

Chapter 1

First-Order Logic

In this chapter we introduce a calculus of logical deduction, called first-order logic, that makes it possible to formalize mathematical proofs. The main theorem about this calculus that we shall prove is Gödel's completeness theorem (1.5.2), which asserts that the unprovability of a sentence must be due to the existence of a counterexample. From the finitary character of a formalized proof we then immediately obtain the Finiteness Theorem (1.5.6), which is fundamental for model theory, and which asserts that an axiom system possesses a model provided that every finite subsystem of it possesses a model.

In (1.6) we shall axiomatize a series of mathematical (in particular, algebraic) theories. In order to show the extent of first-order logic, we shall also give within this framework the Zermelo–Fraenkel axiom system for set theory, a theory that allows us to represent all of ordinary mathematics in it.

1.1 Analysis of Mathematical Proofs

In this section we try, by means of an example, to come closer to an answer to the question, “What is a mathematical proof?”. For the example to be considered, we assume that we find ourselves in an undergraduate mathematics course in which the field of all real numbers is being introduced axiomatically. Let us further assume that the field properties have already been covered and the order properties are just now being introduced by the following axioms:

- (0) \leq is a partial order;
- (1) for all x, y , either $x \leq y$ or $y \leq x$;
- (2) for all x, y with $x \leq y$, we have $x + z \leq y + z$ for all z ; and
- (3) if $0 \leq x$ and $0 \leq y$, then also $0 \leq x \cdot y$.

Then we want to give a proof for the following

Claim: $0 \leq x \cdot x$ for all x .

A proof of this could look something like the following:

Proof: 1. From (1) we obtain $0 \leq x$ or $x \leq 0$.
 2. If $0 \leq x$, then (3) gives $0 \leq x \cdot x$.
 3. If, however, $x \leq 0$, then from (2) follows $0 \leq -x$ (where we set $z = -x$).
 4. Now (3) again gives $0 \leq (-x) \cdot (-x) = x \cdot x$.
 5. Therefore $0 \leq x \cdot x$ holds for all x . □

In view of this example of a proof, several remarks are now in order with regard to an exact definition of the concept of “mathematical proof”.

Remark 1.1.1 (on the level of detail in a proof). The level of detail of a proof is as a rule geared toward the background of those for whom the proof is intended. In our example, this was the background of undergraduate mathematics students. For experts, a proof at this level of detail would not be necessary – usually a proof in such a case would consist of the single word “trivial”. For nonmathematicians, on the other hand, the above proof might be too short, hence hard to understand. A nonmathematician might not be able to follow it, since certain intermediate steps that are clear to the mathematician are simply omitted, or certain conventions are used that only mathematicians are familiar with. For example, the mathematician writes $0 \leq (-x) \cdot (-x) = x \cdot x$, and actually means the expression:

$$0 \leq (-x) \cdot (-x) \text{ and } (-x) \cdot (-x) = x \cdot x \text{ imply } 0 \leq x \cdot x.$$

It should be clear that for an exact definition of proof, we must strive for the greatest possible fullness of detail, so that the question of whether a given sequence of sentences is a proof is checkable by anyone who knows this definition. Moreover, it should even be possible for a suitably programmed computer to make this determination.

Remark 1.1.2 (on the choice of a formal language). The *language* used to write a proof out is, as a rule, likewise chosen according to the intended audience. In mathematics it is common to care less about good linguistic style, and much more about unique readability. The example of a proof above can well be taken as typical. From the standpoint of unique readability, however, let us attempt some improvements. Thus, the words “also” (in Axiom 3) or “however” (in line 3 of the proof) can be viewed as purely ornamental. They possess no additional informational content. On the contrary, such ornamental words often cause ambiguities. In the above proof one could also complain that sometimes a generalization “for all x ” is missing from the beginning of a sentence, and sometimes it appears at the end of such a sentence. This especially can easily lead to ambiguities. In order to be able to give an exact definition for the concept of proof, it is therefore indispensable to agree once and for all upon linguistic conventions that guarantee unique readability.

Remark 1.1.3 (on the layout of a proof). Normally a proof consists of a finite sequence of statements. Often additional hints are given, as, for example, in line 4 of

the above proof. An exact definition of proof should, however, make such things superfluous. Such hints should serve only to promote readability, and should have no influence on whether a given sequence of sentences is a proof or not. As to the utilization of the given space for writing out the proof, it should also be immaterial whether the sequence of these sentences is arranged in a series within one line, or (as in the example above) there is only one sentence per line. For the sake of readability, we shall stick to the latter form.

Considering the criticisms in Remark 1.1.2 above, and using symbolism that is widespread in mathematics (which we shall make precise in the next section), we shall now repeat the above proof. First, however, we want to “formalize” those axioms that occur in the proof:

- (1) $\forall xy (x \leq y \vee y \leq x)$
- (2) $\forall xyz (x \leq y \rightarrow x + z \leq y + z)$
- (3) $\forall xy (0 \leq x \wedge 0 \leq y \rightarrow 0 \leq x \cdot y)$

Now to the claim and the proof:

Claim: $\forall x \ 0 \leq x \cdot x$

- Proof:*
- 1. (1) $\rightarrow (0 \leq x \vee x \leq 0)$
 - 2. $0 \leq x \wedge (3) \rightarrow 0 \leq x \cdot x$
 - 3. $x \leq 0 \wedge (2) \rightarrow 0 \leq -x$
 - 4. $0 \leq -x \wedge (3) \rightarrow 0 \leq (-x) \cdot (-x) = x \cdot x$
 - 5. $\forall x \ 0 \leq x \cdot x$ □

In order to take the criticism in Remark 1.1.1 into account somewhat, we could formulate the proof in more detail – say, as follows:

- 1. (1) $\rightarrow (0 \leq x \vee x \leq 0)$
- 2. $0 \leq x \wedge (3) \rightarrow 0 \leq x \cdot x$
- 3. $0 \leq x \rightarrow 0 \leq x \cdot x$
- 4. $x \leq 0 \wedge (2) \rightarrow x + (-x) \leq 0 + (-x)$
- 5. $x + (-x) \leq 0 + (-x) \wedge x + (-x) = 0 \rightarrow 0 \leq 0 + (-x)$
- 6. $0 \leq 0 + (-x) \wedge 0 + (-x) = -x \rightarrow 0 \leq -x$
- 7. $0 \leq -x \wedge (3) \rightarrow 0 \leq (-x) \cdot (-x)$
- 8. $0 \leq (-x) \cdot (-x) \wedge (-x) \cdot (-x) = x \cdot x \rightarrow 0 \leq x \cdot x$
- 9. $x \leq 0 \rightarrow 0 \leq x \cdot x$
- 10. $(0 \leq x \vee x \leq 0) \rightarrow 0 \leq x \cdot x$
- 11. $\forall x \ 0 \leq x \cdot x$ □

Now we can discuss several typical characteristics of our by now already somewhat formalized proof.

A proof is a sequence of expressions, each of which either contains a universally valid, logical fact, or follows purely logically (with the help of the axioms) from

earlier sentences in the proof. Thus, the first line contains a universally valid fact. Namely, it has (if we suppress the variable x for a moment) the form

$$\forall y \varphi(y) \rightarrow \varphi(0),$$

where $\varphi(y)$ is an expression that, in our case, speaks about arbitrary elements y of the real number field. Likewise, lines 2, 4, 5, 6, 7 and 8 represent universally valid implications. Using the axioms (1)–(3) and also the identities $x + (-x) = 0$, $0 + (-x) = -x$ and $(-x) \cdot (-x) = x \cdot x$ (which are also to be used as axioms), lines 3 and 9 result from previous statements by purely logical deductions. Thus, for example, we obtain line 3 from the rule of inference that says: if we have already proved $(\alpha \wedge \beta) \rightarrow \gamma$ and, in addition, β , then we have also thereby proved $\alpha \rightarrow \gamma$. In our case, β is axiom (3), which, as an axiom, needs no proof, or, in other words, can be assumed to have been proved. We obtain line 10 by applying to lines 3 and 9 the rule of inference: from $\alpha \rightarrow \beta$ and $\gamma \rightarrow \beta$ follows $(\alpha \vee \gamma) \rightarrow \beta$. From lines 1 and 10 we actually obtain, at first, only $0 \leq x \cdot x$. However, since this has been proved for a “fixed but arbitrary” x , we deduce $\forall x \ 0 \leq x \cdot x$. Thus, line 11 is likewise a universally valid logical inference.

In the next two sections we want, first, to fix the linguistic framework exactly, and, second, to give an exact definition of proof. Since we shall later have very much to say about formulae and proofs, and must often use induction to prove metatheorems about them, it behooves us to proceed very economically in our definitions of formula and proof. Therefore, we shall not take a large number of rules of inference as a basis, but rather try to get by with a minimum. This has a consequence that gap-free (formal) proofs become very long. Thus the above proof, for example, would, in gap-free form, swell to about 50 lines. However, once we give an exact definition of proof, we shall agree to relax that definition so as to allow the use, in proofs, of so-called “derived” rules of inference. These methods correspond exactly with mathematical practice: in new proofs one refers back, possibly, to already known proofs, without having to repeat them. All that is important is that all gaps could, if necessary, be filled (at least theoretically!).

1.2 Construction of Formal Languages

For the definition of proof, it is necessary to describe the underlying formal language more precisely; this is the goal of the present section. The objects of our consideration will be an alphabet, and the words and statements formed therefrom. The formal language itself therefore becomes an object of our investigation. On the other hand, we use informal (mathematical) colloquial language (which we might describe as “mathematical English”) to formulate everything that we establish in this investigation of the formal language. This is necessary in order for us to communicate these stipulations and results to the reader. Therefore we must deal with two languages, one being the object of our considerations (which we therefore call

the *object language*, or, in other contexts, the *formal language*), and the other being the language in which we talk *about* the object language (we call this second one the *metalanguage*).

The metalanguage will always be the mathematical colloquial language, in which, for example, we occasionally use common abbreviations (such as “iff” for “if and only if”). In the metalanguage, we shall also use the set theoretical conceptual apparatus, as is usual in mathematics. And especially in the second part of this book (the model theoretic part, Chapters 2 and 3), we shall reason in set theory. When the considerations make it necessary, it is, however, also possible to return to the “finitist standpoint”, in which one speaks only of finite sequences (or “strings”) of symbols (built up from the alphabet of the object language) or of finite sequences of such strings of symbols.

The object language will depend on the subject being considered at the time. For example, if we want to talk about the consistency of mathematics, then we adopt the finitist standpoint and therefore require that the alphabet of the object language considered be finite. If, however, we adopt the model theoretic standpoint, then the alphabet may be an arbitrary set.

Before we come to the definitions, we give yet another hint, this time about a fundamental difficulty. The issues that we pursue here are not common in mathematics. Ordinarily one utilizes only one language: the language in which one communicates something, such as a proof. For a mathematician, writing out a sentence is usually tantamount to claiming that that sentence is true. Thinking about the real numbers, for example, one might write (using the usual abbreviations)

$$\forall x \exists y \ x < y,$$

rather than

the statement $\forall x \exists y \ x < y$ holds.

But if we want to speak *about* a language, then we must necessarily distinguish between the symbol-sequence $\forall x \exists y \ x < y$ and its possible meaning. This and the next section will deal only with *syntactical* questions, i.e. questions such as whether a string of symbols is correctly formed with reference to certain rules of formation.

The *alphabet* of the object language that we consider consists of the following fundamental symbols:

logical symbols:	\neg (not) \wedge (and) \forall (for all) \doteq (equals)	
variables:	$v_0 \ v_1 \ v_2 \ \dots \ v_n \ \dots$ ($n \in \mathbb{N} := \{0, 1, 2, \dots\}$)	
relation symbols:	R_i (for $i \in I$)	(1.2.0.1)
function symbols:	f_j (for $j \in J$)	(1.2.0.2)
constant symbols:	c_k (for $k \in K$)	(1.2.0.3)
punctuation:	$, \quad) \quad ($	

Here I , J and K are arbitrary index sets, which may even be empty. If we wish to adopt the finitist standpoint, we can generate the infinitely many variables v_n ($n \in \mathbb{N}$) by means of finitely many basic symbols, say, v and $'$: then instead of the symbol v_n , we would write

$$\underbrace{v''' \dots'}_{n \text{ times}}.$$

One could rewrite the relation, function and constant symbols in an analogous way, in which case the index sets I , J , K would naturally be at most countable.

From these basic symbols we now want to construct certain strings of symbols, which we call terms. Terms will, via a semantical interpretation given later, designate *things*; they are, therefore, possible names. If one keeps this in mind, the following definition of *terms* becomes understandable:

- (a) All variables v_n and all constant symbols c_k are terms.
- (b) If $t_1, \dots, t_{\mu(j)}$ are terms, then so is $f_j(t_1, \dots, t_{\mu(j)})$.
- (c) No other strings of symbols are terms.

Here μ is a function that, to each $j \in J$, assigns the “arity” (= number of arguments) $\mu(j)$ of the function symbol f_j ; thus, $\mu(j) \geq 1$.

Then Tm, the *set of all terms*, is the smallest set of strings of symbols that contains all v_n and c_k and that, for each $j \in J$, contains $f_j(t_1, \dots, t_{\mu(j)})$ whenever it contains $t_1, \dots, t_{\mu(j)}$.

By convention, we may sometimes write $t_1 f_j t_2$ instead of the official $f_j(t_1, t_2)$, in case $\mu(j) = 2$; for example, $t_1 + t_2$ instead of $+(t_1, t_2)$, if $+$ is a binary function symbol.

Next, we construct *formulae*:

- (a) If t_1 and t_2 are terms, then $t_1 \doteq t_2$ is a formula.
- (b) If $t_1, \dots, t_{\lambda(i)}$ are terms, then $R_i(t_1, \dots, t_{\lambda(i)})$ is a formula.
- (c) If φ and ψ are formulae and v is a variable, then $\neg\varphi$ and $(\varphi \wedge \psi)$ and $\forall v \varphi$ are formulae.
- (d) No other strings of symbols are formulae.

Here λ is a function that, to each $i \in I$, assigns the “arity” $\lambda(i)$ of the relation symbol R_i ; again, $\lambda(i) \geq 1$.

Then Fml, the *set of all formulae*, is the smallest set of strings of symbols that contains all strings of the form $t_1 \doteq t_2$ and $R_i(t_1, \dots, t_{\lambda(i)})$ (these are also called the *atomic formulae*), and that contains $\neg\varphi$ and $(\varphi \wedge \psi)$ as well as $\forall v \varphi$ whenever it contains φ and ψ .

From now on, the notations

t_1, t_2, \dots	will denote <i>terms</i> ,
$\varphi, \psi, \rho, \tau, \alpha, \beta, \gamma$ (possibly with subscripts)	will denote <i>formulae</i> , and
u, v, w, x, y, z (possibly with subscripts)	will denote <i>variables</i> .

We further employ the following *abbreviations*:

$(\varphi \vee \psi)$	stands for	$\neg(\neg\varphi \wedge \neg\psi)$	(or)	
$(\varphi \rightarrow \psi)$	stands for	$\neg(\varphi \wedge \neg\psi)$	(implies)	(1.2.0.4)
$(\varphi \leftrightarrow \psi)$	stands for	$(\neg(\varphi \wedge \neg\psi) \wedge \neg(\psi \wedge \neg\varphi))$	(equivalent)	
$\exists v \varphi$	stands for	$\neg\forall v \neg\varphi$	(there exists)	

And we adopt the following *conventions*, which are customary:

1. \vee and \wedge bind more strongly than \rightarrow and \leftrightarrow ;
2. \neg binds more strongly than \vee and \wedge ;
3. $t_1 \neq t_2$ stands for $\neg t_1 \doteq t_2$; (1.2.0.5)
4. $t_1 R_i t_2$ often stands for $R_i(t_1, t_2)$, in case $\lambda(i) = 2$;
5. $\forall u, v, w, \dots$ stands for $\forall u \forall v \forall w \dots$;
6. $\exists x, y, \dots$ stands for $\exists x \exists y \dots$;
7. $(\varphi_1 \wedge \varphi_2 \wedge \varphi_3)$ stands for $((\varphi_1 \wedge \varphi_2) \wedge \varphi_3)$ (i.e. we group left parentheses together); (1.2.0.6)
8. $(\psi_1 \vee \psi_2 \vee \psi_3 \vee \psi_4)$ stands for $((\psi_1 \vee \psi_2) \vee \psi_3) \vee \psi_4$; and (1.2.0.7)
9. we drop outside parentheses when this can lead to no ambiguity.

Thus, according to these conventions, the string of symbols

$$\forall x, y (\neg\varphi \wedge \psi \rightarrow \alpha \vee \beta \vee \gamma)$$

stands for

$$\forall x \forall y ((\neg\varphi \wedge \psi) \rightarrow ((\alpha \vee \beta) \vee \gamma)).$$

Now we wish to enter into the role of variables in formulae. As an example, we consider a formal language (object language) with a relation symbol and a constant symbol. Thus $I = \{0\}$, $J = \emptyset$ and $K = \{0\}$, say. For R_0 we write $<$, and for c_0 we write 0, for short. Let φ denote the formula

$$\exists v_0 (0 < v_0 \wedge v_0 < v_1) \wedge \forall v_0 (v_0 < 0 \rightarrow v_0 < v_1).$$

If we think of the usual ordering on the real numbers, then we see that the variables v_0 and v_1 play different roles in φ . First, it makes little sense to ask whether φ is true in the real numbers. This would begin to make sense only if for v_1 we think of a particular real number. Obviously φ is true if we think of a positive real number for v_1 . Consider the case where we think of v_1 as 1; then φ remains true if we replace v_0 by, say, v_{13} . The “truth value” of φ does not change if we replace the v_0 s in the first part of the formula by v_{13} and the v_0 s in the second part by v_{17} . This is so not only in the case where v_1 is 1, but also in every case. On the other hand, we may not replace the two occurrences of v_1 in φ by two distinct variables; this would alter

the “sense” of ϕ in an essential way. This distinction involving the occurrence of a variable in a formula is captured formally by the following definitions.

In the recursive construction of a “for all” (or “universal”) formula $\forall v \phi$, we refer to the subformula ϕ as the *scope* (or *effective range*) of the “quantifier” $\forall v$. We call an occurrence of a variable v in a formula ψ *bound* if this occurrence lies within the scope of a quantifier $\forall v$ used in the construction of ψ . Every other occurrence¹ of the variable v in the formula ψ is called *free*. We denote by $\text{Fr}(\psi)$ the set of variables that possess at least one free occurrence in ψ . The following equations are easily checked:

$$\begin{aligned}\text{Fr}(\psi) &= \{v \mid v \text{ occurs in } \psi\}, \text{ if } \psi \text{ is atomic;} \\ \text{Fr}(\neg\psi) &= \text{Fr}(\psi); \\ \text{Fr}(\phi \wedge \psi) &= \text{Fr}(\phi) \cup \text{Fr}(\psi); \text{ and} \\ \text{Fr}(\forall u \phi) &= \text{Fr}(\phi) \setminus \{u\}.\end{aligned}$$

The elements of $\text{Fr}(\phi)$ are called the *free variables* of ϕ .

For example, in the formula

$$\forall v_0 (v_0 < 0 \rightarrow v_0 < v_1) \wedge \exists v_2 (0 < v_2 \wedge v_2 < v_0),$$

the variable v_2 has only bound occurrences, the variable v_1 has only free occurrences and the variable v_0 occurs both bound (in the first half) and free (in the second half). Note that the scope of $\forall v_0$ is only $(v_0 < 0 \rightarrow v_0 < v_1)$, and not everything after that $\forall v_0$ symbol.

Later we shall need yet another syntactic operation: the *replacement of a variable* v in a string of symbols ζ by a term t . Let

$$\zeta(v/t) \tag{1.2.0.8}$$

denote the string obtained by replacing each free occurrence of v in ζ by t . If a free occurrence of v in the formula ϕ falls within the scope of a quantifier $\forall u$, and if u occurs somewhere in t , then after replacement of v in ϕ by t , the variable u will obviously fall within the scope of $\forall u$. If this does not happen for any variable u in t , then t is called *free for* v in ϕ . In other words, t is free for v in ϕ if no free occurrence of v in ϕ lies within the scope of a quantifier $\forall u$ used in the construction of ϕ , where u occurs in t .

By analogy with the replacement of a variable, we define the *replacement of a constant* c_k in ζ by a variable v to mean that every occurrence of c_k in ζ is replaced by v . We denote the result of this replacement by

$$\zeta(c_k/v). \tag{1.2.0.9}$$

This, too, is a syntactic operation, i.e. a manipulation of strings of symbols.

¹ We do not count the “occurrence” of v in $\forall v$ as a true occurrence.

If the formula φ possesses no free variables (i.e. if $\text{Fr}(\varphi) = \emptyset$), then we call φ a *sentence*. We write Sent for the set of sentences:

$$\text{Sent} = \{ \varphi \in \text{Fml} \mid \text{Fr}(\varphi) = \emptyset \}.$$

The following syntactic operation transforms a given formula φ into a sentence: letting n denote the greatest natural number such that v_n occurs free in φ , we write $\forall \varphi$ for the formula $\forall v_0, v_1, \dots, v_n \varphi$, which we call the *universal closure* of φ . Obviously, then, $\forall \varphi$ is a sentence.

The concepts of our formal language that we have introduced in this section depend upon three quantities, which we fixed earlier in this section:

$$\begin{aligned} &\text{the "arity" function } \lambda : I \rightarrow \mathbb{N}, \\ &\text{the "arity" function } \mu : J \rightarrow \mathbb{N}, \text{ and} \\ &\text{the index set } K. \end{aligned} \tag{1.2.0.10}$$

The entire construction of the language depends, therefore, on the triple

$$L = (\lambda, \mu, K). \tag{1.2.0.11}$$

(Observe that the index sets I and J can be recovered as the domains of definition of λ and μ .) When we wish to emphasize this dependence on L , we write

$$\text{Tm}(L), \text{Fml}(L), \text{Sent}(L)$$

instead of

$$\text{Tm}, \text{Fml}, \text{Sent}.$$

Since all of these concepts are already determined by L , we shall often refer to L itself as the “language.” By an *extended language* L' of L we mean a triple

$$L' = (\lambda', \mu', K')$$

such that

1. The function $\lambda' : I' \rightarrow \mathbb{N}$ extends λ , i.e. $I \subseteq I'$ and $\lambda'(i) = \lambda(i)$ for all $i \in I$.
2. The function $\mu' : J \rightarrow \mathbb{N}$ extends μ .
3. $K \subseteq K'$.

The following inclusions follow immediately from the definitions:

$$\text{Tm}(L) \subseteq \text{Tm}(L'), \quad \text{Fml}(L) \subseteq \text{Fml}(L'), \quad \text{Sent}(L) \subseteq \text{Sent}(L').$$

Observe further that the variables are the same in both languages. We write Vbl for the *set of variables*. We shall write $L \subseteq L'$ to indicate that L' is an extended language of L .

In the following chapters we shall often use the following abbreviations: For a finite *conjunction*

$$(\varphi_1 \wedge \cdots \wedge \varphi_n) \quad \text{we write} \quad \bigwedge_{i=1}^n \varphi_i,$$

and for a finite *disjunction*

$$(\psi_1 \vee \cdots \vee \psi_m) \quad \text{we write} \quad \bigvee_{j=1}^m \psi_j \quad (1.2.0.12)$$

(recall (1.2.0.6) and (1.2.0.7), respectively). If a formula φ has the forms

$$\bigwedge_{i=1}^n \bigvee_{j=1}^{m_i} \varphi_{ij} \quad \text{or} \quad \bigvee_{i=1}^n \bigwedge_{j=1}^{m_i} \varphi_{ij},$$

where each $m_i \geq 1$ and each φ_{ij} is an atomic or a negated atomic formula, then φ is said to be in *conjunctive normal form* or in *disjunctive normal form*, respectively. A formula φ is in *prenex normal form* if φ is of the form

$$Q_1 x_1 \cdots Q_n x_n \psi,$$

where each Q_i is either the symbol \forall or the symbol \exists , and ψ is quantifier-free.

1.3 Formal Proofs

Given a (formal) language $L = (\lambda, \mu, K)$, we now want to define the concept of a (formal) proof.

Let Σ be a set of formulae: $\Sigma \subseteq \text{Fml}(L)$. In a proof, we shall allow the elements of this set to appear as “axioms” so to speak. A sequence $\varphi_1, \dots, \varphi_n$ of formulae is called a *proof* (or *deduction*) of φ_n from Σ if, for each $i \in \{1, 2, \dots, n\}$:

$$\varphi_i \text{ belongs to } \Sigma, \quad (1.3.0.1)$$

$$\text{or } \varphi_i \text{ is a logical axiom,} \quad (1.3.0.2)$$

$$\text{or } \varphi_i \text{ arises from the application of a logical rule} \quad (1.3.0.3)$$

to members of the sequence with indices $< i$.

The last line, φ_n , is sometimes called the *end-formula* of the proof.

The concepts “logical axiom” and “logical rule” used in this definition must now be made precise. We subdivide the *logical axioms* into three categories:

tautologies,
quantifier axioms and

equality axioms.

As *logical rules*, we shall allow:

modus ponens and
the generalization rule.

We shall now define these various axioms and rules one by one.

In order to be able to define the concept of a tautology precisely, we must first give a brief introduction to the *language of sentential logic*. Its alphabet consists of

$$\rightarrow \quad \wedge \quad) \quad (\quad A_0 \quad A_1 \quad \dots \quad A_n \quad \dots \quad (n \in \mathbb{N}).$$

From this alphabet we construct *sentential forms*:

- (a) A_0, A_1, \dots are sentential forms.
- (b) If Φ, Ψ are sentential forms, then so are $\neg\Phi$ and $(\Phi \wedge \Psi)$.
- (c) No other strings of symbols are sentential forms.

The symbols A_0, A_1, \dots are called *sentential variables*. By a *truth assignment* \mathcal{H} of the variables A_0, A_1, \dots we mean a function from the set $\{A_0, A_1, \dots\}$ to the set $\{T, F\}$ of truth values T (= true) and F (= false). Thus, for every $n \in \mathbb{N}$, either $\mathcal{H}(A_n) = T$ or $\mathcal{H}(A_n) = F$. This truth assignment extends canonically from the set of variables to the set of all sentential forms, as follows:

$$\mathcal{H}(\neg\Phi) = \neg\mathcal{H}(\Phi) \tag{1.3.0.4}$$

$$\mathcal{H}(\Phi \wedge \Psi) = \mathcal{H}(\Phi) \cap \mathcal{H}(\Psi). \tag{1.3.0.5}$$

Here \neg and \cap are operations defined on the set $\{T, F\}$ by the following tables:

\neg	$\begin{array}{c cc} & T & F \\ \hline T & F & T \\ F & T & F \end{array}$	\cap	$\begin{array}{c cc} & T & F \\ \hline T & T & F \\ F & F & F \end{array}$
--------	--	--------	--

A sentential form Φ is called a *tautological form* if Φ receives the value T for every truth assignment \mathcal{H} . If the sentential form Φ contains exactly n distinct sentential variables, then for a proof that Φ is a tautological form, there are exactly 2^n cases to consider: for each variable there are just the two values T and F to “plug in”. This calculation can be carried out in general according to the schema of the following example. We test the sentential form

$$\neg((A_0 \wedge A_1) \wedge \neg A_0) \tag{1.3.0.6}$$

by means of the following table:

A_0	A_1	$(A_0 \wedge A_1)$	$\neg A_0$	$(A_0 \wedge A_1) \wedge \neg A_0$	$\neg((A_0 \wedge A_1) \wedge \neg A_0)$
T	T	T	F	F	T
T	F	F	F	F	T
F	T	F	T	F	T
F	F	F	T	F	T

Here the first line says that for each truth assignment \mathcal{H} with $\mathcal{H}(A_0) = \mathcal{H}(A_1) = T$, the above sentential form receives the value T . The subsequent lines are to be read similarly. Since the last column has all T s (and no F s), we conclude that (1.3.0.6) is, indeed, a tautological form.

Returning to our formal language L , an *instance* of a tautological form, or simply a *tautology*, is the formula obtained from a tautological form Φ by simply replacing each sentential variable in Φ by a formula of L . It should go without saying that different occurrences of the same sentential variable in Φ must be replaced by the same formula. Thus, if $\varphi, \psi \in \text{Fml}(L)$, then the formula

$$\neg((\varphi \wedge \psi) \wedge \neg\varphi)$$

is an example of a tautology (in view of (1.3.0.6)), and hence of a logical axiom. If we still employ the abbreviations introduced in Section 1.2 (specifically, (1.2.0.1)), then this axiom takes the form

$$(\varphi \wedge \psi) \rightarrow \varphi.$$

It is advisable to utilize such abbreviations whenever one has to check for a tautological form. The following table follows from the definitions:

φ	ψ	$(\varphi \vee \psi)$	$(\varphi \rightarrow \psi)$	$(\varphi \leftrightarrow \psi)$
T	T	T	T	T
T	F	T	F	F
F	T	T	T	F
F	F	F	T	T

With the help of this table, the calculations needed to test for tautological forms are shortened considerably. Another reason to employ such abbreviations is that they enable mathematicians (who are already familiar with most tautologies) to see them in their well-known form.

Next, the *quantifier axioms* are:

$$(A1) \quad \forall x \varphi \rightarrow \varphi(x/t), \text{ in case } t \text{ is free for } x \text{ in } \varphi \quad (1.3.0.7)$$

$$(A2) \quad \forall x (\varphi \rightarrow \psi) \rightarrow (\varphi \rightarrow \forall x \psi), \text{ in case } x \notin \text{Fr}(\varphi) \quad (1.3.0.8)$$

Here φ and ψ are formulae, x is any variable and t is any term. Actually, (A1) and (A2) each represent infinitely many axioms (just as each tautological form gives rise to infinitely many tautologies).

Next, the *equality axioms* are:

- (I1) $x \doteq x$
 (I2) $x \doteq y \rightarrow (x \doteq z \rightarrow y \doteq z)$
 (I3) $x \doteq y \rightarrow (R_i(v, \dots, x, \dots, u) \rightarrow R_i(v, \dots, y, \dots, u))$
 (I4) $x \doteq y \rightarrow f_j(v, \dots, x, \dots, u) \doteq f_j(v, \dots, y, \dots, u)$
- (1.3.0.9)

Note that (I3) and (I4) are actually *families* of axioms, one for each $i \in I$ and $j \in J$, respectively. The arity of R_i is $\lambda(i)$, and that of f_j is $\mu(j)$ (recall p. 10). In (I3) and (I4), x (which is an arbitrary variable) may be the variable in any position (even the first or last) in the list of variables of R_i or f_j , respectively; the variables in the other positions remain unchanged when x gets replaced by y .

Having thus completed our description of the logical axioms, we now describe the logical rules. First, *modus ponens* is a logical rule that can be applied to two lines of a proof in the event that one of those two lines has the form $\phi \rightarrow \psi$ and the other has the form ϕ , where $\phi, \psi \in \text{Fml}(L)$. The result of the application is then the formula ψ . We display this rule in the following form:

$$\frac{\phi \rightarrow \psi \quad \phi}{\psi} \quad (\text{MP})$$

The line ϕ_i of a proof ϕ_1, \dots, ϕ_n arises from application of *modus ponens* in case there are indices $j_1, j_2 < i$ such that ϕ_{j_1} has the form $\phi_{j_2} \rightarrow \phi_i$.

The *generalization rule* allows us to pass from a line of the form ϕ to a line $\forall x \phi$, where x is an arbitrary variable. The line ϕ_i of a proof ϕ_1, \dots, ϕ_n arises from application of the generalization rule in case there is an index $j < i$ such that ϕ_j has the form $\forall x \phi_j$. We display this rule in the following form:

$$\frac{\phi}{\forall x \phi} \quad (\forall)$$

Here, finally, the definition of a formal proof (from an axiom system Σ) is concluded. To elucidate this concept, we give a series of examples. The first examples all have the form of *derived rules* (with a single premise); that is, given a proof with a certain end-formula, these rules show how one could, independent of how the end-formula was obtained, append additional lines to the proof so as to obtain a proof of a certain new line.

First we prove the following derived rule:

$$\frac{\phi \wedge \psi}{\phi} \quad (\wedge B_1)$$

Assume we have a proof of $(\phi \wedge \psi)$ (from some set Σ); let this proof be, say,

$$\begin{array}{c}
\varphi_1 \\
\vdots \\
\varphi_{n-1} \\
(\varphi \wedge \psi) \quad ;
\end{array}$$

then we extend this proof in the following manner:

$$\begin{array}{c}
\varphi_1 \\
\vdots \\
\varphi_{n-1} \\
(\varphi \wedge \psi) \\
(\varphi \wedge \psi) \rightarrow \varphi \qquad (1.3.0.10) \\
\varphi \qquad (1.3.0.11)
\end{array}$$

(1.3.0.10) above is an instance of a tautological form. (1.3.0.11) arises from its two predecessors via an application of *modus ponens*.

By the same argumentation one obtains, in succession, the following derived rules:

$\frac{(\varphi \wedge \psi)}{\psi}$	$(\wedge B_2)$	from the tautology $(\varphi \wedge \psi) \rightarrow \psi$
$\frac{\varphi}{(\varphi \vee \psi)}$	$(\vee B_1)$	from the tautology $\varphi \rightarrow (\varphi \vee \psi)$
$\frac{\psi}{(\varphi \vee \psi)}$	$(\vee B_2)$	from the tautology $\psi \rightarrow (\varphi \vee \psi)$
$\frac{\varphi \rightarrow \psi}{\neg \psi \rightarrow \neg \varphi}$	(CP)	from the tautology $(\varphi \rightarrow \psi) \rightarrow (\neg \psi \rightarrow \neg \varphi)$
$\frac{\varphi \leftrightarrow \psi}{\varphi \rightarrow \psi}$	$(\leftrightarrow B_1)$	from the tautology $(\varphi \leftrightarrow \psi) \rightarrow (\varphi \rightarrow \psi)$
$\frac{\forall x \varphi}{\varphi(x/t)}$	$(\forall B)$	if t is free for x in φ
$\frac{\forall x (\varphi \rightarrow \psi)}{\varphi \rightarrow \forall x \psi}$	$(K\forall)$	if $x \notin Fr(\varphi)$

The last two derived rules above are obtained using the logical axioms (A1) and (A2), respectively, in just the same way that we obtained $(\wedge B_1)$

The following derived rules each have two premises; i.e. we assume that we have already proved two lines – the premises. First we consider the important rule of “chain implication”:

$$\frac{\varphi \rightarrow \psi \quad \psi \rightarrow \sigma}{\varphi \rightarrow \sigma} \quad (\text{KS})$$

This derived rule can be established as follows.

Suppose

$$\begin{array}{ccc} \varphi_1 & & \psi_1 \\ \vdots & \text{and} & \equiv \\ \varphi_{n-1} & & \psi_{m-1} \\ (\varphi \rightarrow \psi) & & (\psi \rightarrow \sigma) \end{array}$$

are proofs (say, from Σ_1 and Σ_2 , respectively). Then we obtain the following proof (from $\Sigma_1 \cup \Sigma_2$):

$$\begin{array}{l} \varphi_1 \\ \vdots \\ \varphi_{n-1} \\ (\varphi \rightarrow \psi) \end{array} \quad (1.3.0.12)$$

$$\begin{array}{l} \psi_1 \\ \equiv \\ \psi_{m-1} \\ (\psi \rightarrow \sigma) \end{array} \quad (1.3.0.13)$$

$$(\varphi \rightarrow \psi) \rightarrow ((\psi \rightarrow \sigma) \rightarrow (\varphi \rightarrow \sigma)) \quad (1.3.0.14)$$

$$(\psi \rightarrow \sigma) \rightarrow (\varphi \rightarrow \sigma) \quad (1.3.0.15)$$

$$(\varphi \rightarrow \sigma) \quad (1.3.0.16)$$

Here (1.3.0.14) is a tautology, (1.3.0.15) is obtained by applying *modus ponens* to (1.3.0.12) and (1.3.0.14), and (1.3.0.16) is obtained by applying *modus ponens* to (1.3.0.13) and (1.3.0.15).

In a similar way one obtains the following derived rules (each with two premises):

$$\frac{\varphi \rightarrow \psi \quad \psi \rightarrow \varphi}{\varphi \leftrightarrow \psi} \quad (\leftrightarrow) \quad \begin{array}{l} \text{from the tautology} \\ (\varphi \rightarrow \psi) \rightarrow ((\psi \rightarrow \varphi) \rightarrow (\varphi \leftrightarrow \psi)) \end{array}$$

$$\frac{\varphi \quad \psi}{(\varphi \wedge \psi)} \quad (\wedge) \quad \begin{array}{l} \text{from the tautology} \\ \varphi \rightarrow (\psi \rightarrow (\varphi \wedge \psi)) \end{array}$$

$$\frac{\varphi \rightarrow \sigma \quad \psi \rightarrow \sigma}{(\varphi \vee \psi) \rightarrow \sigma} \quad (\vee) \quad \begin{array}{l} \text{from the tautology} \\ (\varphi \rightarrow \sigma) \rightarrow ((\psi \rightarrow \sigma) \rightarrow ((\varphi \vee \psi) \rightarrow \sigma)) \end{array}$$

One could extend the list of derived rules arbitrarily. In fact, this is the method to make proofs more “bearable”. Again and again, arguments crop up that one does not want to repeat every time; rather, one incorporates them gradually into the logical system (as derived rules). The more advanced a mathematician is, the more he masters such rules, and the shorter his proofs become.

Before we give a formal proof of our example in Section 1.1, we wish to state, finally, the following derived rules:

$$\frac{t_1 \doteq t_2}{t_2 \doteq t_1} \quad (\text{S}) \qquad \frac{t_1 \doteq t_2 \quad t_2 \doteq t_3}{t_1 \doteq t_3} \quad (\text{Tr}) \quad (1.3.0.17)$$

$$\frac{t' \doteq t''}{R_i(t_1, \dots, t', \dots, t_{\lambda(i)}) \rightarrow R_i(t_1, \dots, t'', \dots, t_{\lambda(i)})} \quad (\text{R}_i) \quad (1.3.0.18)$$

$$\frac{t' \doteq t''}{f_j(t_1, \dots, t', \dots, t_{\mu(j)}) \doteq f_j(t_1, \dots, t'', \dots, t_{\mu(j)})} \quad (\text{f}_j)$$

We leave the proofs of (S), (Tr) and (f_j) to the reader; we now present that of (R_i). We shall carry out the replacement of t' by t'' in the first argument of R_i . It will be clear that we shall be able to carry out the replacement of every arbitrary argument of R_i by the same method. Thus we assume that we are given a proof of $t' \doteq t''$:

$$\begin{array}{c} \vdots \\ t' \doteq t'' \end{array} \quad (1.3.0.19)$$

We extend this proof by the following lines, where we choose the variables $x, y, u_2, \dots, u_{\lambda(i)}$ so that they do not appear in any of the terms $t', t'', t_2, \dots, t_{\lambda(i)}$:

$$x \doteq y \rightarrow (R_i(x, u_2, u_3, \dots) \rightarrow R_i(y, u_2, u_3, \dots)) \quad (1.3.0.20)$$

$$\forall x (x \doteq y \rightarrow (R_i(x, u_2, u_3, \dots) \rightarrow R_i(y, u_2, u_3, \dots)))$$

$$t' \doteq y \rightarrow (R_i(t', u_2, u_3, \dots) \rightarrow R_i(y, u_2, u_3, \dots))$$

$$\forall y (t' \doteq y \rightarrow (R_i(t', u_2, u_3, \dots) \rightarrow R_i(y, u_2, u_3, \dots)))$$

$$t' \doteq t'' \rightarrow (R_i(t', u_2, u_3, \dots) \rightarrow R_i(t'', u_2, u_3, \dots))$$

$$\forall u_2 (t' \doteq t'' \rightarrow (R_i(t', u_2, u_3, \dots) \rightarrow R_i(t'', u_2, u_3, \dots)))$$

$$t' \doteq t'' \rightarrow (R_i(t', t_2, u_3, \dots) \rightarrow R_i(t'', t_2, u_3, \dots))$$

$$\begin{array}{ccc} \vdots & \vdots & \vdots \end{array}$$

$$t' \doteq t'' \rightarrow (R_i(t', t_2, t_3, \dots) \rightarrow R_i(t'', t_2, t_3, \dots)) \quad (1.3.0.21)$$

$$R_i(t', t_2, t_3, \dots) \rightarrow R_i(t'', t_2, t_3, \dots) \quad (1.3.0.22)$$

Here (1.3.0.20) is an equality axiom (I3) (1.3.0.9). Thereafter we alternately used the rules (\forall) and ($\forall B$), until (1.3.0.21). We applied *modus ponens* to this line and (1.3.0.19) to obtain (1.3.0.22).

Now we want to give a formal proof of $\forall x \ 0 \leq x \cdot x$ from the axiom system $\Sigma = \{(1), \dots, (5)\}$:

- (1) $\forall x, y \ (x \leq y \vee y \leq x)$
- (2) $\forall x, y, z \ (x \leq y \rightarrow x + z \leq y + z)$
- (3) $\forall x, y \ (0 \leq x \wedge 0 \leq y \rightarrow 0 \leq x \cdot y)$
- (4) $\forall x, y \ (x + (-x) \doteq 0 \wedge 0 + y \doteq y)$
- (5) $\forall x \ (-x) \cdot (-x) \doteq x \cdot x$

These axioms are sentences of a language L whose symbols are specified as follows:

The index set I contains only one element (say, $I = \{0\}$), and $\lambda(0) = 2$; i.e. R_0 is a binary relation symbol. For the sake of readability we write \leq for R_0 . Here one thinks instinctively of the “less-than-or-equal-to” relation between real numbers, which brings with it the temptation to reason semantically. As agreed, however, we wanted to give a purely formal proof whose correctness could be checked even by a computer.

We want to retain for the function symbols the suggestive notation that we have begun to use for the relation symbol. Here J has three elements, say, $J = \{0, 1, 2\}$, and $\mu(0) = 1$, $\mu(1) = \mu(2) = 2$. For f_0, f_1, f_2 we write $-, +, \cdot$, respectively. Furthermore, we make use of the convention that $x + y$ is written for the term $+(x, y)$. Without these agreements, Axiom (2) would take the following form:

- (2) $\forall x, y, z \ (R_0(x, y) \rightarrow R_0(f_1(x, z), f_1(y, z)))$.

The index set K is again a singleton – say, $K = \{7\}$. For c_7 we write 0, for short.

When we now finally give a formal proof of $\forall x \ 0 \leq x \cdot x$ from Σ , we shall number the lines, and at the end of each line point out how it arose. Thus, for example, “(MP 3, 29)” in line 30 indicates that this line came about via an application of *modus ponens* to lines 3 and 29.

1. $\forall x, y \ (x \leq y \vee y \leq x)$ (Ax(1))
2. $\forall y \ (x \leq y \vee y \leq x)$ ($\forall B$ 1)
3. $(x \leq 0 \vee 0 \leq x)$ ($\forall B$ 2)
4. $\forall x, y \ (0 \leq x \wedge 0 \leq y \rightarrow 0 \leq x \cdot y)$ (Ax(3))
5. $\forall y \ (0 \leq x \wedge 0 \leq y \rightarrow 0 \leq x \cdot y)$ ($\forall B$ 4)
6. $(0 \leq x \wedge 0 \leq x \rightarrow 0 \leq x \cdot x)$ ($\forall B$ 5)
7. $0 \leq x \rightarrow 0 \leq x \wedge 0 \leq x$ (Taut.)
8. $0 \leq x \rightarrow 0 \leq x \cdot x$ (KS, 6, 7)
9. $\forall x, y, z \ (x \leq y \rightarrow x + z \leq y + z)$ (Ax(2))
10. $\forall y, z \ (x \leq y \rightarrow x + z \leq y + z)$ ($\forall B$ 9)
11. $\forall z \ (x \leq 0 \rightarrow x + z \leq 0 + z)$ ($\forall B$ 10)
12. $x \leq 0 \rightarrow x + (-x) \leq 0 + (-x)$ ($\forall B$ 11)

13.	$\forall x, y \ (x + (-x) \doteq 0 \wedge 0 + y \doteq y)$	(Ax(4))
14.	$\forall y \ (x + (-x) \doteq 0 \wedge 0 + y \doteq y)$	(\forall B 13)
15.	$(x + (-x) \doteq 0 \wedge 0 + (-x) \doteq -x)$	(\forall B 14)
16.	$x + (-x) \doteq 0$	(\wedge B ₁ 15)
17.	$x + (-x) \leq 0 + (-x) \rightarrow 0 \leq 0 + (-x)$	(R ₀ 16)
18.	$x \leq 0 \rightarrow 0 \leq 0 + (-x)$	(KS 12, 17)
19.	$0 + (-x) \doteq -x$	(\wedge B ₂ 15)
20.	$0 \leq 0 + (-x) \rightarrow 0 \leq -x$	(R ₀ 19)
21.	$x \leq 0 \rightarrow 0 \leq -x$	(KS 18, 20)
22.	$\forall x \ (0 \leq x \rightarrow 0 \leq x \cdot x)$	(\forall 8)
23.	$0 \leq -x \rightarrow 0 \leq (-x) \cdot (-x)$	(\forall B 22)
24.	$\forall x \ (-x) \cdot (-x) \doteq x \cdot x$	(Ax(5))
25.	$(-x) \cdot (-x) \doteq x \cdot x$	(\forall B 24)
26.	$0 \leq (-x) \cdot (-x) \rightarrow 0 \leq x \cdot x$	(R ₀ 25)
27.	$0 \leq -x \rightarrow 0 \leq x \cdot x$	(KS 23, 26)
28.	$x \leq 0 \rightarrow 0 \leq x \cdot x$	(KS 21, 27)
29.	$(x \leq 0 \vee 0 \leq x) \rightarrow 0 \leq x \cdot x$	(\vee 8, 28)
30.	$0 \leq x \cdot x$	(MP 3, 29)
31.	$\forall x \ 0 \leq x \cdot x$	(\forall 30)

We have thus, finally, transformed the proof given in the usual mathematical style in Section 1.2, into a formal proof. That this transformation has substantially increased the length of the proof is, as already explicitly mentioned earlier, due to the fact that we have presented derived rules only to a limited extent. Our ambition now, however, is not to gain further familiarity with formal proofs or to simplify them step by step until they are, finally, practicable. We want to leave this example as it is. Instead, we now occupy ourselves with the “reach” of such proofs. This will lead us to make claims in our metalanguage about formal proofs, which we then have to prove. The proofs will be carried out in the usual, informal mathematical style. In this section we wish to prove only a couple of small claims and the so-called deduction theorem.

First another definition. Let φ be an L -formula and Σ a set of L -formulae. Then we say that φ is *provable* (or *derivable*) *from* Σ , and write

$$\Sigma \vdash \varphi,$$

if there is a proof $\varphi_1, \dots, \varphi_n$ from Σ whose last formula φ_n is identical with φ .

It is clear that if $\Sigma \vdash \varphi$ holds, then for every set Σ' of formulae with $\Sigma \subseteq \Sigma'$, $\Sigma' \vdash \varphi$ also holds. In the following claims and their proofs, we become acquainted with further properties of the metalinguistic relation \vdash .

Lemma 1.3.1. *Let $\varphi, \psi \in \text{Fml}(L)$, $\Sigma \subseteq \text{Fml}(L)$ and $x \in \text{Vbl}$. Then the following hold:*

- (a) $\Sigma \vdash \phi$ if and only if $\Sigma \vdash \forall x \phi$
 (b) $\Sigma \cup \{\psi\} \vdash \phi$ if and only if $\Sigma \cup \{\forall x \psi\} \vdash \phi$.

Proof: (a) One reasons from left to right, thusly:

$$\text{If } \begin{array}{c} \vdots \\ \phi \end{array} \text{ is a proof from } \Sigma, \text{ then so is } \begin{array}{c} \vdots \\ \phi \\ \forall x \phi \end{array}.$$

From right to left we use (\forall B):

$$\text{If } \begin{array}{c} \vdots \\ \forall x \phi \end{array} \text{ is a proof from } \Sigma, \text{ then so is } \begin{array}{c} \vdots \\ \forall x \phi \\ \phi \end{array}.$$

Observe here that $\phi(x/x)$ is identical with ϕ , and that (\forall B) may be used, since x is naturally free for x in ϕ .

(b) One reasons from left to right, thusly:

$$\text{If } \begin{array}{c} \vdots \\ \psi \\ \equiv \\ \phi \end{array} \text{ is a proof from } \Sigma \cup \{\psi\}, \text{ then } \begin{array}{c} \vdots \\ \forall x \psi \\ \psi \\ \equiv \\ \phi \end{array} \text{ is a proof from } \Sigma \cup \{\forall x \psi\}.$$

Here the rule (\forall B) is again used. The reasoning from right to left goes as follows:

$$\text{If } \begin{array}{c} \vdots \\ \forall x \psi \\ \equiv \\ \phi \end{array} \text{ is a proof from } \Sigma \cup \{\forall x \psi\}, \text{ then } \begin{array}{c} \vdots \\ \psi \\ \forall x \psi \\ \equiv \\ \phi \end{array} \text{ is a proof from } \Sigma \cup \{\psi\}. \quad \square$$

By repeated use of Lemma 1.3.1(a), we see that the derivability of a formula ϕ from Σ is equivalent to the derivability of its universal closure $\forall \phi$ (p. 13) from Σ . Similarly, repeated use of Lemma 1.3.1(b) permits us to replace all formulae in Σ by their universal closures. On the basis of this, we shall often, in the future, limit ourselves to the case where Σ is a set of sentences, and ϕ is a sentence.

Now, however, again let $\phi, \psi \in \text{Fml}(L)$ and $\Sigma \subseteq \text{Fml}(L)$. If

$$\Sigma \vdash (\phi \rightarrow \psi)$$

holds, then one obtains immediately via (MP):

$$\Sigma \cup \{\varphi\} \vdash \psi.$$

Indeed, if

$$\begin{array}{c} \vdots \\ \varphi \rightarrow \psi \end{array}$$

is a proof from Σ , then

$$\begin{array}{c} \vdots \\ \varphi \rightarrow \psi \\ \varphi \\ \psi \end{array}$$

is a proof from $\Sigma \cup \{\varphi\}$. Here it is immaterial whether φ contains free variables or not. If, however, one knows that $\text{Fr}(\varphi) = \emptyset$, then the above implication can be reversed. We have the following theorem, which is very important for applications:

Theorem 1.3.2 (Deduction Theorem). *Let $\Sigma \subseteq \text{Fml}(L)$, $\varphi, \psi \in \text{Fml}(L)$ and $\text{Fr}(\varphi) = \emptyset$. Then from $\Sigma \cup \{\varphi\} \vdash \psi$ we obtain $\Sigma \vdash (\varphi \rightarrow \psi)$.*

Proof: We shall show, by induction on n :

$$\begin{array}{ll} \varphi_1 & \varphi \rightarrow \varphi_1 \\ \text{if } \vdots & \text{is a proof from } \Sigma \cup \{\varphi\}, \text{ then the sequence of formulae } \vdots \\ \varphi_n & \varphi \rightarrow \varphi_n \end{array}$$

can be completed so as to become a proof from Σ whose last line remains $\varphi \rightarrow \varphi_n$. It is clear that the claim of the Deduction Theorem will follow from this.

Induction basis step: Since $n = 1$, we are given a one-line proof from Σ , consisting of the line φ_1 .

Case 1: φ_1 is a logical axiom or a member of Σ . In this case,

$$\begin{array}{c} \varphi_1 \\ \varphi_1 \rightarrow (\varphi \rightarrow \varphi_1) \\ \varphi \rightarrow \varphi_1 \end{array}$$

is clearly a proof from Σ . Here we obtained the last line from φ_1 and the tautology $\varphi_1 \rightarrow (\varphi \rightarrow \varphi_1)$ via (MP).

Case 2: φ_1 is identical with φ . In this case the implication $\varphi \rightarrow \varphi_1$ is a tautology, and hence in particular a proof from Σ .

Step from n to $n + 1$: Assume that

$$\begin{array}{ll} \varphi_1 & \varphi \rightarrow \varphi_1 \\ \vdots & \text{is a proof from } \Sigma \cup \{\varphi\}, \text{ and the sequence of formulae } \vdots \\ \varphi_n & \varphi \rightarrow \varphi_n \end{array}$$

has already been completed to a proof from Σ , with end-formula $\varphi \rightarrow \varphi_n$.

Case 1: φ_{n+1} is a logical axiom or a member of Σ . In this case we extend the right-hand proof with the lines

$$\begin{aligned}\varphi_{n+1} \\ \varphi_{n+1} \rightarrow (\varphi \rightarrow \varphi_{n+1}) \\ \varphi \rightarrow \varphi_{n+1},\end{aligned}$$

and again obtain a proof from Σ .

Case 2: φ_{n+1} is identical with φ . In this case we simply add the tautology $\varphi \rightarrow \varphi_{n+1}$ as the last line.

Case 3: φ_{n+1} is obtained by (MP). In this case there are $i, j \leq n$ such that the formula φ_j has the form $\varphi_i \rightarrow \varphi_{n+1}$. But then the lines $\varphi \rightarrow \varphi_i$ and $\varphi \rightarrow (\varphi_i \rightarrow \varphi_{n+1})$ occur in the right-hand proof. We extend the right-hand proof with the following lines:

$$(\varphi \rightarrow (\varphi_i \rightarrow \varphi_{n+1})) \rightarrow (\varphi \rightarrow \varphi_i) \rightarrow (\varphi \rightarrow \varphi_{n+1}) \quad (1.3.2.1)$$

$$(\varphi \rightarrow \varphi_i) \rightarrow (\varphi \rightarrow \varphi_{n+1}) \quad (1.3.2.2)$$

$$\varphi \rightarrow \varphi_{n+1}. \quad (1.3.2.3)$$

Here (1.3.2.1) is a tautology, and (1.3.2.2) and (1.3.2.3) are obtained via applications of (MP).

Case 4: φ_{n+1} is obtained via (\forall) . In this last case there exists $i \leq n$ such that φ_{n+1} has the form $\forall x \varphi_i$, where x is a variable. If we extend the right-hand proof with the lines $\forall x(\varphi \rightarrow \varphi_i)$ and $\varphi \rightarrow \varphi_{n+1}$ (in that order), then we shall again obtain a proof of $\varphi \rightarrow \varphi_{n+1}$ from Σ , since these two new lines are justified by the rule (\forall) and the derived rule $(K\forall)$, respectively. This application of $(K\forall)$ is correct, since obviously $x \notin \text{Fr}(\varphi)$ (by the hypothesis that $\text{Fr}(\varphi) = \emptyset$) and φ_{n+1} is $\forall x \varphi_i$. \square

1.4 Completeness of First-Order Logic

The preceding sections have shown that it is possible to give a strict, formal definition of the concept of a mathematical proof. It remains to clarify the question of whether this definition really captures what one ordinarily understands by a proof. The unwieldiness that came to light in our earlier example (proving $0 \leq x \cdot x$ from the axioms for ordered fields) can, in principle, be eliminated by the introduction of more and more derived rules. Thus, this unwieldiness is no genuine argument against our formal system. A further possible objection could be that the formal languages we use have inadequate expressive power. But this objection can also be refuted: in Section 1.6 we shall formalize set theory in such a language. Set theory has enough expressive power to express every reasonable mathematical concept.

A completely different objection could be brought against the strength of such formal proofs. It is conceivable that the definition of proof that we gave in the previous section overlooks some valid form of logical reasoning. We wish to show now

that this is not the case – i.e. that our concept of proof *completely* comprehends all valid forms of logical reasoning. A heuristic reflection should elucidate this claim.

Let $\Sigma \subseteq \text{Sent}(L)$ and $\varphi \in \text{Sent}(L)$. We ask:

How might it happen that φ is not provable from Σ – i.e. $\Sigma \not\vdash \varphi$? (1.4.0.1)

One possibility could be that there is a “counterexample”. This should mean that a domain (a mathematical structure – cf. Section 1.5) could exist in which all axioms $\sigma \in \Sigma$ hold, but not φ . Here we implicitly assume that formal proofs are sound, i.e. that everything that is provable holds wherever the axioms hold. (In Section 1.5 we shall make this precise and prove it.) Another possible reason for the unprovability of φ from Σ could be that we forgot some method of reasoning when we defined “proof”, in which case φ could be unprovable even though there is no “counterexample”. We shall show that this second case cannot arise:

Any unprovability rests necessarily on a counterexample. (1.4.0.2)

We shall prove this claim in Theorem 1.5.2 (Gödel’s Completeness Theorem) in the next section, but we shall need several technical preparations, which we would like to carry out in this section. Although we postpone to the next section a precise description of the concept of a structure, and of the definition of the satisfaction of a formula in such a structure, the “counterexample” that we shall construct from the assumption that $\Sigma \not\vdash \varphi$ will take on a clear form already by the end of this section.

First we would like to undertake a small reformulation of the hypothesis. For this we call a set $\Sigma \subseteq \text{Sent}(L)$ *consistent* if there is no L -sentence α for which both

$$\Sigma \vdash \alpha \quad \text{and} \quad \Sigma \vdash \neg \alpha$$

hold simultaneously. If there is such an α , then Σ is called *inconsistent*. Then obviously:

Σ is inconsistent if and only if one can prove every sentence β from Σ . (1.4.0.3)

Indeed, if Σ is inconsistent, then there is a proof of $(\alpha \wedge \neg \alpha)$ from Σ , by rule (\wedge) . We extend this proof by the lines

$$(\alpha \wedge \neg \alpha) \rightarrow \beta \tag{1.4.0.4}$$

$$\beta. \tag{1.4.0.5}$$

Here (1.4.0.4) is a tautology, and (1.4.0.5) is obtained with (MP). Using (1.4.0.3) we now show:

Lemma 1.4.1. *Let $\Sigma \subseteq \text{Sent}(L)$ and $\varphi \in \text{Sent}(L)$. Then $\Sigma \not\vdash \varphi$ if and only if $\Sigma \cup \{\neg \varphi\}$ is consistent.*

Proof: We show that $\Sigma \vdash \varphi$ is equivalent to $\Sigma \cup \{\neg \varphi\}$ being inconsistent. First suppose $\Sigma \vdash \varphi$. Then $\Sigma \cup \{\neg \varphi\} \vdash \varphi$ on the one hand, and in any case $\Sigma \cup \{\neg \varphi\} \vdash \neg \varphi$ on the other, whence $\Sigma \cup \{\neg \varphi\}$ is inconsistent.

Now suppose $\Sigma \cup \{\neg\varphi\}$ is inconsistent. Then $\Sigma \cup \{\neg\varphi\} \vdash \varphi$, by (1.4.0.3). Using the Deduction Theorem (1.3.2) we obtain

$$\Sigma \vdash (\neg\varphi \rightarrow \varphi).$$

We take some particular proof of $(\neg\varphi \rightarrow \varphi)$ from Σ , and extend it with the lines

$$(\neg\varphi \rightarrow \varphi) \rightarrow \varphi \quad (1.4.1.1)$$

$$\varphi. \quad (1.4.1.2)$$

Here (1.4.1.1) is a tautology, and (1.4.1.2) is obtained via (MP). Thus $\Sigma \vdash \varphi$ holds. \square

Our assumption $\Sigma \not\vdash \varphi$ is therefore equivalent to the consistency of the set $\Sigma \cup \{\neg\varphi\}$ of sentences. On the other hand, a “counterexample” to $\Sigma \vdash \varphi$ is exactly a domain in which all $\sigma \in \Sigma \cup \{\neg\varphi\}$ hold. (If φ does not hold, then obviously $\neg\varphi$ holds.) In order to produce the “completeness proof” of (1.4.0.2) that we seek, it therefore clearly suffices to do the following:

$$\begin{aligned} & \text{to construct, for any consistent set}^2 \Sigma \text{ of sentences,} \\ & \text{a domain in which all } \sigma \in \Sigma \text{ hold.} \end{aligned} \quad (1.4.1.3)$$

This is what we would like to do now. For this, we shall pursue the following strategy. By means of a systematic, consistent expansion of the set Σ , we wish to determine the desired domain as far as possible. The steps toward this goal are of a rather technical nature and will be completely motivated and clear only later.

In the *first step* of (1.4.1.3) (which is also the most difficult), we wish to arrive at a stage in which, whenever an existence sentence holds in the domain to be constructed, this can be verified by an example. One such example should be capable of being named by a constant symbol c_k in our language. To achieve this, we shall be forced to extend the given language L by the addition of new constant symbols. We prove the following:

Theorem 1.4.2. *Let $\Sigma \subseteq \text{Sent}(L)$ be consistent. Then there is a language $L' \supseteq L$ with $I' = I$ and $J' = J$, and there is a consistent set $\Sigma' \subseteq \text{Sent}(L')$ with $\Sigma \subseteq \Sigma'$, such that to each L' -existence sentence³ $\exists x \varphi$, there exists a $k \in K'$ such that*

$$(\exists x \varphi \rightarrow \varphi(x/c_k)) \text{ is a member of } \Sigma'.$$

For the proof of this theorem we need the following:

Lemma 1.4.3. *Let $L^{(1)} \subseteq L^{(2)}$ be two languages with $I^{(1)} = I^{(2)}$, $J^{(1)} = J^{(2)}$ and $K^{(2)} = K^{(1)} \cup \{0\}$, where $0 \notin K^{(1)}$. Further, let $\varphi_1, \dots, \varphi_n$ be a proof in $L^{(2)}$ from $\Sigma := \{\varphi_1, \dots, \varphi_m\}$, with $m \leq n$. Then, if y is a variable not occurring in any φ_i ($1 \leq i \leq n$), then $\varphi_1(c_0/y), \dots, \varphi_n(c_0/y)$ is a proof in $L^{(1)}$ from $\{\varphi_1(c_0/y), \dots, \varphi_m(c_0/y)\}$.*

² Previously $\Sigma \cup \{\neg\varphi\}$.

³ Note that this notion differs from that of an “ \exists -sentence” introduced in Theorem 2.5.4.

Proof (of Lemma 1.4.3): We apply induction on the length n of the given formal proof.

Basis step: If, in the case $n = 1$, also $m = 1$, then there is nothing to show. If, on the other hand, $m = 0$ (i.e. $\Sigma = \emptyset$), then φ_n must be a logical axiom. In the case of an equality axiom, there is again nothing to show, since such an axiom contains no constant symbols. If φ_n is an instance of a tautological form, then $\varphi_n(c_0/y)$ is, likewise, clearly an instance of the same tautological form. There remains the case where φ_n is a quantifier axiom. So let φ_n be of the form (A1) (1.3.0.7):

$$\forall x \psi \rightarrow \psi(x/t),$$

where t is a term free for x in ψ . Now one can easily convince oneself that:

$$\psi(x/t)(c_0/y) \text{ is just } \psi(c_0/y)(x/t(c_0/y)).$$

Since y does not occur in φ_n by hypothesis, $t(c_0/y)$ is free for x in $\psi(c_0/y)$. Therefore $\varphi_n(c_0/y)$ takes the form of an axiom (A1), namely,

$$\forall x \psi(c_0/y) \rightarrow \psi(c_0/y)(x/t(c_0/y)).$$

If φ_n has the form of (A2) (1.3.0.8), one may convince oneself just as easily that $\varphi_n(c_0/y)$ is again an axiom of type (A2).

Step from $n - 1$ to n : Already we know that $\varphi_1(c_0/y), \dots, \varphi_{n-1}(c_0/y)$ is a proof in $L^{(1)}$ from $\varphi_1(c_0/y), \dots, \varphi_m(c_0/y)$, where we can assume, without loss of generality, that $m \leq n - 1$. Now if φ_n is a logical axiom, then, as we saw above, $\varphi_n(c_0/y)$ is again a logical axiom. Two cases remain, in which φ_n is obtained by a rule.

Case 1: φ_n is obtained by (MP). In this case there are $i, j \leq n - 1$ such that φ_j has the form $(\varphi_i \rightarrow \varphi_n)$. Then $\varphi_j(c_0/y)$ has the form $\varphi_i(c_0/y) \rightarrow \varphi_n(c_0/y)$. Therefore $\varphi_n(c_0/y)$ is likewise obtained by (MP).

Case 2: φ_n is obtained via (\forall) . In this case there is an $i \leq n - 1$ such that φ_n has the form $\forall x \varphi_i$. Then $\varphi_n(c_0/y)$ has the form $\forall x \varphi_i(c_0/y)$. Therefore $\varphi_n(c_0/y)$ is likewise obtained via (\forall) .

We observe, finally, that for every $L^{(2)}$ -formula ψ , the replacement of c_0 by a variable leads to an $L^{(1)}$ -formula. Thus it is clear that the resulting proof is in $L^{(1)}$. \square

Proof (Theorem 1.4.2): We shall obtain the language L' and the set Σ' of sentences by a countable process. For each $n \in \mathbb{N}$, we recursively construct a language L_n in the following way: let L_0 be the language L . For $n \geq 1$, if L_{n-1} has already been constructed, then we obtain L_n by setting $I_n = I_{n-1}$, $J_n = J_{n-1}$ and $K_n = K_{n-1} \cup M_n$. Here M_n is a set disjoint from K_{n-1} such that there is a bijection

$$g_n : M_n \rightarrow \{ \exists x \varphi \mid \exists x \varphi \text{ is a member of } \text{Sent}(L_{n-1}) \} \quad (1.4.3.1)$$

from M_n to the set of all existence sentences in the language L_{n-1} . This means nothing more than that we “enumerate” all existence sentences in L_{n-1} with new

indices in a one-to-one and onto manner. Sets M_n and bijections g_n of the required kind always exist.

We thereby obtain an ascending chain

$$L_0 \subseteq L_1 \subseteq \cdots \subseteq L_{n-1} \subseteq L_n \subseteq \cdots \quad (1.4.3.2)$$

of languages. Finally we set $L' = \bigcup_{n \in \mathbb{N}} L_n$, i.e. we set $I' = I$, $J' = J$ and $K' = \bigcup_{n \in \mathbb{N}} K_n$. From this one sees immediately that

$$\text{Sent}(L') = \bigcup_{n \in \mathbb{N}} \text{Sent}(L_n)$$

also holds. Therefore if $\exists x \varphi$ is an L' -sentence, then it lies already in a set $\text{Sent}(L_{n-1})$ for some $n \in \mathbb{N}$.

Suppose

$$\Sigma = \Sigma_0 \subseteq \Sigma_1 \subseteq \cdots \subseteq \Sigma_{n-1} \subseteq \Sigma_n \subseteq \cdots \quad (1.4.3.3)$$

is an ascending chain of sets such that for each $n \in \mathbb{N}$:

- (1) $\Sigma_n \subseteq \text{Sent}(L_n)$, and
 - (2) for each $k \in M_n$, $(\exists x \varphi \rightarrow \varphi(x/c_k))$ is a member of Σ_n ,
- (1.4.3.4)

where $\exists x \varphi$ is $g_n(k)$. Once we have such a chain, we shall be able to take Σ' to be $\bigcup_{n \in \mathbb{N}} \Sigma_n$. Then it will remain only to check that Σ' is consistent.

We obtain a chain (1.4.3.3) with satisfying (1.4.3.4) by setting

$$\Sigma_n := \Sigma_{n-1} \cup \{ (\exists x \varphi \rightarrow \varphi(x/c_k)) \mid k \in M_n, g_n(k) \text{ is } \exists x \varphi \}. \quad (1.4.3.5)$$

Since g_n is surjective, each existence sentence $\exists x \varphi$ in $\text{Sent}(L_{n-1})$ gets counted in (1.4.3.5) – $\exists x \varphi$ is, say, $g_n(k)$. Since $\text{Fr}(\varphi) \subseteq \{x\}$, $\varphi(x/c_k)$ is again a sentence (but now in L_n , not L_{n-1}). Therefore $\Sigma_n \subseteq \text{Sent}(L_n)$. The most important property of Σ_n is now its consistency. This results by induction on n .

For $n = 0$, the consistency of Σ_0 is the hypothesis. For $n \geq 1$, suppose that Σ_{n-1} is consistent, but not Σ_n . Then we would obtain an α in $\text{Sent}(L_n)$ with

$$\Sigma_n \vdash (\alpha \wedge \neg \alpha).$$

Since any single proof from Σ_n can be traced back to only finitely many axioms of Σ_n , $(\alpha \wedge \neg \alpha)$ would already be provable from Σ_{n-1} , together with finitely many sentences

$$(\exists x_1 \varphi_1 \rightarrow \varphi_1(x_1/c_{k_1})), \dots, (\exists x_r \varphi_r \rightarrow \varphi_r(x_r/c_{k_r})), \quad (1.4.3.6)$$

where $g_n(k_i)$ is $\exists x_i \varphi_i$, which is a member of $\text{Sent}(L_{n-1})$ for $1 \leq i \leq r$. We briefly write $\sigma_1, \dots, \sigma_r$ for the r sentences in (1.4.3.6). Then we have

$$\Sigma_{n-1} \cup \{\sigma_1, \dots, \sigma_r\} \vdash (\alpha \wedge \neg \alpha).$$

By means of a possible (still finite) expansion of the set $\{\sigma_1, \dots, \sigma_r\}$ we can ensure, in addition, that there is a proof of $(\alpha \wedge \neg\alpha)$ already in the sublanguage $L^{(2)}$ of L_n defined by

$$I^{(2)} = I_n, \quad J^{(2)} = J_n, \quad \text{and} \quad K^{(2)} = K_{n-1} \cup \{g_n^{-1}(\exists x_1 \varphi_1), \dots, g_n^{-1}(\exists x_r \varphi_r)\}.$$

Note that $r \geq 1$, since otherwise Σ_{n-1} would be inconsistent. Thus, by the Deduction Theorem 1.3.2, we obtain

$$\Sigma_{n-1} \cup \{\sigma_2, \dots, \sigma_r\} \vdash (\sigma_1 \rightarrow (\alpha \wedge \neg\alpha)),$$

and, by use of the tautology

$$((\beta \rightarrow \gamma) \rightarrow (\alpha \wedge \neg\alpha)) \rightarrow (\beta \wedge \neg\gamma)$$

and (MP) it follows, finally, that

$$\Sigma_{n-1} \cup \{\sigma_2, \dots, \sigma_r\} \vdash (\exists x_1 \varphi_1 \wedge \neg\varphi_1(x_1/c_{k_1})).$$

If one bears in mind that $\exists x_1$ is an abbreviation for $\neg\forall x_1\neg$, then we obtain via $(\wedge B_1)$, on the one hand,

$$\Sigma_{n-1} \cup \{\sigma_2, \dots, \sigma_r\} \vdash \neg\forall x_1 \neg\varphi_1, \tag{1.4.3.7}$$

and, via $(\wedge B_2)$ on the other hand,

$$\Sigma_{n-1} \cup \{\sigma_2, \dots, \sigma_r\} \vdash \neg\varphi_1(x_1/c_{k_1}). \tag{1.4.3.8}$$

The provabilities asserted in (1.4.3.7) and (1.4.3.8) are meant in the language $L^{(2)}$. If we now define $L^{(1)}$ by

$$I^{(1)} = I_n, \quad J^{(1)} = J_n, \quad \text{and} \quad K^{(1)} = K_{n-1} \cup \{g_n^{-1}(\exists x_2 \varphi_2), \dots, g_n^{-1}(\exists x_r \varphi_r)\},$$

then we recognize that $\neg\forall x_1 \neg\varphi_1$ as well as the set of sentences

$$\Pi := \Sigma_{n-1} \cup \{\sigma_2, \dots, \sigma_r\}$$

already lie in $\text{Sent}(L^{(1)})$. Applying (a suitable version of) Lemma 1.4.3 to the proofs whose existence is asserted by (1.4.3.7) and (1.4.3.8), we obtain, on the one hand, a deduction

$$\Pi \vdash \neg\forall x_1 \neg\varphi_1$$

in $L^{(1)}$, and, on the other hand, a deduction

$$\Pi \vdash \neg\varphi_1(x_1/c_{k_1})(c_{k_1}/y)$$

likewise in $L^{(1)}$. Here y is a suitably chosen “new” variable (i.e. a variable not occurring in the proof of $\neg\varphi_1(x_1/c_{k_1})$ or in the proof of $\neg\forall x_1\neg\varphi_1$ from Π , to which we applied Lemma 1.4.3). Now it is obvious that

$$\varphi_1(x_1/c_{k_1})(c_{k_1}/y) \text{ is } \varphi_1(x_1/y),$$

since φ_1 is a member of $\text{Fml}(L_{n-1})$. We therefore have

$$\Pi \vdash \neg\varphi_1(x_1/y).$$

Via an application of (\forall) on y and $(\forall B)$ we obtain, first,

$$\Pi \vdash \forall y\neg\varphi_1(x_1/y)$$

and then

$$\Pi \vdash \neg\varphi_1(x_1/y)(y/x_1).$$

If one considers that y is new for φ_1 , then one understands immediately that

$$\varphi_1(x_1/y)(y/x_1) \text{ is } \varphi_1.$$

Thus we have $\Pi \vdash \neg\varphi_1$ and thus we finally obtain

$$\Pi \vdash \forall x_1\neg\varphi_1.$$

This derivability, together with the derivability

$$\Pi \vdash \neg\forall x_1\neg\varphi_1$$

shows that Π is inconsistent in $L^{(1)}$.

Just as we have reduced the inconsistency of $\Sigma_{n-1} \cup \{\sigma_1, \dots, \sigma_r\}$ to that of $\Sigma_{n-1} \cup \{\sigma_2, \dots, \sigma_r\}$, we can, through iteration, finally deduce a contradiction already in Σ_{n-1} . Since this contradicts our hypothesis, the consistency of Σ_n follows. In this way, all the Σ_n are recognized as consistent.

Now the consistency of $\Sigma' = \bigcup_{n \in \mathbb{N}} \Sigma_n$ is seen as follows: since the proof of a contradiction from Σ' is a finite sequence of formulae, and both the languages L_n as well as the sets Σ_n form ascending chains, there is an $n \in \mathbb{N}$ such that this proof is already a proof from Σ_n in the language L_n . This is, however, impossible, because of the previously proved consistency of Σ_n . \square

We shall carry out the *second step* in determining a domain in which all sentences in our consistent set Σ will hold (1.4.1.3), in the extension language L' constructed just above. For this we shall apply the following theorem, which we formulate for an arbitrary language (again denoted by L).

Theorem 1.4.4. *To each consistent set $\Sigma \subseteq \text{Sent}(L)$ there is a maximal consistent extension $\Sigma^* \subseteq \text{Sent}(L)$ of Σ ; this means that $\Sigma \subseteq \Sigma^*$, Σ^* is consistent, and whenever $\Sigma^* \subseteq \Sigma_1 \subseteq \text{Sent}(L)$ and Σ_1 is consistent, $\Sigma^* = \Sigma_1$.*

Proof: We consider the system

$$\mathfrak{M} = \{ \Sigma_1 \subseteq \text{Sent}(L) \mid \Sigma \subseteq \Sigma_1, \Sigma_1 \text{ consistent} \}.$$

Since $\Sigma \in \mathfrak{M}$, \mathfrak{M} is not empty. If we are given a subsystem $\mathfrak{M}' \subseteq \mathfrak{M}$ that is linearly ordered by inclusion (i.e. for $\Sigma_1, \Sigma_2 \in \mathfrak{M}'$, either $\Sigma_1 \subseteq \Sigma_2$ or $\Sigma_2 \subseteq \Sigma_1$), then the set

$$\Sigma' := \bigcup_{\Sigma_1 \in \mathfrak{M}'} \Sigma_1$$

is clearly an upper bound for \mathfrak{M}' in \mathfrak{M} . To see this, note first that for each $\Sigma_1 \in \mathfrak{M}'$, $\Sigma_1 \subseteq \Sigma'$, trivially. The consistency of Σ' rests simply on the finiteness of a proof of a hypothetical contradiction from Σ' : such a proof can utilize only finitely many axioms $\sigma_1, \dots, \sigma_n \in \Sigma'$. Each of the σ_i lies in a member of the system \mathfrak{M}' , say, $\sigma_i \in \Sigma_i$. Since, however, the sets $\Sigma_1, \dots, \Sigma_n$ are comparable, one of them, say Σ_n , must contain all the others as subsets. Then the hypothetical contradiction would be deducible already from Σ_n , which is impossible.

We have therefore shown that the system \mathfrak{M} fulfils the hypotheses of Zorn's lemma. Therefore there is a maximal element Σ^* in \mathfrak{M} . Then, according to the definition of \mathfrak{M} , $\Sigma \subseteq \Sigma^*$ and Σ^* is consistent. \square

Remark 1.4.5 *If the language L is countable, i.e. the sets I , J and K are (finite or) countable, then Zorn's lemma can be avoided in the proof of the above lemma.*

Proof (of 1.4.5): In this case we can start with an enumeration $(\varphi_n)_{n \in \mathbb{N}}$ of all L -sentences. Then we recursively define

$$\begin{aligned} \Sigma_0 &= \Sigma \\ \Sigma_{n+1} &= \begin{cases} \Sigma_n & \text{in case } \Sigma_n \cup \{\varphi_n\} \text{ is inconsistent, and} \\ \Sigma_n \cup \{\varphi_n\} & \text{otherwise.} \end{cases} \end{aligned} \quad (1.4.5.1)$$

Thus we obtain an ascending chain

$$\Sigma_0 \subseteq \Sigma_1 \subseteq \dots \subseteq \Sigma_n \subseteq \Sigma_{n+1} \subseteq \dots$$

of consistent sets of sentences. From this it follows, as before, that also

$$\Sigma^* := \bigcup_{n \in \mathbb{N}} \Sigma_n$$

is consistent. Because of $\Sigma = \Sigma_0 \subseteq \Sigma^*$, it remains only to show the maximality of Σ^* . Assume there were a $\varphi \in \text{Sent}(L)$ such that $\Sigma^* \cup \{\varphi\}$ were still consistent. This sentence φ occurs in the enumeration $(\varphi_n)_{n \in \mathbb{N}}$ of all L -sentences. Let us say φ is φ_n . From the consistency of $\Sigma^* \cup \{\varphi\}$ follows that of $\Sigma^* \cup \{\varphi_n\}$ and, *a fortiori*, that of $\Sigma_n \cup \{\varphi_n\}$. Therefore

$$\Sigma_n \cup \{\varphi_n\} = \Sigma_{n+1} \subseteq \Sigma^*,$$

by (1.4.5.1). From this follows $\varphi_n \in \Sigma^*$. Thus Σ^* is maximal (as well as consistent). \square

Now we apply Theorem 1.4.4 to the consistent set $\Sigma' \subseteq \text{Sent}(L')$ obtained in Theorem 1.4.2, in order to obtain a maximal consistent extension $\Sigma^* \subseteq \text{Sent}(L')$ of Σ' . For such a Σ^* we have:

- (I) Σ^* is maximal consistent in $\text{Sent}(L')$;
- (II) for each $\exists x \varphi$ in $\text{Sent}(L')$ there is a $k \in K'$ with $(\exists x \varphi \rightarrow \varphi(x/c_k))$ in Σ^* .

These two properties of Σ^* canonically determine a domain in which all sentences $\sigma \in \Sigma^*$ hold, as we shall see. In particular, all $\sigma \in \Sigma$ will hold there.

We first consider the set of *constant L' -terms*:

$$\text{CT} := \{t \in \text{Tm}(L') \mid \text{no variable occurs in } t\}. \quad (1.4.5.2)$$

CT contains, in particular, all c_k with $k \in K'$. We define a binary relation \approx on CT: for $t_1, t_2 \in \text{CT}$, we set

$$t_1 \approx t_2 \quad \text{iff} \quad \Sigma^* \vdash t_1 \doteq t_2. \quad (1.4.5.3)$$

With the help of axiom (II) (1.3.0.9) and Rules (S) and (Tr), we recognize immediately that \approx is an equivalence relation on CT, i.e. for $t_1, t_2, t_3 \in \text{CT}$:

- (i) $t_1 \approx t_1$,
- (ii) if $t_1 \approx t_2$, then $t_2 \approx t_1$, and
- (iii) if $t_1 \approx t_2$ and $t_2 \approx t_3$, then $t_1 \approx t_3$.

Now the sought-for domain is the set

$$A := \text{CT}/\approx \quad (1.4.5.4)$$

of all equivalence classes \bar{t} of constant terms. Here, as usual, we define for $t \in \text{CT}$:

$$\bar{t} = \{t_1 \in \text{CT} \mid t \approx t_1\}. \quad (1.4.5.5)$$

Then

$$\bar{t}_1 = \bar{t}_2 \quad \text{iff} \quad t_1 \approx t_2. \quad (1.4.5.6)$$

In order to be able to speak meaningfully of the truth of a sentence in the domain (something that we shall make precise only in the next section), we must say which relations, functions and individuals the symbols R_i , f_j and c_k name – i.e. we must give an interpretation of these symbols.

To each $i \in I$ we define a $\lambda(i)$ -place *relation* \mathcal{R}_i on the domain A , by declaring, for term-classes $\bar{t}_1, \dots, \bar{t}_{\lambda(i)}$, that

$$\mathcal{R}_i(\overline{t_1}, \dots, \overline{t_{\lambda(i)}}) \text{ holds if and only if } \Sigma^* \vdash R_i(t_1, \dots, t_{\lambda(i)}). \quad (1.4.5.7)$$

Here the notation $\mathcal{R}_i(\overline{t_1}, \dots, \overline{t_{\lambda(i)}})$ means, as usual, that the relation \mathcal{R}_i holds at the $\lambda(i)$ -tuple $(\overline{t_1}, \dots, \overline{t_{\lambda(i)}})$ of term-classes; i.e. that $(\overline{t_1}, \dots, \overline{t_{\lambda(i)}}) \in \mathcal{R}_i$. One should observe, however, that the above definition of \mathcal{R}_i refers back to a particular choice of *representatives* t_v of the term-classes $\overline{t_v}$. It must be shown that a choice of other representatives leads to the same definition. Thus, suppose $t_1 \approx t'_1, \dots, t_{\lambda(i)} \approx t'_{\lambda(i)}$; then we must show:

$$\Sigma^* \vdash R_i(t_1, \dots, t_{\lambda(i)}) \quad \text{iff} \quad \Sigma^* \vdash R_i(t'_1, \dots, t'_{\lambda(i)}). \quad (1.4.5.8)$$

By symmetry, it obviously suffices to show only one direction. So let us assume that

$$\Sigma^* \vdash R_i(t_1, \dots, t_{\lambda(i)}).$$

Along with this deducibility we have, according to the hypothesis, the deducibilities

$$\Sigma^* \vdash t_v \doteq t'_v, \quad \text{for } 1 \leq v \leq \lambda(i).$$

By piecing these $\lambda(i) + 1$ deductions together, we assemble a deduction from Σ^* that ends with the following lines:

$$\begin{aligned} t_1 &\doteq t'_1 \\ &\vdots \\ t_{\lambda(i)} &\doteq t'_{\lambda(i)} \\ R_i(t_1, \dots, t_{\lambda(i)}) & \end{aligned}$$

Now we extend this proof with the following lines:

$$\begin{aligned} R_i(t_1, t_2, \dots, t_{\lambda(i)}) &\rightarrow R_i(t'_1, t_2, \dots, t_{\lambda(i)}) \\ &R_i(t'_1, t_2, \dots, t_{\lambda(i)}) \\ R_i(t'_1, t_2, t_3, \dots, t_{\lambda(i)}) &\rightarrow R_i(t'_1, t'_2, t_3, \dots, t_{\lambda(i)}) \\ &R_i(t'_1, t'_2, t_3, \dots, t_{\lambda(i)}) \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ R_i(t'_1, t'_2, \dots, t'_{\lambda(i)-1}, t_{\lambda(i)}) &\rightarrow R_i(t'_1, t'_2, \dots, t'_{\lambda(i)-1}, t'_{\lambda(i)}) \\ &R_i(t'_1, t'_2, \dots, t'_{\lambda(i)-1}, t'_{\lambda(i)}). \end{aligned}$$

These lines arise from alternating application of (R_i) (p. 20) and (MP) . Altogether we obtain

$$\Sigma^* \vdash R_i(t'_1, \dots, t'_{\lambda(i)}),$$

proving (1.4.5.8).

Next, for each $j \in J$ we define a $\mu(j)$ -place function $\mathcal{F}_j : A^{\mu(j)} \rightarrow A$ by defining, for term-classes $\overline{t_1}, \dots, \overline{t_{\mu(j)}}$,

$$\mathcal{F}_j(\overline{t_1}, \dots, \overline{t_{\mu(j)}}) := \overline{f_j(t_1, \dots, t_{\mu(j)})}. \quad (1.4.5.9)$$

Here, too, we must show that this definition does not depend on the choice of representative t_v of the class $\overline{t_v}$. By application of the Rules (F_j) and (Tr) (p. 20), we obtain, following the above pattern of argument, a proof from Σ^* of

$$f_j(t_1, \dots, t_{\mu(j)}) \doteq f_j(t'_1, \dots, t'_{\mu(j)}),$$

if one assumes

$$\Sigma^* \vdash t_v \doteq t'_v \quad \text{for } 1 \leq v \leq \mu(j).$$

Finally, for every $k \in K$,

$$\text{we take the class } \overline{c_k} \text{ to be the interpretation of } c_k. \quad (1.4.5.10)$$

The content of the next theorem is that, under the above interpretations, all L' -sentences $\sigma \in \Sigma^*$, and only those, hold in the domain A . This will become conclusively clear, however, only after we have, in the next section, made the notion of satisfaction precise. We place a small technical lemma before the promised theorem.

Lemma 1.4.6. *The maximal consistent set Σ^* of sentences is deductively closed; i.e. for each $\alpha \in \text{Sent}(L')$ with $\Sigma^* \vdash \alpha$, $\alpha \in \Sigma^*$.*

Proof: In view of the maximal consistency of Σ^* , it suffices to show that $\Sigma^* \cup \{\alpha\}$ is consistent whenever $\Sigma^* \vdash \alpha$. But this is clear: namely, if $\Sigma^* \cup \{\alpha\}$ were inconsistent, then we would have, in particular,

$$\Sigma^* \cup \{\alpha\} \vdash \neg\alpha,$$

which, together with the Deduction Theorem 1.3.2, would lead to

$$\Sigma^* \vdash (\alpha \rightarrow \neg\alpha).$$

Because of the tautology

$$(\alpha \rightarrow \neg\alpha) \rightarrow \neg\alpha,$$

this would lead, finally, to $\Sigma^* \vdash \neg\alpha$, contrary to the assumption that Σ^* is consistent. \square

Theorem 1.4.7. *Suppose Σ^* is an arbitrary subset of $\text{Sent}(L')$ possessing properties (I) and (II) (p. 33). Then, for every α , β , and $\forall x \phi$ in $\text{Sent}(L')$, we have:*

- (a) $\neg\alpha \in \Sigma^*$ iff $\alpha \notin \Sigma^*$;
- (b) $(\alpha \wedge \beta) \in \Sigma^*$ iff $(\alpha \in \Sigma^* \text{ and } \beta \in \Sigma^*)$; and
- (c) $\forall x \phi \in \Sigma^*$ iff $\phi(x/t) \in \Sigma^*$ for all $t \in \text{CT}$ (1.4.5.2).

Proof: (a) (\Rightarrow) Since Σ^* is consistent (I), α and $\neg\alpha$ cannot both lie in Σ^* .

(\Leftarrow) From $\alpha \notin \Sigma^*$ we deduce immediately that $\Sigma^* \not\vdash \alpha$, by Lemma 1.4.6. From this it follows that $\Sigma^* \cup \{\neg\alpha\}$ is consistent, by Lemma 1.4.1. But since Σ^* is maximal consistent (I), $\Sigma^* \cup \{\neg\alpha\} = \Sigma^*$, whence $\neg\alpha \in \Sigma^*$.

(b) (\Rightarrow) From $(\alpha \wedge \beta) \in \Sigma^*$ it follows trivially that $\Sigma^* \vdash (\alpha \wedge \beta)$. Then $\Sigma^* \vdash \alpha$ and $\Sigma^* \vdash \beta$, by Rules $(\wedge B_1)$ and $(\wedge B_2)$. Now use Lemma 1.4.6.

(\Leftarrow) Use Rule (\wedge) (p. 19) followed by Lemma 1.4.6.

(c) (\Rightarrow) If $\forall x \varphi \in \Sigma^*$, then $\Sigma^* \vdash \forall x \varphi$. Then for any $t \in \text{CT}$, $\Sigma^* \vdash \varphi(x/t)$, by Rule $(\forall B)$ (p. 18), which applies here since every constant term t is, vacuously, free for x in φ . By (1.4.6), we get $\varphi(x/t) \in \Sigma^*$.

(\Leftarrow) Assume $\forall x \varphi \notin \Sigma^*$. Then $\neg\forall x \varphi \in \Sigma^*$, by (a). From this we would like to deduce $\exists x \neg\varphi \in \Sigma^*$. $\Sigma^* \vdash (\neg\forall x \varphi \rightarrow \exists x \neg\varphi)$, since $\neg\forall x \varphi \rightarrow \exists x \neg\varphi$ is a tautology. From this, $\Sigma^* \cup \{\neg\forall x \varphi\} \vdash \exists x \neg\varphi$ follows immediately, and thence (with Lemma 1.3.1)

$$\Sigma^* \cup \{\forall x \neg\varphi\} \vdash \forall x \varphi.$$

Since $\forall x \varphi$ is a sentence by hypothesis, so is $\forall x \neg\varphi$. Therefore we may apply the Deduction Theorem 1.3.2 to obtain

$$\Sigma^* \vdash (\forall x \neg\varphi \rightarrow \forall x \varphi).$$

From this we obtain, using (CP) (p. 18),

$$\Sigma^* \vdash (\neg\forall x \varphi \rightarrow \neg\forall x \neg\varphi),$$

and, using $\neg\forall x \varphi \in \Sigma^*$, finally $\exists x \neg\varphi \in \Sigma^*$. By hypothesis (II) there is at least one $t \in \text{CT}$ (in fact, t may even be taken to be a constant *symbol*) such that

$$\Sigma^* \vdash (\exists x \neg\varphi \rightarrow \neg\varphi(x/t)).$$

Therefore for this t we conclude $\neg\varphi(x/t) \in \Sigma^*$. But then $\varphi(x/t) \notin \Sigma^*$, by (I). \square

1.5 First-Order Semantics

In this section we want to define what it means for a formula φ of a formal language L to hold in, or to be satisfied by, a particular mathematical structure, and, more generally, what it should mean for such a structure to be a model of an axiom system Σ . In order to be able to carry this out meaningfully, we must first fix the boundaries of a domain of objects to which our quantifiers $\forall u$ and $\exists v$ should refer, i.e. over which the variables v_0, v_1, \dots should “vary”. After that we must define the interpretation of each relation symbol R_i ($i \in I$), each function symbol f_j ($j \in J$), and each constant symbol c_k ($k \in K$).

Suppose we are given a formal language $L = (\lambda, \mu, K)$. An L -structure \mathfrak{A} is determined by the following data:

- $|\mathfrak{A}|$: a nonempty set, the *universe* of \mathfrak{A} ;
- $R_i^{\mathfrak{A}}$: a $\lambda(i)$ -place relation on $|\mathfrak{A}|$ (i.e. a subset of $|\mathfrak{A}|^{\lambda(i)}$), for each $i \in I$;
- $f_j^{\mathfrak{A}}$: a $\mu(j)$ -place function defined on all of $|\mathfrak{A}|$
(i.e. a function $|\mathfrak{A}|^{\mu(j)} \rightarrow |\mathfrak{A}|$), for each $j \in J$;
- $c_k^{\mathfrak{A}}$: a fixed element of $|\mathfrak{A}|$, for each $k \in K$.

We summarize this with the notation

$$\mathfrak{A} = \langle |\mathfrak{A}|; (R_i^{\mathfrak{A}})_{i \in I}; (f_j^{\mathfrak{A}})_{j \in J}; (c_k^{\mathfrak{A}})_{k \in K} \rangle.$$

If we again consider the language L (p. 21) used in our example of a formal proof in Section 1.3, with the relation symbol \leq , the function symbols $-$, $+$, \cdot , and the constant symbol 0 , then the following is an L -structure:

$$\mathfrak{R} = \langle \mathbb{R}; \leq^{\mathbb{R}}, -^{\mathbb{R}}, +^{\mathbb{R}}, \cdot^{\mathbb{R}}; 0^{\mathbb{R}} \rangle. \quad (1.5.0.1)$$

Here \mathbb{R} is the set of real numbers; $\leq^{\mathbb{R}}$ is the usual “less-than-or-equal-to” relation on \mathbb{R} ; $-^{\mathbb{R}}$, $+^{\mathbb{R}}$, and $\cdot^{\mathbb{R}}$ are the usual operations “negative” (unary), “plus” (binary), and “times” (binary) on \mathbb{R} ; and $0^{\mathbb{R}}$ is the real number “zero”.

Let us pursue this example further by considering the formula

$$\exists v_0 (0 \leq v_0 \wedge v_0 \leq v_1). \quad (1.5.0.2)$$

Then the question whether this formula holds in (or is satisfied by) \mathfrak{R} – presupposing for now a definition of satisfaction that agrees with our intuition – can be meaningfully answered only after we assign to v_1 a definite value in \mathbb{R} : for a negative value of v_1 , the answer is no; for other values, the answer is yes. Thus we see in this example that for a meaningful definition of the satisfaction of a formula φ in \mathfrak{R} , each of the free variables of φ must be assigned a value in \mathbb{R} . For certain technical reasons we assign values not only to the free variables of one formula, but to the free variables of all formulae; i.e. to all variables. However, in the definition of satisfaction we must make sure that in the case of the bound variables of the formula φ under consideration, the fixed assignment by h is “unfixed”.

An assignment of values in $|\mathfrak{A}|$ for all variables will be called an *evaluation of the variables in \mathfrak{A}* . Thus, an evaluation in \mathfrak{A} is a function

$$h : \text{Vbl} \rightarrow |\mathfrak{A}|.$$

If h is an evaluation in \mathfrak{A} , then the value $h(x) \in |\mathfrak{A}|$ is assigned to the variable x . For each $a \in |\mathfrak{A}|$, each $x \in \text{Vbl}$, and each evaluation h , the function $h\binom{x}{a}$, defined as follows, is again an evaluation:

$$h_a^{(x)}(v) = \begin{cases} h(v) & \text{for } v \neq x \\ a & \text{for } v = x. \end{cases} \quad (1.5.0.3)$$

The evaluations h and $h_a^{(x)}$ agree with each other at all variables other than x . At x , the value of h is $h(x)$, while that of $h_a^{(x)}$ is a . Obviously, $h_{h(x)}^{(x)} = h$.

Next we define, by recursion on the construction of a term, the *value* $t^{\mathfrak{A}}[h]$ of the term t under the evaluation h (or simply the h -value of t) in \mathfrak{A} :

$$\begin{aligned} v^{\mathfrak{A}}[h] &:= h(v) \\ c_k^{\mathfrak{A}}[h] &:= c_k^{\mathfrak{A}} \\ f_j(t_1, \dots, t_{\mu(j)})^{\mathfrak{A}}[h] &:= f_j^{\mathfrak{A}}(t_1^{\mathfrak{A}}[h], \dots, t_{\mu(j)}^{\mathfrak{A}}[h]). \end{aligned} \quad (1.5.0.4)$$

It is clear that these equations determine the h -value (in $|\mathfrak{A}|$) of each term, by starting the recursive process with the h -values of its simplest subterms – the variables and constant symbols occurring in it.

The *satisfaction of a formula φ under an evaluation h in \mathfrak{A}* will be a ternary relation of our metatheory. If this relation between \mathfrak{A} , φ , and h holds, then we shall write $\mathfrak{A} \models \varphi[h]$ (pronounced: “ φ holds in \mathfrak{A} under h ”, “ φ is true in \mathfrak{A} under h ”, or “ φ is satisfied by \mathfrak{A} under h ”); if this relation does not hold, then we shall write $\mathfrak{A} \not\models \varphi[h]$. This relation will likewise be defined by recursion on the construction of formulae, starting with the simplest formulae, the atomic formulae, and indeed simultaneously for all evaluations.

For atomic formulae $t_1 \doteq t_2$ and $R_i(t_1, \dots, t_{\lambda(i)})$, we declare, for an arbitrary evaluation h in \mathfrak{A} :

$$\mathfrak{A} \models t_1 \doteq t_2 [h] \quad \text{iff} \quad t_1^{\mathfrak{A}}[h] = t_2^{\mathfrak{A}}[h]; \quad (1.5.0.5)$$

$$\mathfrak{A} \models R_i(t_1, \dots, t_{\lambda(i)}) [h] \quad \text{iff} \quad R_i^{\mathfrak{A}}(t_1^{\mathfrak{A}}[h], \dots, t_{\lambda(i)}^{\mathfrak{A}}[h]). \quad (1.5.0.6)$$

Thereafter, for formulae φ and ψ we continue our recursive definition as follows:

$$\mathfrak{A} \models \neg\varphi [h] \quad \text{iff} \quad \mathfrak{A} \not\models \varphi [h]; \quad (1.5.0.7)$$

$$\mathfrak{A} \models (\varphi \wedge \psi) [h] \quad \text{iff} \quad (\mathfrak{A} \models \varphi [h] \text{ and } \mathfrak{A} \models \psi [h]); \quad (1.5.0.8)$$

$$\mathfrak{A} \models \forall x \varphi [h] \quad \text{iff} \quad \mathfrak{A} \models \varphi [h_a^{(x)}], \text{ for all } a \in |\mathfrak{A}|. \quad (1.5.0.9)$$

Observe that in the last case, where x is certainly bound, the prescription by h of a particular value in $|\mathfrak{A}|$ for x is unfixed, since in this case we consider, instead of h itself, an alteration of h at the point x ; here every such alteration is taken into consideration. In this way we ensure that the definition of satisfaction really agrees with our intuition.

Using the definitions of \forall , \rightarrow , \leftrightarrow , and \exists from Section 1.2, and the definition of satisfaction, one obtains immediately the following equivalences:

$$\begin{aligned} \mathfrak{A} \models (\varphi \vee \psi) [h] & \quad \text{iff } (\mathfrak{A} \models \varphi [h] \quad \text{or} \quad \mathfrak{A} \models \psi [h]); \\ \mathfrak{A} \models (\varphi \rightarrow \psi) [h] & \quad \text{iff } (\mathfrak{A} \models \varphi [h] \text{ implies } \mathfrak{A} \models \psi [h]); \end{aligned} \quad (1.5.0.10)$$

$$\begin{aligned} \mathfrak{A} \models (\varphi \leftrightarrow \psi) [h] & \quad \text{iff } (\mathfrak{A} \models \varphi [h] \quad \text{iff} \quad \mathfrak{A} \models \psi [h]); \\ \mathfrak{A} \models \exists x \varphi [h] & \quad \text{iff there is an } a \in |\mathfrak{A}| \text{ such that } \mathfrak{A} \models \varphi [h(x_a^x)]. \end{aligned} \quad (1.5.0.11)$$

Here, the words “and”, “or”, “implies”, “iff”, “for all a ” and “there exists an a ” are to be understood in the usual mathematical sense; in particular, “or” is not used in the exclusive sense, and “implies” is regarded as false only when the premise is true and the conclusion is false; cf. the truth table on page 16.

For the structure \mathfrak{R} in (1.5.0.1) above (with the set of real numbers as universe), and the formula $\exists v_0 (0 \leq v_0 \wedge v_0 \leq v_1)$ (1.5.0.2), we have the following translation:

$$\begin{aligned} \mathfrak{R} \models \exists v_0 (0 \leq v_0 \wedge v_0 \leq v_1) [h] \\ \text{iff there exists an } a \in \mathbb{R} \text{ such that } \mathfrak{R} \models (0 \leq v_0 \wedge v_0 \leq v_1) [h(v_0^a)] \\ \text{iff there exists an } a \in \mathbb{R} \text{ such that } (0 \leq^{\mathfrak{R}} a \wedge a \leq^{\mathfrak{R}} h(v_1)). \end{aligned}$$

This example shows yet again that the definition of satisfaction has fulfilled its purpose: it translates, using the prescription of values of the free variables given by an evaluation h , a string φ of symbols into an assertion in the metalanguage – the interpretation of φ in \mathfrak{R} under h .

In this example we see, in addition, that for the relation \models to hold, the only variables whose h -values are material are the free variables of φ . This can be shown in general:

Lemma 1.5.1. *Let h' and h'' be evaluations in the L -structure \mathfrak{A} . Then:*

- (a) *If h' and h'' agree on the variables in the L -term t , then $t^{\mathfrak{A}}[h'] = t^{\mathfrak{A}}[h'']$.*
- (b) *If h' and h'' agree on the free variables occurring in the L -formula φ , then*

$$\mathfrak{A} \models \varphi [h'] \quad \text{iff} \quad \mathfrak{A} \models \varphi [h'']. \quad (1.5.1.1)$$

Proof: (a) The following equations prove this by induction on the recursive construction of terms:

$$\begin{aligned} v^{\mathfrak{A}}[h'] &= h'(v) = h''(v) = v^{\mathfrak{A}}[h''] \\ c_k^{\mathfrak{A}}[h'] &= c_k^{\mathfrak{A}} = c_k^{\mathfrak{A}}[h''] \\ f_j(t_1, \dots, t_{\mu(j)})^{\mathfrak{A}}[h'] &= f_j^{\mathfrak{A}}(t_1^{\mathfrak{A}}[h'], \dots, t_{\mu(j)}^{\mathfrak{A}}[h']) \\ &= f_j^{\mathfrak{A}}(t_1^{\mathfrak{A}}[h''], \dots, t_{\mu(j)}^{\mathfrak{A}}[h'']) \\ &= f_j(t_1, \dots, t_{\mu(j)})^{\mathfrak{A}}[h'']. \end{aligned}$$

(b) We prove (1.5.1.1) similarly, using (a) and induction on the recursive construction of formulae:

$$\begin{aligned}
\mathfrak{A} \models t_1 \doteq t_2 [h'] & \text{ iff } t_1^{\mathfrak{A}}[h'] = t_2^{\mathfrak{A}}[h'] \\
& \text{ iff } t_1^{\mathfrak{A}}[h''] = t_2^{\mathfrak{A}}[h''] \\
& \text{ iff } \mathfrak{A} \models t_1 \doteq t_2 [h'']. \\
\mathfrak{A} \models R_i(t_1, \dots) [h'] & \text{ iff } R_i^{\mathfrak{A}}(t_1^{\mathfrak{A}}[h'], \dots) \\
& \text{ iff } R_i^{\mathfrak{A}}(t_1^{\mathfrak{A}}[h''], \dots) \\
& \text{ iff } \mathfrak{A} \models R_i(t_1, \dots) [h'']. \\
\mathfrak{A} \models \neg \varphi [h'] & \text{ iff } \mathfrak{A} \not\models \varphi [h'] \\
& \text{ iff } \mathfrak{A} \not\models \varphi [h''] \quad (\text{ind. hyp.}) \\
& \text{ iff } \mathfrak{A} \models \neg \varphi [h'']. \\
\mathfrak{A} \models (\varphi \wedge \psi) [h'] & \text{ iff } (\mathfrak{A} \models \varphi [h'] \text{ and } \mathfrak{A} \models \psi [h']) \\
& \text{ iff } (\mathfrak{A} \models \varphi [h''] \text{ and } \mathfrak{A} \models \psi [h'']) \quad (\text{ind. hyp.}) \\
& \text{ iff } \mathfrak{A} \models (\varphi \wedge \psi) [h'']. \\
\mathfrak{A} \models \forall x \varphi [h'] & \text{ iff } \mathfrak{A} \models \varphi [h'(\frac{x}{a})] \text{ for all } a \in |\mathfrak{A}| \\
& \text{ iff } \mathfrak{A} \models \varphi [h''(\frac{x}{a})] \text{ for all } a \in |\mathfrak{A}| \quad (1.5.1.2) \\
& \text{ iff } \mathfrak{A} \models \forall x \varphi [h''].
\end{aligned}$$

In (1.5.1.2) we applied the inductive hypothesis to the shorter formula φ and the evaluations $h'(\frac{x}{a})$ and $h''(\frac{x}{a})$. Note that these two evaluations agree with each other on all free variables of φ , due to the common alteration of h' and h'' at x . \square

From Lemma 1.5.1 we see, in particular, that the satisfaction of a sentence φ in \mathfrak{A} does not depend on the evaluation h considered. That is, for $\varphi \in \text{Sent}(L)$ and evaluations h' and h'' in \mathfrak{A} , we always have

$$\mathfrak{A} \models \varphi [h'] \quad \text{iff} \quad \mathfrak{A} \models \varphi [h'']. \quad (1.5.1.3)$$

We say that a formula φ *holds in*, or *is true in*, or *is satisfied by*, \mathfrak{A} (and we write $\mathfrak{A} \models \varphi$), in case $\mathfrak{A} \models \varphi [h]$ holds for all evaluations h in \mathfrak{A} . One easily sees that

$$\mathfrak{A} \models \varphi \quad \text{iff} \quad \mathfrak{A} \models \forall \varphi \quad (\text{p. 13}).$$

This follows, by induction, from the equivalence

$$\mathfrak{A} \models \varphi \quad \text{iff} \quad \mathfrak{A} \models \forall x \varphi \quad (1.5.1.4)$$

(for any variable x), which is proved as follows:

$$\begin{aligned}
\mathfrak{A} \models \varphi & \text{ iff } \mathfrak{A} \models \varphi[h] && \text{for every evaluation } h \\
& \text{ iff } \mathfrak{A} \models \varphi[h_a^x] && \text{for every evaluation } h \text{ and every } a \in |\mathfrak{A}| \\
& \text{ iff } \mathfrak{A} \models \forall x \varphi[h] && \text{for every evaluation } h \\
& \text{ iff } \mathfrak{A} \models \forall x \varphi.
\end{aligned}$$

If Σ is a set of L -sentences, then an L -structure \mathfrak{A} is called a *model* of Σ if every $\sigma \in \Sigma$ holds in \mathfrak{A} , i.e. $\mathfrak{A} \models \sigma$ for all $\sigma \in \Sigma$. In this case we write $\mathfrak{A} \models \Sigma$.

Now we come back to the starting point (1.4.0.1) of our inquiry in Section 1.4: How can an “undeducibility” $\Sigma \not\vdash \varphi$ occur? The answer is given by the following theorem (which is the promised, precise version of (1.4.0.2)):

Theorem 1.5.2 (Gödel’s completeness theorem). *Let $\Sigma \subseteq \text{Sent}(L)$ and $\varphi \in \text{Sent}(L)$. Then from $\Sigma \not\vdash \varphi$ follows the existence of a “counterexample”, i.e. an L -structure \mathfrak{A} that is a model of Σ but in which φ does not hold (thus \mathfrak{A} is a model of $\Sigma \cup \{\neg\varphi\}$).*

Proof: By Lemma 1.4.1, the condition $\Sigma \not\vdash \varphi$ is equivalent to the condition that the set $\Sigma_1 := \Sigma \cup \{\neg\varphi\}$ is consistent. Thus we must show that every consistent set $\Sigma \subseteq \text{Sent}(L)$ possesses a model. (Here we have replaced $\Sigma \cup \{\neg\varphi\}$ by Σ .)

So let $\Sigma \subseteq \text{Sent}(L)$ be consistent. We apply first Theorem 1.4.2 and then Theorem 1.4.4 to this Σ . We obtain thereby a $\Sigma^* \subseteq \text{Sent}(L')$ with properties (I) and (II) on page 33. Here, L' is the extension language of L constructed in Theorem 1.4.2. Let

$$A = \text{CT}/\approx \tag{1.5.2.1}$$

be the set of equivalence classes of constant terms of the language L' , constructed in (1.4.5.4). Further, let \mathcal{R}_i (for $i \in I$), \mathcal{F}_j (for $j \in J$) and $\overline{c_k}$ (for $k \in K'$) be the interpretations of the relation symbols R_i , the function symbols f_j and the constant symbols c_k , given in (1.4.5.7), (1.4.5.9) and (1.4.5.10), respectively. Then

$$\mathfrak{A} = \langle A; (\mathcal{R}_i)_{i \in I}; (\mathcal{F}_j)_{j \in J}; (c_k)_{k \in K'} \rangle \tag{1.5.2.2}$$

is, finally, an L' -structure. By definition,

$$\mathcal{R}_i = R_i^{\mathfrak{A}}, \quad \mathcal{F}_j = f_j^{\mathfrak{A}}, \quad \overline{c_k} = c_k^{\mathfrak{A}}. \tag{1.5.2.3}$$

From our special definition of functions, we even have

$$t^{\mathfrak{A}}[h] = \overline{t} \tag{1.5.2.4}$$

for all $t \in \text{CT}$ and all evaluations h in \mathfrak{A} . Indeed, by induction on the construction of a constant term we have

$$\begin{aligned}
f_j(t_1, \dots, t_{\mu(j)})^{\mathfrak{A}}[h] &= f_j^{\mathfrak{A}}(t_1^{\mathfrak{A}}[h], \dots, t_{\mu(j)}^{\mathfrak{A}}[h]) && \text{by (1.5.0.4)} \\
&= \mathcal{F}_j(\overline{t_1}, \dots, \overline{t_{\mu(j)}}) && \text{by (1.5.2.3) and ind. hyp.} \\
&= \overline{f_j(t_1, \dots, t_{\mu(j)})} && \text{by (1.4.5.9).}
\end{aligned}$$

Now we show that for every $\varphi \in \text{Sent}(L')$ and every evaluation h in \mathfrak{A} ,

$$\mathfrak{A} \models \varphi[h] \quad \text{iff} \quad \varphi \in \Sigma^*. \quad (1.5.2.5)$$

Then, in particular, \mathfrak{A} will be a model of Σ^* . We shall prove (1.5.2.5) by induction on the construction of φ – more precisely, by induction on the number of logical symbols \neg, \wedge and \forall used in the construction of φ .

If this number is 0, then we are dealing with an atomic formula. But in this case the definitions furnish us with the required equivalence, since for constant terms t_1, t_2, \dots we have, first,

$$\begin{aligned} \mathfrak{A} \models t_1 \doteq t_2 [h] & \quad \text{iff} \quad t_1^{\mathfrak{A}}[h] = t_2^{\mathfrak{A}}[h] && \text{by (1.5.0.5)} \\ & \quad \text{iff} \quad \bar{t}_1 = \bar{t}_2 && \text{by (1.5.2.4)} \\ & \quad \text{iff} \quad t_1 \doteq t_2 \in \Sigma^* && \text{by (1.4.5.6), (1.4.5.3),} \\ & && \text{and (1.4.6),} \end{aligned}$$

and, second,

$$\begin{aligned} \mathfrak{A} \models R_i(t_1, \dots) [h] & \quad \text{iff} \quad R_i^{\mathfrak{A}}(t_1^{\mathfrak{A}}[h], \dots) && \text{by (1.5.0.6)} \\ & \quad \text{iff} \quad \mathcal{R}_i(\bar{t}_1, \dots) && \text{by (1.5.2.3) and (1.5.2.4)} \\ & \quad \text{iff} \quad R_i(t_1, \dots) \in \Sigma^* && \text{by (1.4.5.7) and (1.4.6).} \end{aligned}$$

Next, if the sentence φ is of the form $\neg\alpha$ or $(\alpha \wedge \beta)$, then α and β are likewise sentences. We have:

$$\begin{aligned} \mathfrak{A} \models \neg\alpha [h] & \quad \text{iff} \quad \mathfrak{A} \not\models \alpha [h] && \text{by (1.5.0.7)} \\ & \quad \text{iff} \quad \alpha \notin \Sigma^* && \text{by ind. hyp. (1.5.2.5)} \\ & \quad \text{iff} \quad \neg\alpha \in \Sigma^* && \text{by (1.4.7)(a), and} \\ \mathfrak{A} \models (\alpha \wedge \beta) [h] & \quad \text{iff} \quad (\mathfrak{A} \models \alpha [h] \text{ and } \mathfrak{A} \models \beta [h]) && \text{by (1.5.0.8)} \\ & \quad \text{iff} \quad (\alpha \in \Sigma^* \text{ and } \beta \in \Sigma^*) && \text{by ind. hyp. (1.5.2.5)} \\ & \quad \text{iff} \quad (\alpha \wedge \beta) \in \Sigma^* && \text{by (1.4.7)(b).} \end{aligned}$$

Finally, if the sentence φ is of the form $\forall x \psi$, then for $t \in \text{CT}$, $\psi(x/t)$ is obviously again a sentence, since $\text{Fr}(\psi) \subseteq \{x\}$. And $\psi(x/t)$ is more simply built than φ , as far as the number of logical symbols. Therefore, using Lemma 1.5.3, proved just below, we have:

$$\begin{aligned} \mathfrak{A} \models \forall x \psi [h] & \quad \text{iff} \quad \mathfrak{A} \models \psi[h(\frac{x}{a})] \quad \text{for all } a \in A && \text{by (1.5.0.9)} \\ & \quad \text{iff} \quad \mathfrak{A} \models \psi[h(\frac{x}{t})] \quad \text{for all } t \in \text{CT} && \text{by (1.5.2.1) \& (1.4.5.5)} \\ & \quad \text{iff} \quad \mathfrak{A} \models \psi(x/t) [h] \quad \text{for all } t \in \text{CT} && \text{by (1.5.3) for } L' \quad (1.5.2.6) \\ & \quad \text{iff} \quad \psi(x/t) \in \Sigma^* \quad \text{for all } t \in \text{CT} && \text{by ind. hyp. (1.5.2.5)} \\ & \quad \text{iff} \quad \forall x \psi \in \Sigma^* && \text{by (1.4.7)(c).} \end{aligned}$$

We have proved, finally (modulo Lemma 1.5.3 below), that all L' -sentences $\varphi \in \Sigma^*$ (and only those) hold in the L' -structure \mathfrak{A} (1.5.2.2). It is therefore clear that all L -sentences $\varphi \in \Sigma$ hold in the L -structure

$$\mathfrak{A}|_L := \langle A; (\mathcal{R}_i)_{i \in I}; (\mathcal{F}_j)_{j \in J}; (c_k)_{k \in K} \rangle.$$

Thus, Σ possesses a model. \square

The following lemma (used in (1.5.2.6) above) is of a technical nature.

Lemma 1.5.3. *Let \mathfrak{A} be L -structure, h an evaluation in \mathfrak{A} , φ an L -formula, x a variable and t an L -term that is free for x in φ . Then, writing $a = t^{\mathfrak{A}}[h]$ (1.5.0.4), we have:*

$$\mathfrak{A} \models \varphi [h(x_a^x)] \quad \text{iff} \quad \mathfrak{A} \models \varphi(x/t) [h]. \quad (1.5.3.1)$$

Proof: For terms t_1 one shows easily, by induction on their construction, that

$$t_1^{\mathfrak{A}} [h(x_a^x)] = t_1(x/t)^{\mathfrak{A}} [h]. \quad (1.5.3.2)$$

Now we shall prove (1.5.3.1) by induction on the construction of the formula φ . In the atomic formula $t_1 \doteq t_2$, for example, we have:

$$\begin{aligned} \mathfrak{A} \models t_1 \doteq t_2 [h(x_a^x)] & \quad \text{iff} \quad t_1^{\mathfrak{A}} [h(x_a^x)] = t_2^{\mathfrak{A}} [h(x_a^x)] \quad \text{by (1.5.0.5)} \\ & \quad \text{iff} \quad t_1(x/t)^{\mathfrak{A}} [h] = t_2(x/t)^{\mathfrak{A}} [h] \quad \text{by (1.5.3.2)} \\ & \quad \text{iff} \quad \mathfrak{A} \models (t_1 \doteq t_2)(x/t) [h] \quad \text{by (1.5.0.5)}. \end{aligned} \quad (1.5.3.3)$$

In (1.5.3.3) we have enclosed $t_1 \doteq t_2$ in parentheses in order to indicate the scope of application of the syntactic operation (x/t) . (Note that the notation (x/t) (1.2.0.8) belongs not to the object language, but to the metalanguage!) It is clear that $(t_1 \doteq t_2)(x/t)$ is $t_1(x/t) \doteq t_2(x/t)$.

The other case (1.5.0.6) of atomic formulae φ is handled analogously.

The cases in which φ is of the form $\neg\varphi$ or $(\alpha \wedge \beta)$ are likewise easily handled, by referring back to the components α , or α and β , respectively.

There remains only the case where φ is of the form $\forall y \psi$. Here we distinguish two (sub)cases:

Case 1: Either y is x , or $x \notin \text{Fr}(\psi)$. Under either of these assumptions, $x \notin \text{Fr}(\forall y \psi)$. Therefore, using Lemma 1.5.1,

$$\mathfrak{A} \models \forall y \psi [h(x_a^x)] \quad \text{iff} \quad \mathfrak{A} \models \forall y \psi [h].$$

This, however, is (1.5.3.1), since $(\forall y \psi)(x/t)$ is clearly $\forall y \psi$.

Case 2: y is not x , and $x \in \text{Fr}(\psi)$. In this case y cannot occur in t , since, by hypothesis, t is free for x in $\forall y \psi$ (p. 12). It follows, using Lemma 1.5.1(a), that for every $a' \in |\mathfrak{A}|$,

$$a = t^{\mathfrak{A}} [h] = t^{\mathfrak{A}} [h(x_{a'}^y)]. \quad (1.5.3.4)$$

Writing $h' = h(x_{a'}^y)$, we then have

$$\begin{aligned}
\mathfrak{A} \models \forall y \psi [h_a^x] &\text{ iff } \mathfrak{A} \models \psi [h_a^x (a')] && \text{for all } a' \in |\mathfrak{A}| && \text{by (1.5.0.9)} \\
&\text{ iff } \mathfrak{A} \models \psi [h_a^x] && \text{for all } a' \in |\mathfrak{A}| && \text{since } y \text{ is not } x \\
&\text{ iff } \mathfrak{A} \models \psi [h_a^x (a')] && \text{for all } a' \in |\mathfrak{A}| && \text{by (1.5.3.4)} \\
&\text{ iff } \mathfrak{A} \models \psi (x/t) [h] && \text{for all } a' \in |\mathfrak{A}| && \text{ind. hyp. (1.5.3.1) for } h' \\
&\text{ iff } \mathfrak{A} \models \psi (x/t) [h_a^y] && \text{for all } a' \in |\mathfrak{A}| && \text{def. of } h' \\
&\text{ iff } \mathfrak{A} \models \forall y \psi (x/t) [h] && && \text{by (1.5.0.9)} \\
&\text{ iff } \mathfrak{A} \models (\forall y \psi) (x/t) [h] && && \text{since } y \text{ is not } x. \quad \square
\end{aligned}$$

The next theorem assures us of the *correctness* or *soundness* of the concept of proof developed in Section 1.3. It asserts that everything that can be proved from an axiom system Σ also holds in every model of Σ . Thus, if we have a model of $\Sigma \cup \{\neg\varphi\}$, then obviously φ cannot be proved from Σ . Considered formally, this is the converse of the implication asserted in Gödel's Completeness Theorem (1.5.2).

Theorem 1.5.4 (Soundness Theorem). *Suppose $\Sigma \subseteq \text{Sent}(L)$, $\varphi \in \text{Sent}(L)$ and $\Sigma \vdash \varphi$. Then φ holds in every model of Σ .*

Proof: Let \mathfrak{A} be a model of Σ , and $\varphi_1, \dots, \varphi_n$ be a proof from Σ . We shall show that $\mathfrak{A} \models \varphi_i$ for $i \in \{1, 2, \dots, n\}$, by induction on i . This will, in particular, prove the theorem.

Basis step of the induction: Let $i = 1$. Then either $\varphi_i \in \Sigma$ (1.3.0.1), or φ_i is a logical axiom (1.3.0.2). In the first case, φ_i holds in \mathfrak{A} by hypothesis.

In the second case, suppose, first, that φ_i is an instance of a tautological form Φ in the sentential variables A_0, \dots, A_m , arising by the replacement of A_j by the L -formula ψ_j , for $0 \leq j \leq m$ (p. 16). Let h be any evaluation in \mathfrak{A} . Define a truth assignment \mathcal{H} on the variables A_0, \dots, A_m as follows:

$$\mathcal{H}(A_j) = \begin{cases} T & \text{if } \mathfrak{A} \models \psi_j [h], \\ F & \text{if } \mathfrak{A} \not\models \psi_j [h]. \end{cases}$$

Then $\mathcal{H}(\Phi) = T$ iff $\mathfrak{A} \models \varphi_i [h]$, since the equations (1.3.0.4) and (1.3.0.5) giving the recursive definition of \mathcal{H} on arbitrary sentential forms Φ agree, respectively, with the equations (1.5.0.7) and (1.5.0.8) giving (the relevant part of) the recursive definition of $\mathfrak{A} \models \varphi_i [h]$. But $\mathcal{H}(\Phi) = T$, since Φ is tautological; therefore $\mathfrak{A} \models \varphi_i [h]$. Since h was arbitrary, we conclude that $\mathfrak{A} \models \varphi_i$ (p. 40).

In the case where φ_i is one of the equality axioms (1.3.0.9), one may convince oneself just as easily that $\mathfrak{A} \models \varphi_i$.

There remain the cases where φ_i is an instance of the quantifier axioms (A1) (1.3.0.7) or (A2) (1.3.0.8).

First let φ_i be of the form

$$\forall x (\alpha \rightarrow \beta) \rightarrow (\alpha \rightarrow \forall x \beta),$$

where x is not free in α . According to the definition of satisfaction (p. 40), we must show, for every evaluation h , that:

the hypothesis

$$\text{for all } a \in |\mathfrak{A}|, \mathfrak{A} \models (\alpha \rightarrow \beta) [h_a^x] \quad (1.5.4.1)$$

implies

$$\text{if } \mathfrak{A} \models \alpha [h], \text{ then for all } a \in |\mathfrak{A}|, \mathfrak{A} \models \beta [h_a^x]. \quad (1.5.4.2)$$

So assume (1.5.4.1), and suppose $\mathfrak{A} \models \alpha [h]$ and $a \in |\mathfrak{A}|$. Then $\mathfrak{A} \models \alpha [h_a^x]$ by Lemma 1.5.1(b), since $x \notin \text{Fr}(\alpha)$. (1.5.4.1) and (1.5.0.10) then give $\mathfrak{A} \models \beta [h_a^x]$, proving (1.5.4.2), as required in this “(A2)” case.

Second, let φ_i be of the form

$$\forall x \alpha \rightarrow \alpha(x/t),$$

where t is free for x in α . For each evaluation h we must show that: the hypothesis

$$\mathfrak{A} \models \alpha [h_a^x] \text{ for all } a \in |\mathfrak{A}| \quad (1.5.4.3)$$

implies

$$\mathfrak{A} \models \alpha(x/t) [h]. \quad (1.5.4.4)$$

By Lemma 1.5.3, (1.5.4.4) is equivalent to $\mathfrak{A} \models \alpha [h_a^x]$ with $a = t^{\mathfrak{A}} [h]$. But this is a special case of (1.5.4.3).

Induction step: Assume, for all $j < i$, that we have already proved $\mathfrak{A} \models \varphi_j$. We must show $\mathfrak{A} \models \varphi_i$.

If φ_i is a member of Σ or a logical axiom, then we obtain $\mathfrak{A} \models \varphi_i$ as in the basis step of the induction, above. If φ_i comes about by means of (MP), then there are $j, k < i$ with φ_k being $\varphi_j \rightarrow \varphi_i$. For each evaluation h we then have $\mathfrak{A} \models (\varphi_j \rightarrow \varphi_i) [h]$ and $\mathfrak{A} \models \varphi_j [h]$. This immediately gives $\mathfrak{A} \models \varphi_i [h]$, using (1.5.0.10). Since h was arbitrary, $\mathfrak{A} \models \varphi_i$.

Finally, if φ_i comes about by means of (\forall) , then there exists a $j < i$ and a variable x such that φ_i is $\forall x \varphi_j$. The inductive hypothesis is that $\mathfrak{A} \models \varphi_j$. From this follows $\mathfrak{A} \models \forall x \varphi_j$ (1.5.1.4). \square

Corollary 1.5.5. *A set $\Sigma \subseteq \text{Sent}(L)$ is consistent if and only if it possesses a model.*

Proof: If Σ is consistent, then Σ possesses a model by Gödel’s Completeness Theorem 1.5.2. Conversely, if Σ possesses a model, then no contradiction can be deduced from Σ , according to Theorem 1.5.4. \square

From this corollary, which actually summarizes Theorems 1.5.2 and 1.5.4, we obtain the most important theorem of model theory (later named the Compactness Theorem):

Theorem 1.5.6 (Finiteness Theorem). *A set $\Sigma \subseteq \text{Sent}(L)$ possesses a model if and only if every finite subset Π of Σ possesses a model.*

Proof: If Σ possesses a model, then every finite subset of Σ possesses the same model. It remains to show the converse. Assume that Σ possesses no model. Then Σ would be inconsistent, by (1.5.5). However, since only finitely many elements of Σ can occur in any proof of any contradiction from Σ , some finite subset $\Pi \subseteq \Sigma$ would already be inconsistent. By (1.5.5) again, Π would have no model. This proves the converse. \square

Remark 1.5.7 (Model theoretic proofs of the Finiteness Theorem). The above proof of the Finiteness Theorem made essential use of Gödel's Completeness Theorem, and hence of the concept of proof. In the above argument we see immediately how the finiteness of the concept of proof has an impact. Observe, however, that the *statement* of the Finiteness Theorem makes no reference to the concept of proof. In fact, there are other proofs of this theorem, which are purely model theoretic. They can be expressed using only the concepts of “formal language” and “model.” In Section 2.6 we shall carry out such a proof, by means of ultraproducts. For a deeper understanding of the Finiteness Theorem, however, the proof using Gödel's Completeness Theorem seems to us to be better.

1.6 Axiomatization of Several Mathematical Theories

In this section we wish to axiomatize several mathematical theories in the framework of the formal languages introduced here. In particular, we shall give an axiom system (in a suitable language) for set theory. First, however, we shall clarify what we mean by a mathematical theory.

The word “theory” in mathematics has many uses, and cannot easily be defined comprehensively. Consider “number theory”, as an example: what can be said with certainty is that in this “theory” one investigates the set \mathbb{N} of natural numbers (or also the set \mathbb{Z} of integers) and the properties of the operations “addition” and “multiplication” defined on those numbers. Usually one also associates with a certain mathematical “theory”, implicitly or explicitly, the methods that are applied in it; for example, one speaks of “analytic number theory” or “algebraic number theory” or even “modern number theory”. In the case of the attribute “modern”, other or new methods are not necessarily meant; often it is just a matter of a representation in a new, “modern” language. All these methods of investigation have one goal in common:

One wishes to produce, as far as possible, all sentences true in \mathbb{N} (or in \mathbb{Z}).

Analogously, in “group theory” one wants to produce, as far as possible, all sentences true in a specified class of groups. We take this goal as the motivation for the following definition.

Let $L = (\lambda, \mu, K)$ be a formal language and \mathcal{M} a nonempty class of L -structures. Then we define *the L -theory of \mathcal{M}* to be the set

$$\text{Th}(\mathcal{M}) := \{ \alpha \in \text{Sent}(L) \mid \mathfrak{A} \models \alpha, \text{ for all } \mathfrak{A} \in \mathcal{M} \}. \quad (1.6.0.1)$$

Thus, $\text{Th}(\mathcal{M})$ consists exactly of all L -sentences that hold in all structures in \mathcal{M} . The set $\text{Th}(\mathcal{M})$ of sentences possesses two important properties:

- (1) $\text{Th}(\mathcal{M})$ is consistent, and
- (2) $\text{Th}(\mathcal{M})$ is deductively closed.

Here, a subset $\Sigma \subseteq \text{Sent}(L)$ is called *deductively closed* (cf. Lemma 1.4.6) if, for every L -sentence α , $\Sigma \vdash \alpha$ implies $\alpha \in \Sigma$. Property (1) follows from Corollary 1.5.5 and the nonemptiness of \mathcal{M} . Property (2) is derived as follows: Suppose $\text{Th}(\mathcal{M}) \vdash \alpha$. Since every $\mathfrak{A} \in \mathcal{M}$ is obviously a model of $\text{Th}(\mathcal{M})$, \mathfrak{A} is also a model of α , by Theorem 1.5.4. Therefore α belongs to $\text{Th}(\mathcal{M})$, by (1.6.0.1).

More generally, we call a subset T of $\text{Sent}(L)$ an *L -theory* if T is consistent and deductively closed. Every L -theory in this sense is actually *the L -theory* of a class of L -structures, namely, just the class \mathcal{M} of all models of T , i.e. the class

$$\mathcal{M} = \{ \mathfrak{A} \mid \mathfrak{A} \text{ is an } L\text{-structure and } \mathfrak{A} \models T \}.$$

This is easy to see: by (1.6.0.1), $T \subseteq \text{Th}(\mathcal{M})$ holds trivially, since each $\alpha \in T$ naturally holds in all models of T . If, conversely, α holds in all models of T , then the Completeness Theorem 1.5.2 immediately yields $T \vdash \alpha$. Therefore $\alpha \in T$, by the deductive closedness of T .

While for any given nonempty class \mathcal{M} of L -structures we can, indeed, define the set $T = \text{Th}(\mathcal{M})$ purely abstractly, as a rule this set will be completely intractable. We shall return to this problem in Appendix A. Sometimes, however, it is possible to give a reasonable system Σ of axioms for T . Here we call Σ an *axiom system* for T if

$$T = \{ \alpha \in \text{Sent}(L) \mid \Sigma \vdash \alpha \}.$$

Observe that the set $\{ \alpha \in \text{Sent}(L) \mid \Sigma \vdash \alpha \}$, which we wish to denote also by $\text{Ded}(\Sigma)$, is deductively closed. Indeed, if $\alpha_1, \dots, \alpha_n \in \text{Ded}(\Sigma)$ and $\{ \alpha_1, \dots, \alpha_n \} \vdash \alpha$, then we have n proofs of $\alpha_1, \dots, \alpha_n$ from Σ , and a proof of α from $\{ \alpha_1, \dots, \alpha_n \}$; from these $n+1$ proofs we can easily assemble a proof of α directly from Σ . By a “reasonable” axiom system Σ for T we mean one that is effectively enumerable, e.g. finite. Appendix A will explain more precisely how the concept of “effective enumerability” can be given a definition. In the following examples we shall write down the corresponding axiom systems concretely.

Usually the most interesting case of a possible axiomatization of the L -theory of a class \mathcal{M} is that in which \mathcal{M} has exactly one L -structure \mathfrak{A} . In this case we also write $\text{Th}(\mathfrak{A})$ for $\text{Th}(\mathcal{M})$, and (1.6.0.1) simplifies to

$$\text{Th}(\mathfrak{A}) = \{ \alpha \in \text{Sent}(L) \mid \mathfrak{A} \models \alpha \}.$$

Here we also speak of the “ L -theory of \mathfrak{A} ”. For example, if

$$\mathfrak{N} = \langle \mathbb{N}; +^{\mathbb{N}}, \cdot^{\mathbb{N}} \rangle,$$

then $\text{Th}(\mathfrak{N})$ is what we want to be understood by the phrase “number theory”, namely, simply the set of all sentences (of the formal language L appropriate for \mathfrak{A}) that are true in \mathfrak{N} .

For an L -structure \mathfrak{A} , the L -theory $T := \text{Th}(\mathfrak{A})$ has an excellent property: since for every L -sentence α ,

$$\mathfrak{A} \models \alpha \quad \text{or} \quad \mathfrak{A} \models \neg\alpha,$$

we obtain, correspondingly,

$$\alpha \in T \quad \text{or} \quad \neg\alpha \in T.$$

We call an arbitrary L -theory T with this property *complete*. More generally we want to call an arbitrary set Σ of L -sentences *complete* if, for every L -sentence α ,

$$\Sigma \vdash \alpha \quad \text{or} \quad \Sigma \vdash \neg\alpha. \tag{1.6.0.2}$$

The latter definition does not clash with the former in the case where Σ itself is already an L -theory, since in that case Σ would be deductively closed. And the latter definition is more general, as one sees by considering either the consistency, or the deductive closedness, of an L -theory T . If Σ is consistent and $T = \text{Ded}(\Sigma)$, then clearly T is complete in the first sense if and only if Σ is in the second sense.

With this we have finally arrived at a purely syntactical concept, which is then also meaningful when one adopts the finitist standpoint. If L is a language permitting us to express all reasonable mathematical concepts, and if some set Σ is a concrete, consistent axiom system in L (we shall, further below, present the Zermelo–Fraenkel axiom system for set theory as one such system), then the completeness of Σ would mean that one could prove (from Σ) every “true” mathematical sentence. Then the truth of a sentence would be nothing other than its provability. In Appendix A we explain that such an axiom system cannot exist. There we shall see that, for example, the Zermelo–Fraenkel axiom system is incomplete, i.e. that there exist sentences α that are neither provable nor refutable from it (where “ α is refutable” means that $\neg\alpha$ is provable).

Just below we shall introduce a series of axiom systems whose completeness we shall prove in Chapter 3. If $\Sigma \subseteq \text{Sent}(L)$ is complete, and \mathfrak{A} is a model of Σ , then, in particular,

$$\text{Th}(\mathfrak{A}) = \text{Ded}(\Sigma), \tag{1.6.0.3}$$

i.e. $\text{Ded}(\Sigma)$ is the theory of the L -structure \mathfrak{A} . Indeed, since \mathfrak{A} is a model of Σ , we have, on the one hand, $\Sigma \subseteq \text{Th}(\mathfrak{A})$ and hence $\text{Ded}(\Sigma) \subseteq \text{Th}(\mathfrak{A})$. On the other hand, both $T_1 := \text{Ded}(\Sigma)$ and $T_2 := \text{Th}(\mathfrak{A})$ are, first, L -theories (for T_1 this requires noticing that Σ is consistent), and second, complete; and whenever we have an inclusion

$T_1 \subseteq T_2$ between two complete L -theories, we even have equality. Indeed, if $\alpha \in T_2$ and $\alpha \notin T_1$, then $\neg\alpha \in T_1$ and thence $\neg\alpha \in T_2$, contradicting the consistency of T_2 . We record this as a Lemma:

Lemma 1.6.1. *If T_1 and T_2 are complete L -theories with $T_1 \subseteq T_2$, then $T_1 = T_2$. \square*

Now we want to introduce a series of L -theories; for each one, we shall, either in this section or in Chapters 3 or 4, investigate whether it is complete. Each theory considered here will be presented as the deductive closure of a concrete axiom system. While setting up this axiom system, we shall usually have a particular L -structure \mathfrak{A} in view. We usually begin with several general properties of \mathfrak{A} , which we make into axioms, and then we try, by a systematic enlargement of this system, to arrive, finally, at an axiomatization of $\text{Th}(\mathfrak{A})$, i.e. at a complete axiom system.

1. Dense linear orderings with no extrema

The first structure whose theory we want to axiomatize is

$$\mathfrak{R} = \langle \mathbb{R}; <^{\mathbb{R}} \rangle.$$

This is a structure in the language L , for which $J = \emptyset$, $K = \emptyset$ and I consists of a single element i with $\lambda(i) = 2$ (recall (1.2.0.10)–(1.2.0.11)). For the relation symbol R_i we write $<$. The following are L -sentences that hold in \mathfrak{R} :

- O₁: $\forall x \quad \neg x < x$
- O₂: $\forall x, y, z \quad (x < y \wedge y < z \rightarrow x < z)$
- O₃: $\forall x, y \quad (x < y \vee x \dot{=} y \vee y < x)$
- O₄: $\forall x, y \quad (x < y \rightarrow \exists z (x < z \wedge z < y))$
- O₅: $\forall x \exists y, z \quad (y < x \wedge x < z)$

O₁ and O₂ are called the axioms of a *partial ordering*; O₁–O₃ are the axioms of a *linear ordering*. The addition of O₄ says that this linear ordering is *dense*; and the addition of O₅ says that it possesses no *extrema*.

In Section 3.2 we shall see that O₁–O₅ is complete. O₁–O₄ is not complete, since, for example, the sentence O₅ is true in \mathfrak{R} , but not in $\mathfrak{R}^+ := \langle \mathbb{R}^+; <^{\mathbb{R}^+} \rangle$, where \mathbb{R}^+ denotes the set of nonnegative real numbers.

2. Torsion-free, divisible, Abelian groups

Next, let us consider the structure

$$\mathfrak{R} = \langle \mathbb{R}; +^{\mathbb{R}}; 0^{\mathbb{R}} \rangle,$$

which we have, for the sake of simplicity, again denoted by \mathfrak{R} . This time, L has a binary function symbol $+$ and the constant symbol 0 . The following L -sentences hold in \mathfrak{R} :

$$\begin{aligned}
G_1: & \quad \forall x, y, z (x + y) + z \doteq x + (y + z) \\
G_2: & \quad \forall x \quad x + 0 \doteq x \\
G_3: & \quad \forall x \exists y \quad x + y \doteq 0 \\
G_4: & \quad \forall x, y \quad x + y \doteq y + x \\
G_{5,n}: & \quad \forall x \quad (nx \doteq 0 \rightarrow x \doteq 0) \\
G_{6,n}: & \quad \forall x \exists y \quad ny \doteq x
\end{aligned}$$

Here, for each natural number $n \geq 1$, the (informal) expression nx abbreviates the term $x + x + \dots + x$, the n -fold sum of x .⁴ G_1 – G_3 are usually called the *group axioms*; G_4 says that this group is *Abelian*. The set $\{G_{5,n} \mid n \geq 1\}$ of axioms expresses the *torsion free* property of a group; the set $\{G_{6,n} \mid n \geq 1\}$ expresses its *divisibility*. In Section 3.2 we shall prove the completeness of the axiom system

$$\{G_1, G_2, G_3, G_4\} \cup \{G_{5,n} \mid n \geq 1\} \cup \{G_{6,n} \mid n \geq 1\} \cup \{\exists x \, x \neq 0\}$$

of nontrivial, torsion-free, divisible, Abelian groups.

3. Ordered, divisible, Abelian groups

If we consider both the additive group structure and the ordering on \mathbb{R} simultaneously, then one arrives, in the correspondingly extended language, at the structure

$$\mathfrak{R} = \langle \mathbb{R}; <^{\mathbb{R}}, +^{\mathbb{R}}, 0^{\mathbb{R}} \rangle.$$

This is an *ordered, divisible, Abelian group*, i.e. it is a model of all of the above-listed O-axioms, all G-axioms, and additionally the axiom

$$OG: \quad \forall x, y, z \, (x < y \rightarrow x + z < y + z).$$

The totality of all these axioms is, as we shall show in Chapter 4, complete. One sees easily that here the axioms O_4 and $G_{5,n}$ are superfluous, and O_5 can be replaced by $\exists x \, x \neq 0$. We shall also show that the axiom system

$$\{O_1, O_2, O_3, G_1, G_2, G_3, G_4, OG\} \cup \{G_{6,n} \mid n \geq 1\} \cup \{\exists x \, x \neq 0\}$$

is complete. Models of the first part of the above union are called *ordered Abelian groups*.

⁴ More precisely, this term should be written as $(\dots((x+x)+x)+\dots)+x$, since the arity of $+$ is 2. No ambiguity will result from dropping these parentheses, thanks to G_1 .

4. Discrete ordered Abelian groups

Another ordered Abelian group whose theory interests us is

$$\mathfrak{Z} = \langle \mathbb{Z}; <^{\mathbb{Z}}; +^{\mathbb{Z}}; 0^{\mathbb{Z}} \rangle.$$

In order to axiomatize its theory, we observe, first, that it is not divisible: it possesses a smallest, positive element, namely the natural number 1. An ordered Abelian group, which, like \mathfrak{Z} , satisfies the axiom

$$\text{DO: } \exists x(0 < x \wedge \forall y(0 < y \rightarrow x \dot{=} y \vee x < y))$$

is called *discrete ordered*. The element x whose existence DO asserts is unique, by O_1 and O_2 .

5. \mathbb{Z} -groups

To obtain a complete axiomatization of \mathfrak{Z} , it is advisable to add to the language of ordered groups a constant symbol for the smallest positive element. Let this constant symbol be 1. Thus, we now consider the structure

$$\mathfrak{Z} = \langle \mathbb{Z}; <^{\mathbb{Z}}; +^{\mathbb{Z}}; 0^{\mathbb{Z}}, 1^{\mathbb{Z}} \rangle.$$

Besides the axioms for discrete ordered Abelian groups, \mathfrak{Z} satisfies the following family of axioms:

$$D_n: \quad \forall x \exists y \left(\bigvee_{v=1}^n x + v1 \dot{=} ny \right). \quad (1.6.1.1)$$

Here $n = 1, 2, \dots$, and for $1 \leq v \leq n$, the (informal) expression $v1$ abbreviates the term $1 + 1 + \dots + 1$, the n -fold sum of 1. As in (1.2.0.12), the (informal) expression $(\bigvee_{v=1}^n \varphi_v)$ is an abbreviation for the formula

$$(\varphi_1 \vee \dots \vee \varphi_n).$$

A discrete ordered Abelian group that is, in addition, a model of $\{D_n \mid n \geq 1\}$ is called a \mathbb{Z} -group. The reason for this name is the fact (proven later, in Chapter 4) that the axiom system for \mathbb{Z} -groups is complete, and \mathfrak{Z} is a model of it.

6. Ordered fields, real closed fields

Next we consider the structure

$$\mathfrak{R} = \langle \mathbb{R}; +^{\mathbb{R}}, -^{\mathbb{R}}, \cdot^{\mathbb{R}}; 0^{\mathbb{R}}, 1^{\mathbb{R}} \rangle,$$

which we wish again to denote by the letter \mathfrak{R} , even though this structure extends previous structures by the same name. Now \mathfrak{R} satisfies the following axioms:

the *field axioms*:

$$\begin{aligned}
 K_1: & \quad 0 \neq 1 & (1.6.1.2) \\
 K_2: & \quad \forall x, y, z \quad x + (y + z) \doteq (x + y) + z \\
 K_3: & \quad \forall x \quad x + 0 \doteq x \\
 K_4: & \quad \forall x \quad x + (-x) \doteq 0 \\
 K_5: & \quad \forall x, y, z \quad x \cdot (y \cdot z) \doteq (x \cdot y) \cdot z \\
 K_6: & \quad \forall x \quad x \cdot 1 \doteq x \\
 K_7: & \quad \forall x \exists y \quad (x \doteq 0 \vee x \cdot y \doteq 1) \\
 K_8: & \quad \forall x, y \quad x \cdot y \doteq y \cdot x \\
 K_9: & \quad \forall x, y, z \quad (x + y) \cdot z \doteq x \cdot z + y \cdot z;
 \end{aligned}$$

the order axioms O_1, O_2, O_3 , together with:

$$\begin{aligned}
 OK_1: & \quad \forall x, y, z \quad (x < y \rightarrow x + z < y + z) \\
 OK_2: & \quad \forall x, y \quad (0 < x \wedge 0 < y \rightarrow 0 < x \cdot y);
 \end{aligned}$$

and finally, in addition,

$$\begin{aligned}
 RK_1: & \quad \forall x \exists y \quad (x < 0 \vee x \doteq y^2) \\
 RK_{2n}: & \quad \forall x_0, x_1, \dots, x_{2n} \exists y \quad y^{2n+1} + x_{2n} \cdot y^{2n} + \dots + x_1 \cdot y + x_0 \doteq 0.
 \end{aligned}$$

In RK_1 and RK_{2n} (where $n = 1, 2, \dots$), the (informal) expression y^m , for $m = 1, 2, \dots$, abbreviates the term $y \cdot y \cdot \dots \cdot y$, the m -fold product of y .⁵ And in RK_{2n} we utilize the convention in algebra that the operation \cdot is performed before the operation $+$, unless parentheses indicate otherwise. A model of the axioms K_1 – K_9 , O_1 – O_3 , OK_1 and OK_2 is called an *ordered field*. If the axioms RK_1 and RK_{2n} join in (for all $n \geq 1$), then one speaks of a *real closed field*. The reason for this name is the fact (to be proved in Section 4.2) that the theory of real closed fields is complete, and \mathfrak{R} is a model of it. It is easy to see that RK_1 and RK_{2n} hold in \mathfrak{R} , by reflecting on the fact that a polynomial with real coefficients that has a sign-change in \mathbb{R} also has a zero in \mathbb{R} , by the intermediate-value theorem for continuous functions.

7. Algebraically closed fields

We wish to close out the round of structures for whose theories we explicitly give an axiom system, by considering the field

$$\mathfrak{C} = \langle \mathbb{C}; +^{\mathbb{C}}, -^{\mathbb{C}}, \cdot^{\mathbb{C}}; 0^{\mathbb{C}}, 1^{\mathbb{C}} \rangle$$

⁵ More precisely, this term should be written as $(\dots((y \cdot y) \cdot y) \cdot \dots) \cdot y$, since the arity of \cdot is 2. No ambiguity will result from dropping these parentheses, thanks to K_5 .

of complex numbers. \mathfrak{C} satisfies not only the field axioms K_1 – K_9 , but also the family of axioms

$$AK_n: \quad \forall x_0, x_1, \dots, x_n \exists y \quad y^{n+1} + x_n \cdot y^n + \dots + x_0 \doteq 0, \quad (1.6.1.3)$$

for $n = 1, 2, \dots$. The satisfaction of AK_n by \mathfrak{C} is equivalent to the “Fundamental Theorem of Algebra”, which asserts that every non-constant polynomial with coefficients in \mathbb{C} has a zero in \mathbb{C} . Fields that satisfy AK_n for all $n \geq 1$ are called *algebraically closed fields*. The axiom system for algebraically closed fields is not yet complete. But, as we shall show in Chapter 3, it will become complete as soon as we specify the characteristic of the field, i.e. when we adjoin as further axioms either, for some single prime number p , the sentence

$$C_p: \quad p1 \doteq 0 \quad (1.6.1.4)$$

(where $p1$ again denotes the p -fold sum $1 + \dots + 1$), or the set

$$\{ \neg C_q \mid q \in \mathbb{N}, q \text{ prime} \}.$$

In the latter case one obtains a complete axiomatization of the theory of \mathfrak{C} .

8. The natural numbers

For all hitherto considered structures \mathfrak{A} , we were able to give an axiom system $\Sigma \subseteq \text{Sent}(L)$ that, first, axiomatized $\text{Th}(\mathfrak{A})$ (i.e. had the property that $\text{Ded}(\Sigma) = \text{Th}(\mathfrak{A})$), and, second, was “decidable”. This should mean that there is an effective procedure that allows us, for an arbitrary, given L -sentence α , to decide whether α is an axiom in Σ or not. A glance at the various axiom systems we have given will easily convince the reader of the truth of this statement.

As we shall explain in Appendix A, there can be no decidable system of axioms for the structure

$$\mathfrak{N} = \langle \mathbb{N}; +^{\mathbb{N}}, \cdot^{\mathbb{N}}; 0^{\mathbb{N}}, 1^{\mathbb{N}} \rangle.$$

This means, in other words, that every “decidable” system $\Sigma \subseteq \text{Sent}(L)$ of sentences that hold in \mathfrak{N} must necessarily be incomplete. One such system is the following:

- $P_1: \quad \forall x \quad x + 1 \neq 0$
- $P_2: \quad \forall x, y \quad (x + 1 \doteq y + 1 \rightarrow x \doteq y)$
- $P_3: \quad \forall x \quad x + 0 \doteq x$
- $P_4: \quad \forall x, y \quad x + (y + 1) \doteq (x + y) + 1$
- $P_5: \quad \forall x \quad x \cdot 0 \doteq 0$
- $P_6: \quad \forall x, y \quad x \cdot (y + 1) \doteq x \cdot y + x$
- $P_\varphi: \quad \varphi(v_0/0) \wedge \forall v_0 (\varphi(v_0/v_0) \rightarrow \varphi(v_0/v_0 + 1)) \rightarrow \forall v_0 \varphi(v_0/v_0).$

The axiom system

$$\Sigma_{\text{PA}} = \{P_1, \dots, P_6\} \cup \{\forall P_\varphi \mid \varphi \in \text{Fml}(L)\}$$

is usually called the system of *Peano axioms* of arithmetic. The fact that $\forall P_\varphi$ holds in \mathfrak{N} simply means that for every evaluation h in \mathfrak{N} , the set

$$\{a \in \mathbb{N} \mid \mathfrak{N} \models \varphi[h(\frac{v_0}{a})]\}$$

has the following property: if 0 belongs to this set, and if whenever n belongs to it, so does $n + 1$, then every natural number belongs to it.

9. Set theory

Finally, we wish to present an axiom system for set theory – in fact, for *Zermelo–Fraenkel set theory*. The underlying language L has (as in our first axiom system above, namely, in the case of orderings) only one, single relation symbol, of arity 2, which we write here as ε .⁶ The following axioms are oriented toward the membership relation between sets. In order to prevent the axioms from becoming too unreadable, we shall at times introduce common abbreviations:

ZF₁: Axiom of extensionality

$$\forall x, y \ (\forall z (z \varepsilon x \leftrightarrow z \varepsilon y) \rightarrow x \doteq y)$$

If we write $x \subseteq y$ for $\forall z (z \varepsilon x \rightarrow z \varepsilon y)$, then ZF₁ is clearly equivalent to

$$\forall x, y \ (x \subseteq y \wedge y \subseteq x \rightarrow x \doteq y).$$

ZF₁ says that two sets possessing the same elements are equal.

ZF₂: Null set axiom

$$\exists x \forall z \neg z \varepsilon x.$$

This axiom says that a set possessing no elements exists.

ZF₃: Pair set axiom

$$\forall u, v \exists x \forall z (z \varepsilon x \leftrightarrow (z \doteq u \vee z \doteq v))$$

ZF₄: Union set axiom

$$\forall u \exists x \forall z (z \varepsilon x \leftrightarrow \exists v (v \varepsilon u \wedge z \varepsilon v))$$

⁶ Note the distinction between the symbol ε , which belongs to the object language of set theory, and the symbol \in , which we reserve for the metalanguage.

ZF₅: Power set axiom

$$\forall u \exists x \forall z (z \varepsilon x \leftrightarrow z \subseteq u)$$

These last four axioms assert the existence of certain sets x . By ZF₁, each of these sets is uniquely determined, though in the case of ZF₃–ZF₅, x depends on one or two other sets (or “parameters”) u and v . The x in ZF₂, ZF₃, ZF₄ and ZF₅ is usually “denoted” by \emptyset (the “empty set” or “null set”), $\{u, v\}$, $\bigcup u$ and $P(u)$, respectively. This means that one adjoins to the language names (i.e. a constant symbol in the case of ZF₂, and function symbols in the cases of ZF₃–ZF₅) for these uniquely determined objects. This corresponds exactly to the usual progression in mathematics; it contributes to a faster and better understanding of formulae. But it is, in principle, not necessary. Instead of using the name of such a set, one could simply utilize a uniquely characterizing property of it, as follows.

Let ψ be a formula in which x occurs free, and for which we can prove from the axioms that ψ holds for exactly one x (though possibly depending on other variables occurring free in ψ , as mentioned above). For this uniquely determined set we introduce a new symbol, such as \emptyset or $\{u, v\}$, as above. But we do not adjoin this new symbol to the language of set theory – the ε -language. Then, if φ is an ε -formula with free variable y , we shall write $\varphi(y/\emptyset)$ to stand for the following ε -formula:

$$\exists x (\psi(x) \wedge \varphi(y/x)),$$

where $\psi(x)$ is the formula $\forall z \neg z \in x$. Here we assume that x is free for y in φ ; otherwise, instead of x we would utilize some variable that occurs in neither ψ nor φ . In the case of $\{u, v\}$, ψ would be $\forall z (z \varepsilon x \leftrightarrow (z \doteq u \vee z \doteq v))$. As the last example shows, ψ can have, besides x , additional free variables, as mentioned above. In the following axioms we shall make use of these abbreviations.

From the previous axioms it follows that if x and y are sets, then so are $\{x, x\}$ and $\bigcup\{x, y\}$. In the case of $\{x, x\}$ we also write $\{x\}$, and in the case of $\bigcup\{x, y\}$ we also write $x \cup y$. Furthermore, it is customary to write x' for $x \cup \{x\}$. Now the next axiom can be formulated using these notations:

ZF₆: Infinity axiom

$$\exists x (\emptyset \varepsilon x \wedge \forall z (z \varepsilon x \rightarrow z' \varepsilon x)).$$

This axiom ensures the existence of an infinite set. This set is, however, not uniquely determined.

The next axiom says that the image of a set under a function is again a set. Here the function is described (implicitly) by a formula φ , in which the variables u and v occur free. Here and elsewhere we shall use notation such as $\varphi(u, v)$ to emphasize the possibility that u and v occur free in φ ; this notation does not exclude the possibility that other variables might also occur free in φ .

ZF₇: Replacement axiom

$$\forall y (\forall u (u \varepsilon y \rightarrow \exists^1 v \varphi(u, v)) \rightarrow \exists x \forall v (v \varepsilon x \leftrightarrow \exists u (u \varepsilon y \wedge \varphi(u, v)))).$$

Thus x is the image of y under the mapping described by φ . By ZF₁, this x is uniquely determined. If ψ is a formula in which the variable v occurs free, then the abbreviation $\exists^1 v \psi(v)$, used in ZF₇ above, stands for

$$\exists v (\psi(v) \wedge \forall u (\psi(v/u) \rightarrow u \doteq v)),$$

where u is a variable not occurring in ψ .

The last axiom of Zermelo–Fraenkel set theory is usually

ZF₈: Foundation axiom

$$\forall x (x \neq \emptyset \rightarrow \exists z (z \varepsilon x \wedge z \cap x \doteq \emptyset)).$$

Here the notation $z \cap x$ (for “intersection”) denotes a set whose existence is proved using ZF₇, and whose uniqueness is guaranteed by ZF₁.

Usually the so-called Axiom of Choice is added to the above ZF-axioms; we present it in the following form.

ZF₉: Axiom of Choice

$$\forall y, u (u \subseteq P(y) \wedge \forall z_1, z_2 (z_1 \varepsilon u \wedge z_2 \varepsilon u \rightarrow z_1 \neq \emptyset \wedge (z_1 \doteq z_2 \vee z_1 \cap z_2 \doteq \emptyset)) \rightarrow \exists x (x \subseteq y \wedge \forall z (z \varepsilon u \rightarrow \exists v x \cap z \doteq \{v\}))).$$

This axiom says that to every system u of nonempty, pairwise disjoint subsets of a given set y , there exists a “system x of representatives”, i.e. x contains, for every element z of u , exactly one element $v \varepsilon z$, the representative for z .⁷

By consulting the appropriate literature (e.g. [Levy, 1979–2002]), the reader should be able to convince himself that the entirety of mathematics may actually be built up in Zermelo–Fraenkel set theory.

1.7 Exercises for Chapter 1

- 1.7.1. Notation:** $F(x)$ x is female
 $C(x, y)$ x is a child of y
 $M(x, y)$ x and y are married.

Formalize the following statements:

- (a) x is single.
 (b) x is the brother-in-law of y .

⁷ The system x may also contain other elements of y , not belonging to any $z \varepsilon u$.

(c) No two of the three grandchildren of x are siblings.

Here, the variables x, y, \dots range over all people.

1.7.2. Notation: $P(x)$ x is a point
 $G(x)$ x is a straight line
 $E(x)$ x is a plane
 $I(x, y)$ x lies on y

Formalize the following statements:

(a) For every two points, there is a straight line on which these two points lie.

(b) Through two distinct points there is at most one straight line.

(c) For every three points that do not lie on a common line,
 there is exactly one plane on which these three points lie.

Here, the variables x, y, \dots range over all points, lines and planes simultaneously.

1.7.3. Notation: $R(x, y)$ x is less than or equal to y
 $b(x)$ absolute value of x
 $p(x, y)$ x times y
 $s(x, y)$ x plus y
 $d(x, y)$ x minus y
 $f(x)$ one-place function of x
 e one
 n zero

Formalize the following statements:

(a) f is differentiable at 1.

(b) f is uniformly continuous on the interval $[0, 1]$.

Here, the variables x, y, \dots vary over \mathbb{R} .

1.7.4. Let L be a formal language with the one-place function symbol $|\cdot|$, the two-place function symbols $f, +, \cdot, -$, the two-place relation symbol $<$ and the constant symbols $0, 1, x_0, y_0$.

Formalize in L the Implicit Function Theorem: If the function f vanishes at the point (x_0, y_0) , but the partial derivative of f with respect to the second variable does not, then there are open neighbourhoods U of x_0 and V of y_0 such that, for each $x \in U$, there is exactly one $y \in V$ with $f(x, y) = 0$.

1.7.5. Which of the following implications are tautologies?

- (a) $(\alpha \rightarrow (\beta \rightarrow \gamma)) \rightarrow (\alpha \wedge \beta \rightarrow \gamma)$
- (b) $(\alpha \rightarrow (\beta \rightarrow \gamma)) \rightarrow (\alpha \vee \beta \rightarrow \gamma)$
- (c) $(\alpha \rightarrow (\beta \rightarrow \gamma)) \rightarrow ((\alpha \rightarrow \beta) \rightarrow \gamma)$
- (d) $(\alpha \wedge (\beta \rightarrow \gamma)) \rightarrow (\alpha \wedge \beta \rightarrow \gamma)$
- (e) $(\alpha \rightarrow \gamma) \wedge (\beta \rightarrow \gamma) \rightarrow (\alpha \vee \beta \rightarrow \gamma)$

1.7.6. Let L be a formal language with the one-place function symbol $-$, the two-place function symbols $+$ and \cdot , and the constant symbols 0 and 1 . Let Σ be the set of the following four axioms:

$$\begin{aligned}\forall x, y, z \quad (x + y) + z &\doteq x + (y + z) \\ \forall x, y, z \quad (x + y) \cdot z &\doteq x \cdot z + y \cdot z \\ \forall x \quad x + 0 &\doteq x \\ \forall x \quad x + (-x) &\doteq 0\end{aligned}$$

Show that $\Sigma \vdash (0 \cdot 0 \doteq 0)$.

1.7.7. Show:

- (i) $\emptyset \vdash \forall x(\varphi \rightarrow \psi) \rightarrow (\forall x\varphi \rightarrow \forall x\psi)$.

Hint: It could be helpful to apply the tautologies

$$\begin{aligned}(\alpha \rightarrow (\varphi \rightarrow \psi)) &\rightarrow ((\beta \rightarrow \varphi) \rightarrow (\alpha \wedge \beta \rightarrow \psi)) \quad \text{and} \\ (\alpha \wedge \beta \rightarrow \psi) &\rightarrow (\alpha \rightarrow (\beta \rightarrow \psi))\end{aligned}$$

with suitable formulae α and β .

- (ii) $\emptyset \vdash \forall x(\varphi \rightarrow \psi) \rightarrow (\exists x\varphi \rightarrow \psi)$ in case $x \notin \text{Fr}(\psi)$.

1.7.8. Show that for every quantifier-free formula φ there is a likewise quantifier-free formula ψ , with the same free variables, in conjunctive (respectively, disjunctive) normal form, for which $\emptyset \vdash (\varphi \leftrightarrow \psi)$ holds.

1.7.9. Show that

$$\Sigma \vdash (\varphi \rightarrow \psi)$$

does not follow from $\Sigma \cup \{\varphi\} \vdash \psi$.

Hint: In order to show that a formula is not provable from an axiom system, one can give a mathematical domain (such as the real numbers) in which all the axioms hold, but not the formula in question. Furthermore, by the Deduction Theorem, φ must contain a free variable, so that one can construct such a counterexample.

1.7.10. For formulae φ and ψ , prove:

- (a) $\emptyset \vdash (\exists x\varphi \vee \exists x\psi) \leftrightarrow \exists x(\varphi \vee \psi)$
- (b) $\emptyset \vdash (\forall x\varphi \vee \forall y\psi) \rightarrow \forall x(\varphi \vee \psi)$
- (c) $\emptyset \not\vdash \forall x(\varphi \vee \psi) \rightarrow (\forall x\varphi \vee \forall x\psi)$

1.7.11. Show that for every formula φ there is a formula ψ , in prenex normal form, for which $\text{Fr}(\varphi) = \text{Fr}(\psi)$ and $\emptyset \vdash (\varphi \leftrightarrow \psi)$.

1.7.12. Give a formula φ in the language $(+, \cdot, 0, 1)$ of field theory, with $\text{Fr}(\varphi) = \{x_0, \dots, x_n\}$, such that an $(n+1)$ -tuple (a_0, \dots, a_n) of elements of an arbitrary infinite field satisfies the formula if and only if the polynomial $a_0 + a_1X + \dots + a_nX^n$ is irreducible over that field.

1.7.13. Let $L = (+, \cdot, 0, 1)$ be the language of field theory. For which fields \mathcal{A} is the universe $|\mathcal{A}|$ equal to the set

$$\{t^{\mathcal{A}} \mid t \text{ is a constant term}\}?$$

1.7.14. Let L be a formal language. We say that a class \mathcal{M} of L -structures is (*finitely*) *axiomatized* by a (finite) set Σ of L -sentences if \mathcal{M} consists exactly of the models of Σ , i.e. \mathcal{M} is the model class of Σ as defined in Section 3.1.

(i) Show that the class \mathcal{M} of all finite L -structures cannot be axiomatized with L -sentences.

Hint: Suppose that there were an axiomatization. Add suitable additional sentences to these axioms, and apply the Finiteness Theorem to obtain a contradiction.

(ii) Show that the class of infinite L -structures can be axiomatized with infinitely many L -sentences, but not, however, with only finitely many.

1.7.15. Let the language L consist only of the two-place relation symbol $>$. We call an L -structure \mathfrak{A} a *well-ordered set* if $>^{\mathfrak{A}}$ is a strict linear ordering on $|\mathfrak{A}|$, with respect to which there is no infinite descending sequence

$$a_0 >^{\mathfrak{A}} a_1 >^{\mathfrak{A}} \dots >^{\mathfrak{A}} a_n >^{\mathfrak{A}} \dots.$$

Show that the class \mathcal{M} of well-ordered sets cannot be axiomatized⁹ by a set of L -sentences.

Hint: Extend the language by adding constants, if necessary, and then proceed similarly as in Exercise 1.7.14(i).

1.7.16. Let E be a two-place relation symbol, and $L = (E)$ the *language of graphs*. By a *graph* we mean a (possibly infinite) L -structure \mathfrak{A} , in which the two-place relation $E^{\mathfrak{A}}$ is symmetric and irreflexive. (One imagines the universe as a set of points in which any two points x, y are connected by a line whenever $E^{\mathfrak{A}}(x, y)$ holds.)

Which of the following classes of L -structures are axiomatizable?⁹ Either give an axiom system, or show that there can be none.

- (i) The class of all graphs.
- (ii) The class of all L -structures that are not graphs.
- (iii) The class of all finite graphs (i.e. graphs with a finite universe).
- (iv) The class of all infinite graphs.
- (v) The class of all connected graphs.

Here, a graph is called *connected* if for every two points x and y , there is a finite path from x to y , i.e. a finite sequence of edges that begins at x and ends at y .

1.7.17. Let $L = (+, \cdot, 0, 1)$ be the language of rings. Add a three-place relation symbol M . For an arbitrary ring \mathfrak{A} and $a, b, c \in |\mathfrak{A}|$, we now define $M^{\mathfrak{A}}(a, b, c) :\Leftrightarrow a \cdot^{\mathfrak{A}} b = c$.

Give a formula without function symbols that expresses precisely the commutativity of the ring multiplication.

⁹ Recall the definition of “axiomatized” given in Exercise 1.7.14.

1.7.18. Let L be a formal language. For each function symbol f_i of L , we add a new relation symbol F_i with arity increased by one. In this way we obtain the extended language L' . We transform an L -structure \mathfrak{A} into an L' -structure \mathfrak{A}' by the following definition:

$$F_i^{\mathfrak{A}'}(a_1, \dots, a_s, a) :\Leftrightarrow f_i^{\mathfrak{A}}(a_1, \dots, a_s) = a,$$

for elements $a_1, \dots, a_s, a \in |\mathfrak{A}|$.

Show that one can assign to each L -formula φ an L' -formula φ' with the following properties:

- (i) No function symbols occur in φ' .
- (ii) For arbitrary L -structures \mathfrak{A} and arbitrary evaluations h of the variables, we have:

$$\mathfrak{A} \models \varphi[h] \Leftrightarrow \mathfrak{A}' \models \varphi'[h].$$

Hint: First consider prime formulae φ of the form $t \doteq x$, with t a term and x a variable, and define φ' recursively according to the construction of t . Then we define the assignment $\varphi \mapsto \varphi'$ recursively according to the construction of formulae.



<http://www.springer.com/978-1-4471-2175-6>

Mathematical Logic and Model Theory

A Brief Introduction

Prestel, A.; Delzell, C.

2011, X, 194 p., Softcover

ISBN: 978-1-4471-2175-6