_0]$. Otherwise, either $y_1 \in \mathscr{U}$ or $y_2 \,\dot\cup \cdots \dot\cup\, y_M \in \mathscr{U}$, and we can proceed as before. So, we finally find a $y \in \mathscr{U}$ and a k with $0 \leq k \leq M$ such that $[h] = [c_k]$.

Thus, we have $[g] + [c_k] = [g']$, and since $\mathbb{N}^*_\omega \vDash \mathsf{PA}$, this shows that

$$\big\{ [f] \in \omega^* : [g] \leq [f] \leq [g'] \big\} = \big\{ [g] + [c_l] \in \omega^* : 0 \leq l \leq k \big\},$$

which is a finite set.

Case 2: $\forall M \in \omega\, \forall x \in \mathscr{U}\, \exists i \in x \big(h(i) > M\big)$.

For each $r \in \mathbb{R}$ with $0 \leq r \leq 1$ let $h_r : \omega \to \omega$ be defined by stipulating $h_r(n) := \lceil r \cdot h(n) \rceil$, where

$$\lceil r \cdot h(n) \rceil := \min \big\{ k \in \omega : k \geq r \cdot h(n) \big\}.$$

Recall that by the Solution to Exercise 13.8, the set of reals $r \in \mathbb{R}$ with $0 \leq r \leq 1$ is uncountable. Thus, it is enough to show that for all $r, s \in \mathbb{R}$ with $0 \leq r < s \leq 1$ we have $[h_r] < [h_s]$. Since we obviously have $[h_r] \leq [h_s]$, we just have to show that $[h_r] \neq [h_s]$, i.e., there is no $x \in \mathscr{U}$ such that for all $i \in x$, $h_r(i) = h_s(i)$. Since $r < s$ we find an $M_0 \in \omega$ such that

$$r + \frac{1}{M_0} < s\,.$$

Then $r \cdot M_0 + 1 < s \cdot M_0$ and for every $i \in \omega$ with $h(i) > M_0$ we have:

$$\begin{aligned} h_s(i) - h_r(i) &= \lceil s \cdot h(i) \rceil - \lceil r \cdot h(i) \rceil \\ &= \lceil \underbrace{s \cdot M_0}_{> r \cdot M_0 + 1} + \underbrace{s \cdot (h(i) - M_0)}_{> r \cdot (h(i) - M_0)} \rceil - \lceil r \cdot M_0 + r \cdot (h(i) - M_0) \rceil \\ &\geq 1 \end{aligned}$$

In particular, we have $h_s(i) \neq h_r(i)$ for all $i \in \omega$ with $h(i) > M_0$. Now, assume towards a contradiction that there is an $x_0 \in \mathscr{U}$ such that for all $i \in x_0$, $h_r(i) = h_s(i)$. Then by our assumption, for M_0 and x_0 there exists an $i \in x_0$ such that $h(i) > M_0$, which is a contradiction to the choice of x_0.

Thus, for each $r \in \mathbb{R}$ with $0 \leq r \leq 1$ we have $[g] \leq [g] + [h_r] \leq [g']$, which shows that

$$\big\{[f] \in \omega^* : [g] \leq [f] \leq [g']\big\} \supseteq \big\{[g] + [h_r] \in \omega^* : 0 \leq r \leq 1\big\}$$

is an uncountable set.

Chapter 17

17.0 Assume that $[\lambda] < [\mu]$ and take λ', μ' with $\lambda' \sim \lambda$ and $\mu' \sim \mu$. Furthermore, let $M_\lambda, M_\mu \in \mathcal{N}$ be such that for all $n \in \mathbb{Z}_\mathcal{N}$ we have

$$|\lambda(n) - \lambda'(n)| \leq M_\lambda \quad \text{and} \quad |\mu(n) - \mu'(n)| \leq M_\mu \,.$$

Since $\lambda < \mu$, we have $\mathrm{pos}(\mu - \lambda)$, i.e.,

$$\forall N_0 \in \mathcal{N}\, \exists n_0 \in \mathcal{N} \big(\mu(n_0) - \lambda(n_0) > N_0\big) \,.$$

Let $N_1 := N_0 + (M_\lambda + M_\mu)$, and let n_1 be such that

$$\mu(n_1) - \lambda(n_1) > N_1 \,.$$

Then we have

$$\mu'(n_1) - \lambda'(n_1) \;\geq\; \underbrace{\underbrace{\mu(n_0) - \lambda(n_0)}_{>N_1} - (M_\lambda + M_\mu)}_{>N_0}$$

which shows that for every $N_0 \in \mathcal{N}$ we find an $n_1 \in \mathcal{N}$ such that $\mu'(n_1) - \lambda'(n_1) > N_0$, hence, $\lambda' < \mu'$.

17.1 (a) By the definition of addition in $\mathbb{R}^\mathscr{S}_\mathcal{N}$, we have

$$\mathit{\Gamma}([\lambda] + [\mu]) = \mathit{\Gamma}([\lambda + \mu]) = [(a_n^{\lambda+\mu})]$$

where

$$a_n^{\lambda+\mu} = \begin{cases} 0 & \text{if } n = 0 \\ \frac{(\lambda+\mu)(n)}{n} & \text{otherwise.} \end{cases}$$

Now, since

$$\frac{(\lambda+\mu)(n)}{n} = \frac{\lambda(n) + \mu(n)}{n} = \frac{\lambda(n)}{n} + \frac{\mu(n)}{n} = a_n^\lambda + a_n^\mu \,,$$

we obtain $\mathit{\Gamma}([\lambda]+[\mu]) = [(a_n^{\lambda+\mu})] = [(a_n^\lambda)] + [(a_n^\mu)] = \mathit{\Gamma}\big([\lambda]\big) + \mathit{\Gamma}\big([\mu]\big)$.

(b) We have

$$\mathit{\Gamma}(0_{[\mathscr{S}]}) = \mathit{\Gamma}([\lambda_0]) = [(a_n^{\lambda_0})] \,,$$

where

$$a_n^{\lambda_0} = \begin{cases} 0 & \text{if } n = 0, \\ \frac{\lambda_0(n)}{n} = \frac{0}{n} = 0 & \text{otherwise.} \end{cases}$$

Hence, $(a_n^{\lambda_0}) = (0_n)$, which shows that $\mathit{\Gamma}(0_{[\mathscr{S}]}) = 0_{[\mathscr{C}]}$.

(c) By the definition of multiplication in $\mathbb{R}^{\mathscr{F}}_{\mathcal{N}}$, we have

$$\Gamma([\lambda]\cdot[\mu]) = \Gamma([\lambda\circ\mu]) = [(a_n^{\lambda\circ\mu})]\,,$$

where

$$a_n^{\lambda\circ\mu} = \begin{cases} 0 & \text{if } n=0,\\ \frac{(\lambda\circ\mu)(n)}{n} = \frac{\lambda(\mu(n))}{n} & \text{otherwise.}\end{cases}$$

Assume first that $\mu \approx \lambda_0$. Then

$$\Gamma([\lambda]\cdot[\mu]) = \Gamma([\lambda]\cdot[\lambda_0]) = \Gamma([\lambda\circ\lambda_0]) = [(a_n^{\lambda\circ\lambda_0})]\,,$$

where

$$a_n^{\lambda\circ\lambda_0} = \begin{cases} 0 & \text{if } n=0,\\ \frac{\lambda(\lambda_0(n))}{n} = \frac{\lambda(0)}{n} & \text{otherwise.}\end{cases}$$

Let $M_0 := \lambda(0)$. Since for every $\varepsilon \in \mathbb{Q}^+_{\mathcal{N}}$ and for every $n \geq |\frac{M_0}{\varepsilon}|$ we have $|a_n^{\lambda\circ\lambda_0} - 0| = |\frac{M_0}{n}| \leq \varepsilon$, we get that $(a_n^{\lambda\circ\lambda_0}) \approx (0_n)$. Therefore, $\Gamma([\lambda]\cdot[\lambda_0]) = 0_{[\mathscr{C}]}$. On the other hand,

$$\Gamma([\lambda])\cdot\Gamma([\lambda_0]) = [(a_n^{\lambda})]\cdot[(0_n)] = [(a_n^{\lambda}\cdot 0_n)] = [(0_n)] = 0_{[\mathscr{C}]}\,.$$

Hence, $\Gamma([\lambda]\cdot[\lambda_0]) = \Gamma([\lambda])\cdot\Gamma([\lambda_0])$.

Now, assume that $\mu \not\approx \lambda_0$, and without loss of generality assume that $\mathrm{pos}(\mu)$ and that $\mu(n) > 0$ for all $n \in \mathcal{N}$ with $n > 0$. Since $\mathrm{pos}(\mu)$, for every $M \in \mathcal{N}$ we find an $m \in \mathcal{N}$ such that $\mu(m) > N + M_\mu$. Let $M = N + M_\mu$, where $N \in \mathcal{N}$. We claim that for every $n \geq m$ we have $\mu(n) > N$. To see this, assume towards a contradiction that there is some $n_0 \geq m$ such that $\mu(n_0) < N$. Then, by the properties of μ, we get

$$\big|\underbrace{\mu(n_0)}_{<N} - \big(\underbrace{\mu(m)}_{\geq N+M_\mu} + \underbrace{\mu(n_0-m)}_{>0}\big)\big| \;>\; M_\mu\,,$$

which is a contradiction to the definition of M_μ.
Let the sequence $(a_n^{\tilde{\lambda}})$ be defined by stipulating

$$a_n^{\tilde{\lambda}} := \frac{\lambda(\mu(n))}{\mu(n)}\,.$$

Then, since (a_n^{λ}) is a Cauchy sequence, also $(a_n^{\tilde{\lambda}})$ is a Cauchy sequence for which we have $(a_n^{\tilde{\lambda}}) \approx (a_n^{\lambda})$.

Finally, by definition, for $n \neq 0$ we have:

$$\frac{(\lambda \circ \mu)(n)}{n} = \frac{\lambda(\mu(n))}{n} = \frac{\lambda(\mu(n))}{\mu(n)} \cdot \frac{\mu(n)}{n}$$

Thus, we obtain

$$\Gamma([\lambda]\cdot[\mu]) = [(a_n^{\bar{\lambda}})\cdot(a_n^{\mu})] = [(a_n^{\lambda})\cdot(a_n^{\mu})] = [(a_n^{\lambda})]\cdot[(a_n^{\mu})] = \Gamma([\lambda])\cdot\Gamma([\mu])\,.$$

(d) We have

$$\Gamma(1_{[\mathscr{S}]}) = \Gamma([\lambda_1]) = [(a_n^{\lambda_1})]\,,$$

where

$$a_n^{\lambda_1} = \begin{cases} 0 & \text{if } n = 0, \\ \frac{\lambda_1(n)}{n} = \frac{n}{n} = 1 & \text{otherwise.} \end{cases}$$

Hence, $(a_n^{\lambda_1}) \approx (1_n)$, which shows that $\Gamma(1_{[\mathscr{S}]}) = 1_{[\mathscr{C}]}$.

17.2 First we consider the elements of $\mathbb{R}_\omega^{\mathscr{C}}$, $(\mathbb{R}_\omega^{\mathscr{C}})^*$ and $\mathbb{R}_{\omega^*}^{\mathscr{C}}$, respectively.

The elements of set $\mathbb{R}_\omega^{\mathscr{C}}$ are equivalence classes of Cauchy sequences $\langle a_n : n \in \omega\rangle$ in $\mathbb{Q}_\omega$, i.e., the elements of $\mathbb{R}_\omega^{\mathscr{C}}$ are of the form $[(a_n)]$, where (a_n) is a Cauchy sequence in $\mathbb{Q}_\omega$.

By definition, $(\mathbb{R}_\omega^{\mathscr{C}})^*$ is the ultrapower of $\mathbb{R}_\omega^{\mathscr{C}}$ with respect to some non-trivial ultrafilter $\mathscr{U} \subseteq \mathscr{P}(\omega)$. Thus, the elements of $(\mathbb{R}_\omega^{\mathscr{C}})^*$ are equivalence classes of functions $f : \omega \to \mathbb{R}_\omega^{\mathscr{C}}$, where each function f is of the form $\langle [(a_{n,i})_i] : i \in \omega\rangle$ for some Cauchy sequences $\langle a_{n,i} : n \in \omega\rangle$ in $\mathbb{Q}_\omega$. For the sake of simplicity we define

$$(a_{n,i})_i := \big\langle \langle a_{n,i} : n \in \omega\rangle : i \in \omega\big\rangle\,.$$

Now, the elements of $\mathbb{R}_{\omega^*}^{\mathscr{C}}$ are equivalence classes of Cauchy sequences in $\mathbb{Q}_{\omega^*}$, where we can consider $\mathbb{Q}_{\omega^*}$ as the ultrapower of $\mathbb{Q}_\omega$ with respect to the same ultrafilter as above. Thus, the elements of $\mathbb{R}_{\omega^*}^{\mathscr{C}}$ are equivalence classes of Cauchy sequences of the form

$$(a_{n,i})_n := \big\langle \langle a_{n,i} : i \in \omega\rangle : n \in \omega\big\rangle\,,$$

where for all $n, i \in \omega$ we have $a_{n,i} \in \mathbb{Q}_\omega$ and

$$\Big\{ i \in \omega : \langle a_{n,i} : n \in \omega\rangle \text{ is a Cauchy sequence in } \mathbb{Q}_\omega \Big\} \in \mathscr{U}\,.$$

Notice that in the former case, each element of $\mathbb{R}_\omega^{\mathscr{C}}$ is an equivalence class of Cauchy sequences, i.e., for each $i \in \omega$ we have that $(a_{n,i})_i$ is a

Cauchy sequence in $\mathbb{Q}_\omega$, whereas in the latter case it is enough that the set $\{i \in \omega : (a_{n,i})_i$ is a Cauchy sequence in $\mathbb{Q}_\omega\}$ is in $\mathscr{U}$.

Define the function $F : (\mathbb{R}^{\mathscr{C}}_\omega)^* \to \mathbb{R}^{\mathscr{C}}_{\omega^*}$ by stipulating

$$F\big([(a_{n,i})_i]\big) := [(a_{n,i})_n]\,.$$

It remains to show that the function F is well-defined and bijective.

F is well-defined: Recall that by Łoś's Theorem 15.2, $(a_{n,i})_i \sim (b_{n,i})_i$ if and only if there exists an $x \in \mathscr{U}$ such that for each $i_0 \in x$, $(a_{n,i_0})_{i_0} \approx (b_{n,i_0})_{i_0}$. In particular, for each $i_0 \in x$ we have that $(a_{n,i_0})_{i_0}$ and $(b_{n,i_0})_{i_0}$ are Cauchy sequence in $\mathbb{Q}_\omega$, which shows that $[(a_{n,i})_n]$ and $[(b_{n,i})_n]$ are elements of $\mathbb{R}^{\mathscr{C}}_{\omega^*}$ and that $(a_{n,i})_n \approx (b_{n,i})_n$ — the latter follows again by Łoś's Theorem 15.2.

F is injective: Since $(a_{n,i})_i \not\approx (b_{n,i})_i$ if and only if there exists an $x \in \mathscr{U}$ such that for each $i_0 \in x$, $(a_{n,i_0})_{i_0} \not\approx (b_{n,i_0})_{i_0}$, by similar arguments as above we obtain $(a_{n,i})_n \not\approx (b_{n,i})_n$.

F is surjective: Let $[(a_{n,i})_n]$ be an arbitrary element in $\mathbb{R}^{\mathscr{C}}_{\omega^*}$ and let

$$x_0 := \Big\{i \in \omega : \langle a_{n,i} : n \in \omega\rangle \text{ is a Cauchy sequence in } \mathbb{Q}_\omega\Big\} \in \mathscr{U}\,.$$

Now, for each $i \in \omega \setminus x_0$ we choose a Cauchy sequence $\langle b_{n,i} : n \in \omega\rangle$ in $\mathbb{Q}_\omega$ and define the Cauchy sequence $(c_{n,i})_n$ in $\mathbb{Q}_{\omega^*}$ by stipulating

$$c_{n,i} := \begin{cases} a_{n,i} & \text{if } i \in x_0,\\ b_{n,i} & \text{otherwise.}\end{cases}$$

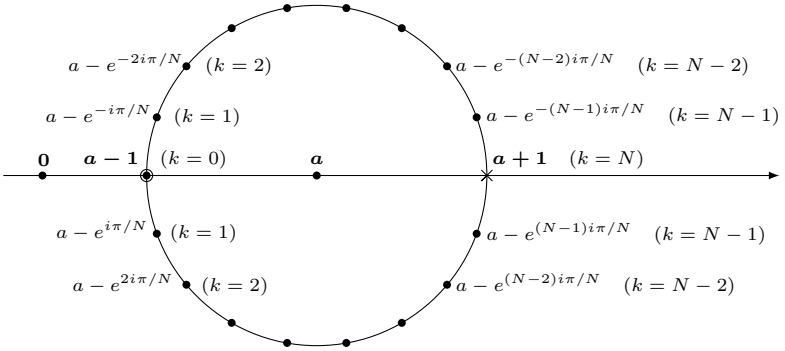
Then, by construction we have $(c_{n,i})_n \approx (a_{n,i})_n$, i.e., $[(c_{n,i})_n] = [(a_{n,i})_n]$, $[(c_{n,i})_i] \in (\mathbb{R}^{\mathscr{C}}_\omega)^*$, and $F\big([(c_{n,i})_i]\big) = [(c_{n,i})_n]$.

17.3 First notice that by Proposition 17.6.(c), for $a^2 \neq 1$ we have

$$\begin{aligned}\int_0^\pi \log\big(1 - 2a\cos(x) + a^2\big)dx \;&=\; \mathrm{st}\Big(\frac{\pi}{N}\sum_{k=0}^{n-1}\log\big(1 - 2a\cos\big(\tfrac{k\pi}{N}\big) + a^2\big)\Big)\\ &=\; \mathrm{st}\Big(\frac{\pi}{N}\log\prod_{k=0}^{N-1}\underbrace{\big(1 - a(e^{ik\pi/N} + e^{-ik\pi/N}) + a^2\big)}_{(a - e^{ik\pi/N})(a - e^{-ik\pi/N})}\Big)\,,\end{aligned}$$

where we used the fact that $\cos(x) = \frac{e^{ix}+e^{-ix}}{2}$.

The following figure shows the complex numbers $a - e^{\pm ki\pi/N}$ in the Argand diagram for $0 \le k \le N$.



Notice that for any k with $0 \le k \le N$ we have $e^{ki\pi/N} = -e^{-(N-k)i\pi/N}$. Thus, taking together products of opposite terms, i.e., products of the form

$$(a-e^{ki\pi/N})(a-e^{-(N-k)i\pi/N}) = (a-e^{ki\pi/N})(a+e^{ki\pi/N}) = (a^2-e^{2ki\pi/N})\,,$$

we obtain

$$\prod_{k=1}^{N-1} \big(a - e^{ki\pi/N}\big)\big(a - e^{-ki\pi/N}\big) = \prod_{k=1}^{N-1} \big(a - e^{ki\pi/N})(a - e^{-(N-k)i\pi/N)}\big) = \prod_{k=1}^{N-1} \big(a^2 - e^{2ki\pi/N}\big)\,.$$

Now, since in the original product $\prod_{k=0}^{N-1}\big(a - e^{ki\pi/N}\big)\big(a - e^{-ki\pi/N}\big)$, we do not have the factor $a+1$ but for $k=0$ we get the factor $a-1$ twice, we finally obtain

$$\prod_{k=0}^{N-1} \big(a - e^{ki\pi/N}\big)\big(a - e^{-ki\pi/N}\big) = \frac{a-1}{a+1} \cdot \prod_{k=0}^{N-1} \big(a^2 - e^{2ki\pi/N}\big)\,.$$

Furthermore, if we replace a^2 by the variable z, then the N pairwise distinct values $\{e^{2ki\pi/N} : 0 \le k < N\}$ are the N roots of the polynomial $z^N - 1$, which shows that

$$\prod_{k=0}^{N-1} (a^2 - e^{2ik\pi/N}) = a^{2N} - 1\,.$$

Thus, we finally have:

$$I = \mathrm{st}\Big(\frac{\pi}{N}\,\log\Big(\frac{a-1}{a+1} \cdot (a^{2N}-1)\Big)\Big)$$

If $|a| < 1$, then, since $a^{2N} - 1 \approx -1$, we have $0 < \frac{a-1}{a+1} \cdot (a^{2N} - 1)$ and obtain:

$$I = \operatorname{st}\Big(\frac{\pi}{N} \underbrace{\log\Big(\frac{a-1}{a+1} \cdot (a^{2N} - 1)\Big)}_{\text{bounded by } \log\left(\frac{2-a}{1+a}\right)}\Big) = 0$$

If $|a| > 1$, then $\operatorname{st}\big(\frac{\pi}{N} \log \frac{a-1}{a+1}\big) = 0$ and we obtain:

$$I = \operatorname{st}\Big(\frac{\pi}{N} \underbrace{\log(a^{2N} - 1)}_{\log a^{2N} + \log(1 - \frac{1}{a^{2N}})}\Big) = \operatorname{st}\Big(\pi \log a^2 + \frac{\pi}{N} \underbrace{\log(1 - \frac{1}{a^{2N}})}_{\text{bounded}}\Big) = \pi \cdot \log a^2$$

Thus, we have

$$\int_0^\pi \log\big(1 - 2a\cos(x) + a^2\big)dx = \begin{cases} 0 & \text{if } |a| < 1, \\ \pi \cdot \log a^2 & \text{if } |a| > 1. \end{cases}$$

References

1. Norbert A'Campo, *A natural construction for the real numbers*, ***Elemente der Mathematik***, vol. 76 (2021), 89–105.
2. Aristotle, ***Topics***, Athens, *published by Andronikos of Rhodos around* 40 b.c.
3. John L. Bell and Alan B. Slomson, ***Models and Ultraproducts: An Introduction***, North-Holland, Amsterdam, 1969.
4. Paul J. Cohen, *The independence of the continuum hypothesis I.*, ***Proceedings of the National Academy of Sciences (U.S.A.)***, vol. 50 (1963), 1143–1148.
5. _____, *The independence of the continuum hypothesis II.*, ***Proceedings of the National Academy of Sciences (U.S.A.)***, vol. 51 (1964), 105–110.
6. Richard Dedekind, ***Was sind und was sollen die Zahlen***, Friedrich Vieweg & Sohn, Braunschweig, 1888 (see also [7, pp. 335–390]).
7. _____, ***Gesammelte mathematische Werke III***, ed. by R. Fricke, E. Noether, and Ö. Ore, Friedrich Vieweg & Sohn, Braunschweig, 1932.
8. Herbert Enderton, ***A mathematical introduction to logic***, Academic Press, New York-London, 1972.
9. Euclid, ***The Thirteen Books of the Elements***, Volume I: Books I & II [translated with introduction and commentary by Sir Thomas L. Heath], Dover, 1956.
10. Adolf Fraenkel, *Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre*, ***Mathematische Annalen***, vol. 86 (1922), 230–237.
11. Gerhard Gentzen, *Untersuchungen über das logische Schließen I*, ***Mathematische Zeitschrift***, vol. 39 (1935), 176–210.
12. _____, *Untersuchungen über das logische Schließen II*, ***Mathematische Zeitschrift***, vol. 39 (1935), 405–431.
13. _____, *Die Widerspruchsfreiheit der reinen Zahlentheorie*, ***Mathematische Annalen***, vol. 112 (1936), 493–565.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3

14. Kurt Gödel, *Über die Vollständigkeit des Logikkalküls*, Dissertation (1929), University of Vienna (Austria), (reprinted and translated into English in [18].
15. ———, *Die Vollständigkeit der Axiome des logischen Funktionenkalküls*, ***Monatshefte für Mathematik und Physik***, vol. 37 (1930), 349–360 (see [56, 18] for a translation into English).
16. ———, *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme*, ***Monatshefte für Mathematik und Physik***, vol. 38 (1931), 173–198 (see [56, 18] for a translation into English).
17. ———, *The consistency of the axiom of choice and of the generalized continuum-hypothesis*, ***Proceedings of the National Academy of Sciences (U.S.A.)***, vol. 24 (1938), 556–557 (reprinted in [19]).
18. ———, ***Collected Works, Volume I: Publications 1929–1936***, edited by S. Feferman (Editor-in-chief), J. W. Dawson, Jr., S. C. Kleene, G. H. Moore, R. M. Solovay, J. van Heijenoort, Oxford University Press, New York, 1986.
19. ———, ***Collected Works, Volume II: Publications 1938–1974***, edited by S. Feferman (Editor-in-chief), J. W. Dawson, Jr., S. C. Kleene, G. H. Moore, R. M. Solovay, J. van Heijenoort, Oxford University Press, New York, 1990.
20. Hermann Grassmann, ***Lehrbuch der Arithmetik für höhere Lehranstalten***, Th. Chr. Fr. Enslin, Berlin, 1861.
21. Lorenz Halbeisen, *A framework for metamathematics*, Axiomatic thinking II, Springer, Cham, 2022, pp. 3–8.
22. Lorenz J. Halbeisen, ***Combinatorial Set Theory, with a gentle introduction to forcing***, 2nd ed., Springer Monographs in Mathematics, Springer, London, 2017.
23. Leon Henkin, *The completeness of the first-order functional calculus*, ***The Journal of Symbolic Logic***, vol. 14 (1949), 159–166.
24. ———, *A problem concerning provability*, ***Journal of Symbolic Logic***, vol. 17 (1952), 160.
25. ———, *The discovery of my completeness proofs*, ***The Bulletin of Symbolic Logic***, vol. 2 (1996), 127–158.
26. David Hilbert, *Mathematische Probleme*, ***Vortrag, gehalten auf dem internationalen Mathematiker-Kongreß zu Paris 1900*** (1900), 253–297.
27. ———, *Die Grundlagen der Mathematik. Vortrag, gehalten auf Einladung des Mathematischen Seminars im Juli 1927 in Hamburg.*, ***Abhandlungen aus dem mathematischen Seminar der Hamburgischen Universität***, vol. 6 (1928), 65–85 (see [56] for a translation into English).
28. David Hilbert and Paul Bernays, ***Grundlagen der Mathematik***, Vol. II, Springer, Berlin, 1939.

29. Richard Kaye, ***Models of Peano Arithmetic***, [Oxford Logic Guides 15], The Clarendon Press Oxford University Press, New York, 1991.
30. Peter Koepke, *Models of Set Theory I*, Lecture Notes, University of Bonn (Germany).
31. Kenneth Kunen, ***Set Theory, an Introduction to Independence Proofs***, [Studies in Logic and the Foundations of Mathematics 102], North-Holland, Amsterdam, 1983.
32. Martin Hugo Löb, *Solution of a problem of Leon Henkin*, ***The Journal of Symbolic Logic***, vol. 20 (1955), 115–118.
33. Jerzy Łoś, *Quelques remarques, théorèmes et problèmes sur les classes définissables d'algèbres*, Mathematical interpretation of formal systems, North-Holland Publishing Co., Amsterdam, 1955, pp. 98–113.
34. Leopold Löwenheim, *Über Möglichkeiten im Relativkalkül*, ***Mathematische Annalen***, vol. 76 (1915), 447–470.
35. Elliott Mendelson, ***Introduction to Mathematical Logic***, 6th ed., [Discrete Mathematics and its Applications], CRC Press, Boca Raton, FL, 2015.
36. Leila Mizrahi, *Thoroughly formalizing an uncommon construction of the real numbers*, Master Thesis (2015), University of Zürich (Switzerland).
37. Roman Murawski, *Undefinability of truth. The problem of priority: Tarski vs. Gödel*, ***History and Philosophy of Logic***, vol. 9 (1998), 153–160.
38. John von Neumann, *Eine Axiomatisierung der Mengenlehre*, ***Journal für die Reine und Angewandte Mathematik***, vol. 154 (1925), 219–240 (see [56] for a translation into English).
39. _____, *Die Axiomatisierung der Mengenlehre*, ***Mathematische Zeitschrift***, vol. 27 (1928), 669–752.
40. _____, *Über eine Widerspruchfreiheitsfrage in der axiomatischen Mengenlehre*, ***Journal für die Reine und Angewandte Mathematik***, vol. 160 (1929), 227–241.
41. Lawrence Paulson, *A machine-assisted proof of gödel's incompleteness theorems for the theory of hereditarily finite sets*, ***Review of Symbolic Logic***, vol. 7 (2014), 484–498.
42. Giuseppe Peano, ***Arithmetices principia, nova methoda exposita***, Fratres Bocca, Torino, 1889 (see [56] for a translation into English).
43. Mojżesz Presburger, *Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt*, Comptes Rendus I Congès des Mathématiciens des Pays Slaves 92–101, Zusatz ebenda, 395 (1930).
44. Alain M. Robert, ***Nonstandard Analysis***, Dover Publications, Mineola, New York, 2003.

45. Abraham Robinson, ***Non-standard analysis***, North-Holland Publishing Co., Amsterdam, 1966.
46. Raphael M. Robinson, *An essentially undecidable axiom system*, ***Proceedings of the International Congress of Mathematics*** (1950), 729–730.
47. Barkley Rosser, *Extensions of some theorems of gödel and church*, ***The Journal of Symbolic Logic***, vol. 1 (1936), no. 03, 87–91.
48. Joseph R. Shoenfield, ***Mathematical Logic***, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1967.
49. Thoralf Skolem, *Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theorem über dichte Mengen*, Krist. Vid. Selsk. Skr. I, 1920, Nr. 4, 36 S. (1922).
50. _____, *Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre*, Matematikerkongressen i Helsingfors den 4–7 Juli 1922, Den femte skandinaviska matematikerkongressen (Helsingfors), Akademiska Bokhandeln, 1923, pp. 217–232 (see [56] for a translation into English).
51. _____, *Über einige Satzfunktionen in der Arithmetik*, ***Skrifter Vitenskapsakademiet i Oslo***, vol. 7 (1931), 1–28.
52. _____, *Über die Unmöglichkeit einer vollständigen Charakterisierung der Zahlenreihe mittels eines endlichen Axiomensystems*, ***Fundamenta Mathematicae***, vol. 23 (1934), 150–161.
53. Raymond M. Smullyan, ***A beginner's guide to mathematical logic***, Dover Publications, Inc., Mineola, NY, 2014.
54. Stanisław S. Świerczkowski, *Finite sets and gödel's incompleteness theorems*, ***Dissertationes Mathematicae***, vol. 422 (2003), 1–58.
55. Alfred Tarski, *Der Wahrheitsbegriff in den formalisierten Sprachen*, ***Studia Philosophica***, vol. 1 (1936), 261–405.
56. Jean van Heijenoort, ***From Frege to Gödel. A Source Book in Mathematical Logic, 1879–1931***, [Source Books in the History of Science], Harvard University Press, Cambridge, Massachusetts, 1967.
57. William Weiss, **An Introduction to Set Theory**, CreateSpace Independent Publishing Platform, 2014.
58. Alfred North Whitehead and Bertrand Russell, ***Principia Mathematica, Vol. I–III***, Cambridge University Press, Cambridge, 1910–1913.
59. Ernst Zermelo, *Beweis, dass jede Menge wohlgeordnet werden kann*, ***Mathematische Annalen***, vol. 59 (1904), 514–516 (see [56, 63] for a translation into English).
60. _____, *Neuer Beweis für die Möglichkeit einer Wohlordnung*, ***Mathematische Annalen***, vol. 65 (1908), 107–128 (see [56, 63] for a translation into English).
61. _____, *Untersuchungen über die Grundlagen der Mengenlehre. I.*, ***Mathematische Annalen***, vol. 65 (1908), 261–281 (see [56, 63] for a translation into English).

62. _____, *Über Grenzzahlen und Mengenbereiche. Neue Untersuchungen über die Grundlagen der Mengelehre*, ***Fundamenta Mathematicae***, vol. 16 (1930), 29–47 (see [63] for a translation into English).
63. _____, ***Collected Works / Gesammelte Werke***, *Volume I: Set Theory, Miscellanea / Band I: Mengenlehre, Varia*, [Schriften der Mathematisch-naturwissenschaftlichen Klasse der Heidelberger Akademie der Wissenschaften, Nr. 21 (2010)], edited by Heinz-Dieter Ebbinghaus, Craig G. Fraser, and Akihiro Kanamori, Springer-Verlag, Berlin · Heidelberg, 2010.

Symbols

Logic

$\exists$ (exists), 9
$\forall$ (for all), 9
$\neg$ (not), 9
$\rightarrow$ (implies), 9
$\vee$ (or), 9
$\wedge$ (and), 9
$\equiv$, 13
$\mathrm{free}(\varphi)$, 12
$\varphi(\nu/\tau)$, 12
$\varphi(\tau)$, 13
$\varphi \Leftrightarrow \psi$, 18
$(\forall)$, 16
(MP), 16
(DT), 21
$\boxbslash$, 29
$\boldsymbol{\Phi} \vdash\!\!\!\dashv \psi$, 25
$\boldsymbol{\Phi} \nvdash \psi$, 16
$\boldsymbol{\Phi} \vdash \psi$, 16
CNF, 38
DNF, 30
NNF, 30
PNF, 32
sPNF, 32
$\mathrm{Con}(\boldsymbol{\Phi})$, 36
$\neg\,\mathrm{Con}(\boldsymbol{\Phi})$, 36
$\equiv_{\boldsymbol{e}}$, 45
$\mathbf{Th}(\mathsf{T})$, 49
$\mathbf{Th}(\mathbf{M})$, 56
$\overline{\varphi}$, 44
$\mathbf{I}\frac{a}{\nu}$, 42
$\mathbf{I} \vDash \varphi$, 42–43
$\mathbf{M} \vDash \varphi$, 43
$\mathbf{M} \nvDash \varphi$, 43
$j\frac{a}{\nu}$, 42

Peano Arithmetic

$\beta(c, i)$, 109
$\#\zeta$, 115
$\ulcorner \zeta \urcorner$, 116
$\mathrm{con}_{\mathsf{PA}}$, 139
$\mathrm{fml}(f)$, 117
$\mathrm{gn}(n)$, 127
$\mathrm{lh}(c)$, 112
$\lceil \zeta \rceil_V^{\mathrm{gn}}$, 143
$\lceil \zeta \rceil$, 142
$\mathrm{prv}(f)$, 121
$\mathrm{sb\_fml}(v, t_0, f, f')$, 119
$\mathrm{sb\_term}(v, t_0, t, t')$, 119
$\mathrm{seq}(s)$, 112
$\mathrm{nat}(n, x)$, 126
$\mathrm{term}(t)$, 117
$\underline{n}$, 103
$\mathrm{var}(v)$, 117
c_i, 112

Axioms

DLO, 50
GT, 15
PA, 15, 83

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3

Persons

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3

Subjects

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

L. Halbeisen, R. Krapf, *Gödel's Theorems and Zermelo's Axioms*,
https://doi.org/10.1007/978-3-031-85106-3