Logic I

Course by Wilfried Buchholz

Winter Semester 2001/02, Dept. of Mathematics, LMU München

§1 Syntax und semantics of 1st-order predicate logic

Mathematical logic is the study of the type of reasoning done in mathematics. The conspicuous feature of mathematics, as opposed to other sciences, is the use of proofs instead of observations. In course of time it has turned out that almost all mathematical proofs can be represented in comparatively simple formal language, the language of *1st-order predicate logic*. Mathematical statements are modeled in this language by so-called *formulas*, which are finite strings of basic symbols formed according to certain grammatical rules. We start from a fixed set of *logical* symbols which will be supplemented by a set $\mathcal{L}$ of *nonlogical* (or *mathematical*) symbols where $\mathcal{L}$ depends on the respective branch of mathematics we are going to formalize. Such a set $\mathcal{L}$ will be briefly called a *language*.

Basic logical symbols:

1. Object or individual variables: $v_0, v_1, v_2, \ldots$
2. $\bot$ (falsehood), $\neg$ (negation), $\wedge$ (conjunction), $\vee$ (disjunction), $\to$ (implication)
3. $\forall$ (universal quantifier), $\exists$ (existential quantifier)
4. $\approx$ (equality symbol)

Definition

A *language* is a set $\mathcal{L}$ of symbols (different from the logical symbols) such that every $p \in \mathcal{L}$ is either introduced as a *function symbol* or as a *relation symbol*. Further for each $p \in \mathcal{L}$ an *arity* (i.e., number of argument places) $\#(p) \in \mathbb{N}$ is fixed, where $\#(p) \geq 1$ for relation symbols p. If $\#(p) = n$ we say that p is an *n-ary* symbol. The 0-ary function symbols are also called *(individual) constants*.

An *$\mathcal{L}$-Struktur*) is a pair $\mathcal{M} = (M, (p^{\mathcal{M}})_{p\in\mathcal{L}})$ consiting of a nonempty set M (the *universe* of $\mathcal{M}$) and a family $(p^{\mathcal{M}})_{p\in\mathcal{L}}$ such that:

(i) $p^{\mathcal{M}} \subseteq M^n$, if p is an n-ary relations symbol,

(ii) $p^{\mathcal{M}} : M^n \to M$, if p is an n-ary function symbol with $n \geq 1$,

(iii) $p^{\mathcal{M}} \in M$, if p is a constant.

The universe of an $\mathcal{L}$-structure $\mathcal{M}$ is denoted by $|\mathcal{M}|$.

If the elements of a language $\mathcal{L}$ are given in a certain order, say $\mathcal{L} = \{p_0, p_1, \ldots\}$, then $\mathcal{L}$-structures are usually presented in the form $(M, p_0^{\mathcal{M}}, p_1^{\mathcal{M}}, \ldots)$.

Every $\mathcal{L}$-structure $\mathcal{M}$ assigns a certain meaning to the symbols of $\mathcal{L}$; moreover it fixes the range of the quantifiers: with respect to $\mathcal{M}$ the meaning of $\forall x$ [$\exists x$, resp.] will be declared as "for all elements a of $|\mathcal{M}|$ holds" ["there exists an element a of $|\mathcal{M}|$ such that", resp.].

We are now going to define the $\mathcal{L}$-formulas, i.e., those strings of basic symbols which represent propositions. Before that we have to define the $\mathcal{L}$-terms; these are strings which occur as parts of formulas and, given an $\mathcal{L}$-structure $\mathcal{M}$, represent elements of $|\mathcal{M}|$ (the universe of $\mathcal{M}$).

The set of all $\mathcal{L}$-terms (as well as the set of $\mathcal{L}$-formulas and many other sets in this course) is introduced by a so-called *inductive definition*. This definition principle is of extreme importance in mathematical logic and computer science. An inductive definition of a set Q is given by certain rules $\mathcal{R}_1, \ldots, \mathcal{R}_n$, which regulate the ways in which elements of Q are to be generated. This always includes the silent agreement, that Q should

consist *exactly* of those objects which can generated by a finite number of successive applications of (some of the rules) $\mathcal{R}_1, ..., \mathcal{R}_n$. Equivalent to this explication is to characterize Q as the *least* set X which is *closed* under $\mathcal{R}_1, ..., \mathcal{R}_n$.

Inductive definition of $\mathcal{L}$-terms

1. Every variable v_i is an $\mathcal{L}$-term;
2. if $f \in \mathcal{L}$ is an n-ary functions symbol ($n \geq 0$), and if $t_1, \ldots, t_n$ are $\mathcal{L}$-terms,
   then the string $ft_1 \ldots t_n$ is an $\mathcal{L}$-Term too.

Remark:

Rule 2. includes the following agreement (for $n = 0$): "Every constant $c \in \mathcal{L}$ is an $\mathcal{L}$-Term".

Definition (Atomic formulas): Strings of the kind $\bot$ or $\approx st$ or $Rt_1...t_n$, where $s, t, t_1, ..., t_n$ are $\mathcal{L}$-terms and $R \in \mathcal{L}$ is an n-ary relation symbol, are called *atomic* $\mathcal{L}$-formulas or *prime* formulas of $\mathcal{L}$.

Inductive definition of $\mathcal{L}$-formulas

1. Each atomic $\mathcal{L}$-formula is an $\mathcal{L}$-formula;
2. if A and B are $\mathcal{L}$-formulas, then $\neg A$, $\wedge AB$, $\vee AB$, and $\rightarrow AB$ are $\mathcal{L}$-formulas;
3. if A is an $\mathcal{L}$-formula and x is a variable, then $\forall xA$ and $\exists xA$ are $\mathcal{L}$-formulas.

Notation: Usually we write $(s \approx t)$, $(A \wedge B)$, $(A \rightarrow B)$, etc. for $\approx st$, $\wedge AB$, $\rightarrow AB$, etc. *(infix notation)*
Example:
$\mathcal{L} := \{\oplus, \otimes, \mathbf{0}, \mathbf{1}, \prec\}$,
where $\oplus, \otimes$ are 2-ary function symbols, $\mathbf{0}, \mathbf{1}$ are constants, and $\prec$ is a 2-ary relation symbol.

Some $\mathcal{L}$-terms: $\oplus v_2 \oplus v_1 v_0$, $\oplus\oplus v_2 v_1 v_2$, $\oplus\otimes v_0 v_1 \otimes v_0 v_2$, $\otimes v_0 \oplus v_1 v_2$.

Some $\mathcal{L}$-formulas: $\forall v_0 \forall v_1 \forall v_2 \wedge \approx \oplus v_0 \oplus v_1 v_2 \oplus\oplus v_0 v_1 v_2 \approx \oplus\otimes v_0 v_1 \otimes v_0 v_2 \otimes v_0 \oplus v_1 v_2$,
$\forall v_0 \rightarrow\neg \approx v_0 \mathbf{0} \exists v_1 \approx \otimes v_0 v_1 \mathbf{1}$, $\rightarrow\approx v_0 v_1 \rightarrow\prec v_1 v_2 \rightarrow\approx v_2 v_3 \prec v_3 v_4$.

Using *infix notation* and applying the "usual" rules for saving parentheses these terms and formulas become:
$v_2 \oplus (v_1 \oplus v_0)$, $(v_2 \oplus v_1) \oplus v_2$, $(v_0 \otimes v_1) \oplus (v_0 \otimes v_2)$, $v_0 \otimes (v_1 \oplus v_2)$,
$\forall v_0 \forall v_1 \forall v_2 (v_0 \oplus (v_1 \oplus v_2) \approx (v_0 \oplus v_1) \oplus v_2 \wedge (v_0 \otimes v_1) \oplus (v_0 \otimes v_2) \approx v_0 \otimes (v_1 \oplus v_2))$,
$\forall v_0 (\neg (v_0 \approx \mathbf{0}) \rightarrow \exists v_1 (v_0 \otimes v_1 \approx \mathbf{1}))$,
$v_0 \approx v_1 \rightarrow (v_1 \prec v_2 \rightarrow (v_2 \approx v_3 \rightarrow v_3 \prec v_4))$.

For obviuos reasons this $\mathcal{L}$ is often called "the language of ordered fields".
The ordered field of real numbers $(\mathbb{R}, +_{\mathbb{R}}, \times_{\mathbb{R}}, 0, 1, <_{\mathbb{R}})$ is an $\mathcal{L}$-structure; but by no means every $\mathcal{L}$-structure is an ordered field, consider e.g. $(\mathbb{N}, +_{\mathbb{N}}, \times_{\mathbb{N}}, 0, 1, <_{\mathbb{N}})$.

Abbreviations:

VARS $:= \{v_0, v_1, v_2, \ldots\}$;

TER$_{\mathcal{L}}$:= set of all $\mathcal{L}$-terms;

FOR$_{\mathcal{L}}$:= set of all $\mathcal{L}$-Formulas.

Definition

Now we also assign an arity $\#(p)$ to each logical symbol p:

$\#(v_i) := 0$, $\#(\bot) := 0$, $\#(\neg) := 1$, $\#(\forall) := \#(\exists) := \#(\approx) := \#(\wedge) := \#(\vee) := \#(\rightarrow) := 2$.

Terms and formulas are called *expressions*.

Remark

If u is an $\mathcal{L}$-expression, then exactly one of the following cases holds

(i) $u \in \text{VARS}$

(ii) $u = \mathsf{Q}xA$ with $\mathsf{Q} \in \{\forall, \exists\}$, $x \in \text{VARS}$, $A \in \text{FOR}_{\mathcal{L}}$

(iii) $u = pu_1...u_n$ with $p \in \mathcal{L} \cup \{\approx, \neg, \wedge, \vee, \rightarrow\}$, $n = \#(p)$ and uniquely determined $\mathcal{L}$-expressions $u_1, ..., u_n$.

The uniqueness of $u_1, ..., u_n$ in (iii) is a consequence of the following lemma.

Lemma 1.1 (Unique readability)

If $u_1, ..., u_n, w_1, ..., w_m$ are expressions such that $u_1...u_n = w_1...w_m$, then $m = n$ and $u_i = w_i$ for $i = 1, ..., n$.

Proof by induction on the length of $u_1...u_n$:

We have $u_1 = p\tilde{u}_1...\tilde{u}_k$, $w_1 = p\tilde{w}_1...\tilde{w}_k$ with $\#(p) = k$. Then obviously $\tilde{u}_1...\tilde{u}_k u_2...u_n = \tilde{w}_1...\tilde{w}_k w_2...w_m$. By I.H. from this we get $m = n$, $\tilde{u}_j = \tilde{w}_j$ for $j = 1, ..., k$, and $u_i = w_i$ for $i = 2, ..., n$. Hence $u_i = w_i$ for $i = 1, ..., n$.

We will use the following *syntactical variables*:

x, y, z for variables $(v_0, v_1, ...)$, f, f_i for function symbols, R, R_i for relation symbols, $\mathcal{L}$ for formal languages, $\mathcal{M}$ for ($\mathcal{L}$-)structures, ξ, η for ($\mathcal{M}$-)assignments, s, t for terms, A, B, C, D for formulas, Γ, Σ for sets of formulas, u for expressions, p for symbols from $\mathcal{L} \cup \{\approx, \bot, \neg, \wedge, \vee, \rightarrow\}$

Definition

An $\mathcal{M}$-*assignment* (or $\mathcal{M}$-*environment*) is a mapping $\eta : \text{VARS} \rightarrow |\mathcal{M}|$.

An $\mathcal{L}$-*Interpretation* is a pair $\mathcal{I} = (\mathcal{M}, \eta)$ consisting of an $\mathcal{L}$-Structure $\mathcal{M}$ and an $\mathcal{M}$-assignment η.

Definition of the *value* $[\![t]\!]^{\mathcal{M}}_{\eta}$ of an $\mathcal{L}$-Terms t with respect to an $\mathcal{L}$-Interpretation $(\mathcal{M}, \eta)$

(The definition proceeds by recursion on the build-up of t and makes essential use of lemma 1.1.)

1. $[\![x]\!]^{\mathcal{M}}_{\eta} := \eta(x)$
2. $[\![ft_1...t_n]\!]^{\mathcal{M}}_{\eta} := f^{\mathcal{M}}([\![t_1]\!]^{\mathcal{M}}_{\eta}, \ldots, [\![t_n]\!]^{\mathcal{M}}_{\eta})$

Example: Let $\mathcal{L} := \{\oplus, \otimes, \mathbf{0}, \mathbf{1}\}$, $\mathcal{M} := (\mathbb{N}, +_{\mathbb{N}}, \times_{\mathbb{N}}, 0, 1)$, and $\eta(v_i) := i + 3$.
$[\![\oplus\otimes v_0 v_1 \otimes v_0 v_2]\!]^{\mathcal{M}}_{\eta} = \oplus^{\mathcal{M}}([\![\otimes v_0 v_1]\!]^{\mathcal{M}}_{\eta}, [\![\otimes v_0 v_2]\!]^{\mathcal{M}}_{\eta}) = \oplus^{\mathcal{M}}(\otimes^{\mathcal{M}}(\eta(v_0), \eta(v_1)), \otimes^{\mathcal{M}}(\eta(v_0), \eta(v_2))) =$
$\oplus^{\mathcal{M}}(\otimes^{\mathcal{M}}(3, 4), \otimes^{\mathcal{M}}(3, 5)) = \oplus^{\mathcal{M}}(12, 15) = 27.$

Definition

If η is an $\mathcal{M}$-assignment, $a \in |\mathcal{M}|$, and $x \in \text{VARS}$, then the modified $\mathcal{M}$-assignment η^a_x is defined as follows:
$\eta^a_x(y) := \begin{cases} a & \text{if } x = y \\ \eta(y) & \text{otherwise} \end{cases}$.
The numbers $0, 1$ are also used as *truth values*, namely 1 for "true" and 0 for "false".

Definition of the *truth value* $[\![A]\!]^{\mathcal{M}}_{\eta}$ of an $\mathcal{L}$-formula A with respect to an $\mathcal{L}$-interpretation $(\mathcal{M}, \eta)$

(The definition proceeds by recursion on the build-up of A.)

1. $[\![s \approx t]\!]^{\mathcal{M}}_{\eta} := \begin{cases} 1 & \text{if } [\![s]\!]^{\mathcal{M}}_{\eta} = [\![t]\!]^{\mathcal{M}}_{\eta} \\ 0 & \text{otherwise} \end{cases}$
2. $[\![Rt_1...t_n]\!]^{\mathcal{M}}_{\eta} := \begin{cases} 1 & \text{if } ([\![t_1]\!]^{\mathcal{M}}_{\eta}, \ldots, [\![t_n]\!]^{\mathcal{M}}_{\eta}) \in R^{\mathcal{M}} \\ 0 & \text{otherwise} \end{cases}$

3. $[\![\neg A]\!]^{\mathcal{M}}_{\eta} := \begin{cases} 1 & \text{if } [\![A]\!]^{\mathcal{M}}_{\eta} = 0 \\ 0 & \text{otherwise} \end{cases}$

4. $[\![A \wedge B]\!]^{\mathcal{M}}_{\eta} := \begin{cases} 1 & \text{if } [\![A]\!]^{\mathcal{M}}_{\eta} = 1 \text{ and } [\![B]\!]^{\mathcal{M}}_{\eta} = 1 \\ 0 & \text{otherwise} \end{cases}$

5. $[\![A \vee B]\!]^{\mathcal{M}}_{\eta} := \begin{cases} 0 & \text{if } [\![A]\!]^{\mathcal{M}}_{\eta} = 0 \text{ and } [\![B]\!]^{\mathcal{M}}_{\eta} = 0 \\ 1 & \text{otherwise} \end{cases}$

6. $[\![A \to B]\!]^{\mathcal{M}}_{\eta} := \begin{cases} 0 & \text{if } [\![A]\!]^{\mathcal{M}}_{\eta} = 1 \text{ and } [\![B]\!]^{\mathcal{M}}_{\eta} = 0 \\ 1 & \text{otherwise} \end{cases}$

7. $[\![\forall x A]\!]^{\mathcal{M}}_{\eta} := \begin{cases} 1 & \text{if } [\![A]\!]^{\mathcal{M}}_{\eta^a_x} = 1 \text{ for all } a \in |\mathcal{M}| \\ 0 & \text{otherwise} \end{cases}$

8. $[\![\exists x A]\!]^{\mathcal{M}}_{\eta} := \begin{cases} 1 & \text{if } [\![A]\!]^{\mathcal{M}}_{\eta^a_x} = 1 \text{ for at least one } a \in |\mathcal{M}| \\ 0 & \text{otherwise} \end{cases}$

Example: Let $\mathcal{M} = (\mathbb{N}, +_{\mathbb{N}}, \times_{\mathbb{N}}, 0, 1)$ as before.
$[\![\forall x(\neg(x \approx \mathbf{0}) \to \exists y(x \otimes y \approx \mathbf{1}))]\!]^{\mathcal{M}}_{\eta} = 1 \iff$
$[\![\neg(x \approx \mathbf{0}) \to \exists y(x \otimes y \approx \mathbf{1}))]\!]^{\mathcal{M}}_{\eta^a_x} = 1$ for all $a \in \mathbb{N} \iff$
(if $[\![\neg(x \approx \mathbf{0})]\!]^{\mathcal{M}}_{\eta^a_x} = 1$, then $[\![\exists y(x \otimes y \approx \mathbf{1})]\!]^{\mathcal{M}}_{\eta^a_x} = 1$) for all $a \in \mathbb{N} \iff$
(iff $[\![\neg(x \approx \mathbf{0})]\!]^{\mathcal{M}}_{\eta^a_x} = 1$, then $[\![x \otimes y \approx \mathbf{1}]\!]^{\mathcal{M}}_{(\eta^a_x)^b_y} = 1$ for some $b \in \mathbb{N}$) for all $a \in \mathbb{N} \iff$
(if $a \neq 0$, dann $(a \times_{\mathbb{N}} b = 1$ for some $b \in \mathbb{N})$) for all $a \in \mathbb{N} \iff$
For all $a \in \mathbb{N} \setminus \{0\}$ there exists a $b \in \mathbb{N}$ such that $a \times_{\mathbb{N}} b = 1$.

In the following we assume a fixed language $\mathcal{L}$ and briefly speak of terms, formulas, interpretations, etc., instead of $\mathcal{L}$-terms, $\mathcal{L}$-formulas, $\mathcal{L}$-interpretations, etc.

Definition

Let $(\mathcal{M}, \eta)$ be an interpretation.

$\mathcal{M} \models A[\eta] \iff [\![A]\!]^{\mathcal{M}}_{\eta} = 1$ $\begin{cases} A \textit{ is valid (or holds) in } (\mathcal{M}, \eta) & \text{or} \\ (\mathcal{M}, \eta) \textit{ satisfies } A & \text{or} \quad (\mathcal{M}, \eta) \textit{ is a model of } A \end{cases}$

$\mathcal{M} \not\models A[\eta] \iff$ not $\mathcal{M} \models A[\eta]$ ($\iff [\![A]\!]^{\mathcal{M}}_{\eta} = 0$).

Let Γ be a set of formulas.

We say that $(\mathcal{M}, \eta)$ is a *model* of Γ (in symbols $\mathcal{M} \models \Gamma[\eta]$), if $(\mathcal{M}, \eta)$ is a model of each $A \in \Gamma$.

A is a *(logical) consequence* of Γ (in symbols $\Gamma \models A$), if A is valid in each model of Γ:

$$\boxed{\Gamma \models A \iff \forall \mathcal{M}, \eta [\, \mathcal{M} \models \Gamma[\eta] \;\Rightarrow\; \mathcal{M} \models A[\eta] \,]}$$

A is *(logically) valid* (in symbols $\models A$), if A is valid in each interpretation.

Remark: $\models A \iff \emptyset \models A$.

A (set of) formula(s) X is *satisfiable*, if it has at least one model, i.e.,
if there exists an interpreation $(\mathcal{M}, \eta)$ such that $\mathcal{M} \models X[\eta]$.

Two formulas A, B are *(logically) equivalent*, if $[\![A]\!]^{\mathcal{M}}_{\eta} = [\![B]\!]^{\mathcal{M}}_{\eta}$ for each interpretation $(\mathcal{M}, \eta)$.

Remark. Let $A \leftrightarrow B := (A \to B) \wedge (B \to A)$. Then we have: A, B equivalent $\iff \models A \leftrightarrow B$.

Remark

For each interpretation $(\mathcal{M}, \eta)$ the following holds:

$\mathcal{M} \models (s \approx t)[\eta] \quad \Longleftrightarrow [\![s]\!]^{\mathcal{M}}_{\eta} = [\![t]\!]^{\mathcal{M}}_{\eta}$

$\mathcal{M} \models (Rt_1 ... t_n)[\eta] \Longleftrightarrow ([\![t_1]\!]^{\mathcal{M}}_{\eta}, \ldots, [\![t_n]\!]^{\mathcal{M}}_{\eta}) \in R^{\mathcal{M}}$

$\mathcal{M} \models (\neg A)[\eta] \quad \Longleftrightarrow \mathcal{M} \not\models A[\eta]$

$\mathcal{M} \models (A \wedge B)[\eta] \Longleftrightarrow \mathcal{M} \models A[\eta]$ and $\mathcal{M} \models B[\eta]$

$\mathcal{M} \models (A \vee B)[\eta] \Longleftrightarrow \mathcal{M} \models A[\eta]$ or $\mathcal{M} \models B[\eta]$

$\mathcal{M} \models (A \to B)[\eta] \Longleftrightarrow \mathcal{M} \models A[\eta]$ implies $\mathcal{M} \models B[\eta]$ (i.e., if $\mathcal{M} \models A[\eta]$, then $\mathcal{M} \models B[\eta]$)

$\mathcal{M} \models (\forall x A)[\eta] \quad \Longleftrightarrow \mathcal{M} \models A[\eta^a_x]$ for all $a \in |\mathcal{M}|$

$\mathcal{M} \models (\exists x A)[\eta] \quad \Longleftrightarrow \mathcal{M} \models A[\eta^a_x]$ for some $a \in |\mathcal{M}|$

Remark

A function $\Phi : \{0,1\}^n \to \{0,1\}$ is called *truth function* oder *boolean function*.

The boolean functions $\mathsf{W}_\neg, \mathsf{W}_\wedge, \mathsf{W}_\vee, \mathsf{W}_\to$ are defined as follows

$\mathsf{W}_\neg(a) := 1 - a$, $\mathsf{W}_\wedge(a,b) := \min\{a,b\}$, $\mathsf{W}_\vee(a,b) := \max\{a,b\}$, $\mathsf{W}_\to(a,b) := \max\{1-a, b\}$.

In addition we set $\mathsf{W}_\bot := 0$, and finally

$p^{\mathcal{M}} := \mathsf{W}_p$, if $p \in \{\bot, \neg, \wedge, \vee, \to\}$;

$p^{\mathcal{M}}(a_1, ..., a_n) := \begin{cases} 1 & \text{if } (a_1, ..., a_n) \in p^{\mathcal{M}} \\ 0 & \text{otherwise} \end{cases}$, if $p \in \mathcal{L}$ is an n-ary relation symbol;

$\approx^{\mathcal{M}}(a,b) := \begin{cases} 1 & \text{if } a = b \\ 0 & \text{otherwise} \end{cases}$.

Then the definitions of $[\![t]\!]^{\mathcal{M}}_{\eta}$ and $[\![A]\!]^{\mathcal{M}}_{\eta}$ can be condensed as follows:

1. $[\![x]\!]^{\mathcal{M}}_{\eta} := \eta(x)$
2. $[\![pu_1 ... u_n]\!]^{\mathcal{M}}_{\eta} := p^{\mathcal{M}}([\![u_1]\!]^{\mathcal{M}}_{\eta}, \ldots, [\![u_n]\!]^{\mathcal{M}}_{\eta})$ $(p \in \mathcal{L} \cup \{\approx, \bot, \neg, \wedge, \vee, \to\},\ n = \#(p))$
3. $[\![\forall x A]\!]^{\mathcal{M}}_{\eta} := \min_{a \in |\mathcal{M}|} [\![A]\!]^{\mathcal{M}}_{\eta^a_x}$, $[\![\exists x A]\!]^{\mathcal{M}}_{\eta} := \max_{a \in |\mathcal{M}|} [\![A]\!]^{\mathcal{M}}_{\eta^a_x}$.

Definition of the sets of variables $\mathrm{FV}(u)$, $\mathrm{BV}(u)$

1. $\mathrm{FV}(x) := \{x\}$ and $\mathrm{BV}(x) := \emptyset$.
2. $\mathrm{FV}(pu_1 ... u_n) := \mathrm{FV}(u_1) \cup \ldots \cup \mathrm{FV}(u_n)$ and $\mathrm{BV}(pu_1 ... u_n) := \mathrm{BV}(u_1) \cup \ldots \cup \mathrm{BV}(u_n)$.
3. $\mathrm{FV}(\mathsf{Q}xA) := \mathrm{FV}(A) \setminus \{x\}$ and $\mathrm{BV}(\mathsf{Q}xA) := \mathrm{BV}(A) \cup \{x\}$ $\quad (\mathsf{Q} \in \{\forall, \exists\})$.

The elements of $\mathrm{FV}(u)$ ($\mathrm{BV}(u)$, resp.) are called the *free* (*bound*, resp.) variables of u. If $\mathrm{FV}(u) = \emptyset$, then u is called *closed*; a closed formula is called a *sentence*.

Abbreviation: $\mathrm{vars}(u) := \mathrm{FV}(u) \cup \mathrm{BV}(u)$

Remark: $\mathrm{vars}(u)$ is the set of *all* variables occurring in u. $\mathrm{FV}(u)$ and $\mathrm{BV}(u)$ need not be disjunct. If u is a quantifier-free expression, then $\mathrm{BV}(u) = \emptyset$ and $\mathrm{vars}(u) = \mathrm{FV}(u)$.

Lemma 1.2 (Coincidence Lemma)
$[\![u]\!]^{\mathcal{M}}_{\eta} = [\![u]\!]^{\mathcal{M}}_{\xi}$, if $\eta(x) = \xi(x)$ for all $x \in \mathrm{FV}(u)$.

Proof by induction on the build-up of u: For brevity we omit the $^{\mathcal{M}}$ in $[\![\cdot]\!]^{\mathcal{M}}_{\eta}$.

1. $u \in$ VARS: Then $u \in \mathrm{FV}(u)$ und $[\![u]\!]_{\eta} = \eta(u) = \xi(u) = [\![u]\!]_{\xi}$.

2. $u = pu_1 ... u_n$: Then $[\![u]\!]_{\eta} = p^{\mathcal{M}}([\![u_1]\!]_{\eta}, ..., [\![u_n]\!]_{\eta}) \overset{\text{IH}}{=} p^{\mathcal{M}}([\![u_1]\!]_{\xi}, ..., [\![u_n]\!]_{\xi}) = [\![pu_1 ... u_n]\!]_{\xi} = [\![u]\!]_{\xi}$.

3. $u = \mathsf{Q}xA$: For each $y \in \mathrm{FV}(A)$ we have $y = x$ or $x \neq y \in \mathrm{FV}(u)$ and thus $\eta^a_x(y) = a = \xi^a_x(y)$ or $\eta^a_x(y) = \eta(y) = \xi(y) = \xi^a_x(y)$. By I.H. from this we get $[\![A]\!]_{\eta^a_x} = [\![A]\!]_{\xi^a_x}$ (for all $a \in |\mathcal{M}|$), and thus $[\![u]\!]_{\eta} = \mathsf{Q}_{a \in M}[\![A]\!]_{\eta^a_x} = \mathsf{Q}_{a \in M}[\![A]\!]_{\xi^a_x} = [\![u]\!]_{\xi}$.

Remark

As Lemma 1.2 shows the value $[\![u]\!]^{\mathcal{M}}_{\eta}$ of a closed expression does not depend on η; therefore one usually writes $[\![u]\!]^{\mathcal{M}}$ instead of $[\![u]\!]^{\mathcal{M}}_{\eta}$.

Correspondingly for closed formulas A and sets of closed formulas Γ we set:

$\mathcal{M}$ is a model of $A \iff \mathcal{M} \models A \iff [\![A]\!]^{\mathcal{M}} = 1$;

$\mathcal{M}$ is a model of $\Gamma \iff \mathcal{M} \models \Gamma \iff [\![A]\!]^{\mathcal{M}} = 1$ for all $A \in \Gamma$.

Substitution

Definition of $u(x/t)$

1. $y(x/t) := \begin{cases} t & \text{if } x = y \\ y & \text{otherwise} \end{cases}$, 2. $(pu_1 ... u_n)(x/t) := pu_1(x/t) \ldots u_n(x/t)$,

3. $(\mathsf{Q}yA)(x/t) := \begin{cases} \mathsf{Q}yA & \text{if } x \notin \mathrm{FV}(\mathsf{Q}yA) \\ \mathsf{Q}yA(x/t) & \text{otherwise} \end{cases}$

Notation: Instead of $u(x/t)$ we also write $u_x(t)$.

Informally said, $u(x/t)$ results from u by replacing each free occurrence of x by t.

Definition of $\mathsf{subst}(u, x, t)$

1. $\mathsf{subst}(y, x, t)$ holds for each y.

2. $\mathsf{subst}(pu_1 ... u_n, x, t) :\Leftrightarrow \mathsf{subst}(u_1, x, t)$ & ... & $\mathsf{subst}(u_n, x, t)$.

3. $\mathsf{subst}(\mathsf{Q}yA, x, t) :\Leftrightarrow x \notin \mathrm{FV}(\mathsf{Q}yA)$ or $(y \notin \mathrm{FV}(t)$ & $\mathsf{subst}(A, x, t))$.

Informally said, $\mathsf{subst}(u, x, t)$ holds iff by the substitution $u \mapsto u(x/t)$ no variable $y \in \mathrm{FV}(t)$ gets into the range of a quantifier $\mathsf{Q}y$. If $\mathsf{subst}(u, x, t)$ holds we say that "*t is substitutable* (or *free*) *for x in u*".

Remark: $\mathrm{FV}(t) \cap \mathrm{BV}(u) = \emptyset \implies \mathsf{subst}(u, x, t)$.

Lemma 1.3 (Substitution Lemma)
$\mathsf{subst}(u, x, t) \implies [\![u_x(t)]\!]^{\mathcal{M}}_{\eta} = [\![u]\!]^{\mathcal{M}}_{\eta^b_x}$ with $b := [\![t]\!]^{\mathcal{M}}_{\eta}$.

Proof by induction on the build-up of u:

1. $u = x$: $[\![u_x(t)]\!]_{\eta} = [\![t]\!]_{\eta} = b = [\![u]\!]_{\eta^b_x}$.

2. $u \in \text{VARS} \setminus \{x\}$: $[\![u_x(t)]\!]_{\eta} = [\![u]\!]_{\eta} = \eta(u) = [\![u]\!]_{\eta^b_x}$.

3. $u = pu_1 ... u_n$: $[\![u(x/t)]\!]_{\eta} = [\![pu_1(x/t) \ldots u_n(x/t)]\!]_{\eta} = p^{\mathcal{M}}([\![u_1(x/t)]\!]_{\eta}, \ldots, [\![u_n(x/t)]\!]_{\eta}) \overset{\text{I.V.}}{=}$

$= p^{\mathcal{M}}([\![u_1]\!]_{\eta_x^b}, \ldots, [\![u_n]\!]_{\eta_x^b}) = [\![pu_1 \ldots u_n]\!]_{\eta_x^b}$.

4. $u = \mathsf{Q}yA$ and $x \notin \mathrm{FV}(u)$: $[\![u_x(t)]\!]_\eta = [\![u]\!]_\eta \overset{\text{L.1.2}}{=} [\![u]\!]_{\eta_x^b}$.

5. $u = \mathsf{Q}yA$ and $x \in \mathrm{FV}(u)$: Then $x \neq y$ & $y \notin \mathrm{FV}(t)$ & $\mathsf{subst}(A, x, t)$.

Since $y \notin \mathrm{FV}(t)$, we have $[\![t]\!]_{\eta_y^a} = [\![t]\!]_\eta = b$ for arbitrary a $(*)$.

$[\![u_x(t)]\!]_\eta = [\![\mathsf{Q}yA_x(t)]\!]_\eta = \mathsf{Q}_{a\in M}[\![A_x(t)]\!]_{\eta_y^a} \overset{\text{IH}+\mathsf{subst}(A,x,t)+(*)}{=} \mathsf{Q}_{a\in M}[\![A]\!]_{(\eta_y^a)_x^b} \overset{x\neq y}{=} \mathsf{Q}_{a\in M}[\![A]\!]_{(\eta_x^b)_y^a} = [\![u]\!]_{\eta_x^b}$.

[Here $\mathsf{Q}_{a\in M}$ stands for $\min_{a\in M}$ or $\max_{a\in M}$, respectively.]

Corollary. $\mathsf{subst}(A, x, t)$ & $b = [\![t]\!]_\eta^{\mathcal{M}} \implies (\ \mathcal{M} \models A_x(t)[\eta] \iff \mathcal{M} \models A[\eta_x^b]\ )$.

Counterexample: $A = \forall y(x \approx y)$, $t = y$.

$\mathcal{M} \models A_x(t)[\eta] \iff \mathcal{M} \models \forall y(y \approx y)[\eta]$.

$\mathcal{M} \models A[\eta_x^b] \iff \mathcal{M} \models (x \approx y)[(\eta_x^b)_y^a]$ for all $a \in |\mathcal{M}| \iff b = a$ for all $a \in |\mathcal{M}|$.

An Example from Analysis:

$\int_0^3 (x^2 + y)dx = [\frac{1}{3}x^3 + xy]_0^3 = 9 + 3y$

$\int_0^3 (x^2 + x)dx = [\frac{1}{3}x^3 + \frac{1}{2}x^2]_0^3 = 9 + \frac{9}{2}$, but $(9 + 3y)_y(x) = 9 + 3x$

Lemma 1.4

(a) $x \notin \mathrm{FV}(u) \implies \mathsf{subst}(u, x, t)$ & $u_x(t) = u$.

(b) $\mathrm{FV}(u_x(t)) \subseteq (\mathrm{FV}(u) \setminus \{x\}) \cup \mathrm{FV}(t)$, where "=" holds, if $x \in \mathrm{FV}(u)$ & $\mathsf{subst}(u, x, t)$.

(c) $x \neq y \implies (\mathsf{Q}yA)_x(t) = \mathsf{Q}yA_x(t)$.

(d) $\mathsf{subst}(u, x, x)$ & $u_x(x) = u$.

(e) $y \notin \mathrm{FV}(u)$ & $\mathsf{subst}(u, x, y) \implies u_x(y)_y(t) = u_x(t)$ and $(\mathsf{subst}(u_x(y), y, t) \Leftrightarrow \mathsf{subst}(u, x, t))$.

(f) $x \neq y$ & $y \notin \mathrm{FV}(t)$ & $\mathsf{subst}(u, y, s) \implies u_y(s)_x(t) = u_x(t)_y(s_x(t))$.

Proof by induction on the build-up of u:

(a) 1. $u = y$: $\mathsf{subst}(u, x, t)$ holds by definition. From $x \notin \mathrm{FV}(u)$ we get $y \neq x$ and thus $u_x(t) = u$.

2. $u = pu_1 ... u_n$: The claim follows immediately from the I.H.

3. $u = \mathsf{Q}yA$: The claim holds by definition.

(b),(d),(e) Exercises.

(c) Assume $x \notin \mathrm{FV}(\mathsf{Q}yA)$ (otherwise the assertion holds by definition). Since $x \neq y$, we then also have $x \notin \mathrm{FV}(A)$ and thus – by (a) – $\mathsf{Q}yA_x(t) = \mathsf{Q}yA = (\mathsf{Q}yA)_x(t)$.

(f) 1. $y \notin \mathrm{FV}(u)$: By (b) we have $\mathrm{FV}(u_x(t)) \subseteq (\mathrm{FV}(u) \setminus \{x\}) \cup \mathrm{FV}(t)$, and so also $y \notin \mathrm{FV}(u_x(t))$ which yields $u_y(s)_x(t) = u_x(t) = u_x(t)_y(s_x(t))$.

2. $y \in \mathrm{FV}(u)$ & $x \notin \mathrm{FV}(u_y(s))$: Since $y \in \mathrm{FV}(u)$ & $\mathsf{subst}(u, y, s)$, by (b) we have $\mathrm{FV}(u_y(s)) = (\mathrm{FV}(u)\setminus\{y\}) \cup \mathrm{FV}(s)$. Hence from $y \neq x \notin \mathrm{FV}(u_y(s))$ we get $x \notin \mathrm{FV}(u) \cup \mathrm{FV}(s)$, and thus $u_y(s)_x(t) = u_y(s) = u_x(t)_y(s_x(t))$.

3. Assume $y \in \mathrm{FV}(u)$ & $x \in \mathrm{FV}(u_y(s))$:

3.1. $u \in \textsc{Vars}$: Then $u = y$ and consequently $u_y(s)_x(t) = s_x(t) = u_y(s_x(t)) \overset{x\neq y}{=} u_x(t)_y(s_x(t))$.

3.2. $u = \mathsf{Q}zA$: Since $y \in \mathrm{FV}(u)$, we have $z \neq y$ and $u_y(s) = \mathsf{Q}zA_y(s)$. Using $x \in \mathrm{FV}(u_y(s))$ we further get $z \neq x$ and $u_y(s)_x(t) = \mathsf{Q}zA_y(s)_x(t)$. From $z \neq x, y$ it follows by (c) that $u_x(t)_y(s_x(t)) = \mathsf{Q}zA_x(t)_y(s_x(t))$. From $\mathsf{subst}(u, y, s)$ & $y \in \mathrm{FV}(u)$ we get $\mathsf{subst}(A, y, s)$ and therefore, $A_y(s)_x(t) = A_x(t)_y(s_x(t))$ by I.H.

§2 The completeness theorem of 1st-order predicate logic

In this section we assume that $\neg$, $\wedge$, $\vee$, $\exists$ are defined in terms of $\bot$, $\to$, $\forall$:
$\neg A := A \to \bot$, $A \wedge B := \neg(A \to \neg B)$, $A \vee B := A \to \neg B$, $\exists x A := \neg\forall x \neg A$.

Further we assume that some arbitrary language $\mathcal{L}$ is fixed.

The calculus $\mathcal{K}_{\{\to,\forall,\bot\}}$

Logical axioms:

All formulas of the form $\forall x_1 ... \forall x_m F$, where $m \geq 0$ and F is one of the following:

($\to$ 1) $A \to A$

($\to$ 2) $A \to (B \to A)$

($\to$ 3) $(C \to (A \to B)) \to ((C \to A) \to (C \to B))$

($\to$ 4) $\neg\neg A \to A$, where A is atomic

($\forall$1) $\forall x A \to A_x(t)$ with $\mathsf{subst}(A, x, t)$

($\forall$2) $\forall x(A \to B) \to (\forall x A \to \forall x B)$

($\forall$3) $A \to \forall x A$ with $x \notin \mathrm{FV}(A)$

(G1) $t \approx t$

(G2) $x \approx y \to (A \to A_x(y))$, where A is atomic.

The inference rule "modus ponens"

(mp) A , $A \to B \vdash B$

Definitions

AX := set of all logical axioms(of the calculus $\mathcal{K}_{\{\to,\forall,\bot\}}$).

A *derivation of* A *from* Γ (in the calculus $\mathcal{K}_{\{\to,\forall,\bot\}}$) is a finite sequence of formulas $(A_i)_{i\leq n}$, such that the following holds:

(i) $A_n = A$,

(ii) For each $k \leq n$ we have $A_k \in \mathsf{AX} \cup \Gamma$ oder $A_j = A_i \to A_k$ for some $i, j < k$.

$\Gamma \vdash A :\Leftrightarrow$ A is *derivable* (or *provable*) from $\Gamma :\Leftrightarrow$ there exists a derivation of A from Γ.

$\vdash A \;\; :\Leftrightarrow$ A is *derivable* (or *provable*) $:\Leftrightarrow$ $\emptyset \vdash A$.

Proposition

(a) $\Gamma_0 \vdash A$ & $\Gamma_1 \vdash A \to B \;\Rightarrow\; \Gamma_0 \cup \Gamma_1 \vdash B$

(b) $\Gamma_0 \vdash A$ & $\Gamma_0 \subseteq \Gamma \;\Rightarrow\; \Gamma \vdash A$

(c) $\Gamma \vdash A \;\Rightarrow$ there exists a finite subset $\Gamma_0 \subseteq \Gamma$ such that $\Gamma_0 \vdash A$.

(d) $\vdash s \approx t \to (A_x(s) \to A_x(t))$, for atomic A.

Proof of(d): Choose $x, y \notin \mathrm{FV}(s, t)$ with $x \neq y$ and $y \notin \mathrm{FV}(A)$. Then $\forall x \forall y(x \approx y \to (A \to A_x(y)))$ is an axiom. From this we obtain $s \approx t \to (A_x(s) \to A_x(y)_y(t))$ by ($\forall$1) and (mp). But $A_x(y)_y(t) = A_x(t)$.

Remark

Let Γ be a set of formulas. The set of all formulas which are derivable from Γ is the smallest set X such that $\mathsf{AX} \cup \Gamma \subseteq X$ and X is closed under "modus ponens", i.e., whenever $A \in X$ and $A \to B \in X$ then also $B \in X$. Hence, for proving that a set X contains all formulas derivable from Γ it suffices to prove that $\mathsf{AX} \cup \Gamma \subseteq X$ and X is closed under "modus ponens". *(Induction on derivations).*

Theorem 2.1 (Soundness Theorem)

$\Gamma \vdash A \implies \Gamma \models A$

Proof:

HS 1: If A is one of the formulas listed under $(\to 1), ..., (\mathrm{G}2)$, then $\models A$.

Proof: For $(\to 1), \ldots, (\to 4), (G1)$ the assertion is trivial.

Now let $(\mathcal{M}, \eta)$ be an arbitrary $\mathcal{L}$-interpretation.

$(\forall 1)$: Let $a := [\![t]\!]^{\mathcal{M}}_{\eta}$ and assume $[\![\forall x A]\!]^{\mathcal{M}}_{\eta} = 1$. Then $[\![A_x(t)]\!]^{\mathcal{M}}_{\eta} \overset{\text{L.1.3}}{=} [\![A]\!]^{\mathcal{M}}_{\eta^a_x} = 1$.

$(\forall 2)$: Assume $[\![\forall x(A \to B)]\!]^{\mathcal{M}}_{\eta} = [\![\forall x A]\!]^{\mathcal{M}}_{\eta} = 1$. Then $[\![A \to B]\!]^{\mathcal{M}}_{\eta^a_x} = 1 = [\![A]\!]^{\mathcal{M}}_{\eta^a_x}$ for all $a \in |\mathcal{M}|$.

Hence $[\![B]\!]^{\mathcal{M}}_{\eta^a_x} = 1$ for all $a \in |\mathcal{M}|$, i.e. $[\![\forall x B]\!]^{\mathcal{M}}_{\eta} = 1$.

$(\forall 3)$: From $[\![A]\!]^{\mathcal{M}}_{\eta} = 1$ and $x \notin \mathrm{FV}(A)$ we obtain by L.1.2 $[\![A]\!]^{\mathcal{M}}_{\eta^a_x} = 1$ for all $a \in |\mathcal{M}|$, i.e. $[\![\forall x A]\!]^{\mathcal{M}}_{\eta} = 1$.

(G2): Let $b := \eta(y)$ and assume $[\![x \approx y]\!]^{\mathcal{M}}_{\eta} = 1$. Then $\eta^b_x = \eta$ and thus $[\![A_x(y)]\!]^{\mathcal{M}}_{\eta} \overset{\text{L.1.3}}{=} [\![A]\!]^{\mathcal{M}}_{\eta^b_x} = [\![A]\!]^{\mathcal{M}}_{\eta}$.

HS 2: If A is a logical axiom, then $\models A$.

Proof: Let $A = \forall x_1 ... \forall x_m F$, where F does not start with $\forall$. Let $\mathcal{M}$ be an arbitrary $\mathcal{L}$-structure.

By induction on m we prove $\mathcal{M} \models A[\eta]$ for each $\mathcal{M}$-assignment η.

1. $m = 0$: HS1.

2. $m > 0$: Then also $B := \forall x_2 ... \forall x_m F$ is a logical axiom and by I.H. we have $\mathcal{M} \models B[\eta^a_{x_1}]$ for all η and all $a \in |\mathcal{M}|$. Hence $\mathcal{M} \models (\forall x_1 B)[\eta]$, i.e. $\mathcal{M} \models A[\eta]$, for all η.

The theorem now follows from HS 2 by induction on the length of the derivation of A from Γ.

Theorem 2.2 (Deduction Theorem)

$\Gamma \cup \{C\} \vdash B \implies \Gamma \vdash C \to B$.

Proof by induction on derivations:

1. $B \in \mathsf{AX} \cup \Gamma$: Then $(B\,,\, B \to (C \to B)\,,\, C \to B)$ is a derivation of $C \to B$ from Γ.

2. $B = C$: Then $C \to B$ is a logical axiom.

3. $\Gamma \cup \{C\} \vdash A$ and $\Gamma \cup \{C\} \vdash A \to B$:

In this case we have $\Gamma \vdash C \to (A \to B)$ and $\Gamma \vdash C \to A$ by I.H.

From this we obtain $\Gamma \vdash C \to B$ by $(\to 3)$ and (mp). $\triangle$

Corollary

$\Gamma_0 \vdash A \;\&\; \Gamma_1 \cup \{A\} \vdash B \implies \Gamma_0 \cup \Gamma_1 \vdash B$.

Lemma 2.3

(a) $\Gamma \vdash A \;\&\; x \notin \mathrm{FV}(\Gamma) \implies \Gamma \vdash \forall x A$.

(b) $y \notin \mathrm{FV}(A) \,\&\, \mathsf{subst}(A, x, y) \implies \vdash \forall y A_x(y) \to \forall x A$.

(c) $\Gamma \vdash A_x(y) \,\&\, y \notin \mathrm{FV}(\Gamma, \forall x A) \,\&\, \mathsf{subst}(A, x, y) \implies \Gamma \vdash \forall x A$.

Proof:

(a) Induction on derivations:

1. $A \in \Gamma$: Then $x \notin \mathrm{FV}(A)$, and $A \to \forall x A$ is a logical axiom.

2. $A \in \mathsf{AX}$: Then also $\forall x A \in \mathsf{AX}$.

3. $\Gamma \vdash C \;\&\; \Gamma \vdash C \to A$: I.H. $\Rightarrow \Gamma \vdash \forall x C \;\&\; \Gamma \vdash \forall x (C \to A) \overset{(\forall 2)}{\Longrightarrow} \Gamma \vdash \forall x A$.

(b) From the premise by L.1.4e,d we obtain $A_x(y)_y(x) = A_x(x) = A$ and $\mathsf{subst}(A_x(y), y, x)$.

Therefore $\forall y A_x(y) \to A$ is an axiom ($\forall 1$), and we have $\{\forall y A_x(y)\} \vdash A$.

From this the assertion follows by (a), L.1.4b and Theorem 2.2.

(c) For $y = x$ the claim follows by L.1.4d and (a). Now let $y \neq x$. Then from the premise by (a) we obtain $\Gamma \vdash \forall y A_x(y) \;\;\&\;\; y \notin \mathrm{FV}(A) \;\&\; \mathsf{subst}(A, x, y)$. Hence $\Gamma \vdash \forall x A$ by (b). $\triangle$

Lemma 2.4

If the constant c does not occur in Γ, then:

(a) $\Gamma \vdash A_x(c) \implies \Gamma \vdash \forall x A$, if c does not occur in A.

(b) $\Gamma \vdash \bot \implies$ there exists a derivation of $\bot$ from Γ, in which c does not occur.

Proof:

(a) For any formula B and variable z let $B(c/z)$ denote the result of replacing in B all occurrences of c by z.

HS: If B is a logical axiom and $z \notin \mathrm{vars}(B)$, then also $B(c/z)$ is a logical axiom.

Not let $H = \big(B_i\big)_{i \leq n}$ be a derivation of $A_x(c)$ from Γ with c not occurring in A. W.o.l.g. Γ is finite. We choose a variable z not occurring in H, Γ, A. From HS it follows that $\big(B_i(c/z)\big)_{i \leq n}$ is a derivation of $A_x(c)(c/z)$ from Γ. Further $A_x(c)(c/z) = A_x(z)$. Hence $\Gamma \vdash A_x(z)$ and $z \notin \mathrm{vars}(\Gamma, A)$. By L.2.3c from this we get $\Gamma \vdash \forall x A$. (Note that $z \notin \mathrm{BV}(A)$ implies $\mathsf{subst}(A, x, z)$.)

(b) If $H = \big(B_i\big)_{i \leq n}$ is a derivation of $\bot$ from Γ and z is a variable not occurring in H, then $\big(B_i(c/z)\big)_{i \leq n}$ is a derivation of $\bot$ from Γ in which c does not occur.

Lemma 2.5

(a) $\vdash \neg\neg A \to A$,

(b) $\Gamma \cup \{\neg A\} \vdash \bot \;\Rightarrow\; \Gamma \vdash A$,

(c) $\Gamma \vdash \bot \;\Rightarrow\; \Gamma \vdash A$,

(d) $\vdash \exists x (x \approx s)$, if $x \notin \mathrm{FV}(s)$.

Proof:

(a) Proof by induction on the build-up of A:

For atomic A the formula $\neg\neg A \to A$ is a logical axiom. — Assume now $\vdash \neg\neg B \to B$.

$$\dfrac{\dfrac{A \quad A\to B}{B}\quad \neg B}{\dfrac{\bot}{\dfrac{\neg(A\to B)\quad \neg\neg(A\to B)}{\dfrac{\bot}{\dfrac{\neg\neg B}{\dfrac{B}{A\to B}}}\quad \neg\neg B\to B}}}$$

$$\dfrac{\dfrac{\forall x B}{B}\quad \neg B}{\dfrac{\bot}{\dfrac{\neg\forall x B\quad \neg\neg\forall x B}{\dfrac{\bot}{\dfrac{\neg\neg B}{\dfrac{B}{\forall x B}}}\quad \neg\neg B\to B}}}$$

(b) $\Gamma \cup \{\neg A\} \vdash \bot \overset{2.2}{\Rightarrow} \Gamma \vdash \neg\neg A \overset{(a)}{\Rightarrow} \Gamma \vdash A$.

(c) $\Gamma \vdash \bot \;\Rightarrow\; \Gamma \cup \{\neg A\} \vdash \bot \overset{(b)}{\Rightarrow} \Gamma \vdash A$.

(d) Assume $x \notin \mathrm{FV}(s)$. Then $\forall x \neg(x \approx s) \to \neg(s \approx s)$ is an axiom ($\forall$1), and thus $\forall x \neg(x \approx s) \vdash \neg(s \approx s)$ holds. Using (G1) and (mp) we then obtain $\forall x \neg(x \approx s) \vdash \bot$ and finally (by Theorem 2.2) the assertion $\vdash \neg\forall x \neg(x \approx s)$. $\triangle$

Definition. Γ is *consistent* $:\Leftrightarrow \Gamma \nvdash \bot$.

Theorem 2.6

Let T be the set of all Terms and $\mathrm{T}_0 := \{t \in \mathrm{T} : \mathrm{FV}(t) = \emptyset\}$.

Further let Σ be a *consistent* set of formulas, such that:

(i) $A \notin \Sigma \;\Rightarrow\; \Sigma \cup \{A\} \vdash \bot$,

(ii) $\neg\forall x A \in \Sigma \;\Rightarrow\; \neg A_x(t) \in \Sigma$, for some $t \in \mathrm{T}_0$.

We define:

$\sim := \{(s,t) : (s \approx t) \in \Sigma\}$, $\overline{t} := \{s \in \mathrm{T} : s \sim t\}$

$\mathcal{M} := \big(\{\overline{t} : t \in \mathrm{T}\}, (p^{\mathcal{M}})_{p\in\mathcal{L}}\big)$, where

$$p^{\mathcal{M}} := \begin{cases} \overline{p} & \text{if } p \text{ is a constant} \\ \{((\overline{t_1},\ldots,\overline{t_n}), \overline{pt_1\ldots t_n}) : t_1,\ldots,t_n \in \mathrm{T}\} & \text{if } p \text{ is an } n\text{-ary fuction symbol } (n \geq 1) \\ \{(\overline{t_1},\ldots,\overline{t_n}) : pt_1\ldots t_n \in \Sigma\} & \text{if } p \text{ is a relation symbol} \end{cases}$$

$\xi : \mathrm{VARS} \to |\mathcal{M}|$, $x \mapsto \overline{x}$.

Then $\mathcal{M}$ is an $\mathcal{L}$-structure and the following holds for all t and A:

(a) $[\![t]\!]^{\mathcal{M}}_{\xi} = \overline{t}$,

(b) $\mathcal{M} \models A[\xi] \;\Leftrightarrow\; A \in \Sigma$.

Proof:

HS 1

(a) $\Sigma \vdash A \;\Leftrightarrow\; A \in \Sigma$.

(b) $A\to B \in \Sigma \;\Leftrightarrow\; A \notin \Sigma$ or $B \in \Sigma$

(c) $\forall x A \in \Sigma \;\Leftrightarrow\; A_x(t) \in \Sigma$ for all $t \in \mathrm{T}_0$.

(d) $s \sim t$ & $A_x(s) \in \Sigma$ & A atomic $\;\Rightarrow\; A_x(t) \in \Sigma$.

(e) $\sim$ is an equivalence relation on T.

(f) For each $s \in \mathrm{T}$ there exists a $t \in \mathrm{T}_0$ such that $\overline{s} = \overline{t}$.

Proof:

(a) "⇐": trivial.

"⇒": $\Sigma \vdash A \,\&\, A \notin \Sigma \overset{\text{(i)}}{\Rightarrow} \Sigma \vdash A \,\&\, \Sigma \cup \{A\} \vdash \bot \Rightarrow \Sigma \vdash \bot$. *Contradiction.*

(b) $A{\to}B \in \Sigma \,\&\, A \in \Sigma \Rightarrow \Sigma \vdash B \Rightarrow B \in \Sigma$.

$B \in \Sigma \Rightarrow \Sigma \cup \{A\} \vdash B \Rightarrow \Sigma \vdash A \to B$.

$A \notin \Sigma \overset{\text{(i)}}{\Rightarrow} \Sigma \cup \{A\} \vdash \bot \overset{\text{2.5c}}{\Rightarrow} \Sigma \cup \{A\} \vdash B \Rightarrow \Sigma \vdash A \to B$.

(c) $\forall xA \in \Sigma \overset{(\forall 1)}{\Rightarrow} \Sigma \vdash A_x(t)$ for all $t \in \mathrm{T}_0$.

$A_x(t) \in \Sigma\ (\forall t \in \mathrm{T}_0) \overset{\Sigma \text{ cons.}}{\Rightarrow} \neg A_x(t) \notin \Sigma\ (\forall t \in \mathrm{T}_0) \overset{\text{(ii)}}{\Rightarrow} \neg\forall xA \notin \Sigma \overset{\text{(i)}}{\Rightarrow} \Sigma \cup \{\neg\forall xA\} \vdash \bot \overset{\text{2.5b}}{\Rightarrow} \Sigma \vdash \forall xA$.

(d) From $s \approx t \in \Sigma \,\&\, A_x(s) \in \Sigma$ one obtains $\Sigma \vdash A_x(t)$ by Prop.(d) above.

(e) (G1) $\Rightarrow \Sigma \vdash t \approx t \overset{\text{(a)}}{\Rightarrow} t \sim t$. $\quad s \sim t \,\&\, \Sigma \vdash s \approx s \overset{\text{(a),(d)}}{\Rightarrow} t \approx s \in \Sigma \Rightarrow t \sim s$.

$s \sim t \,\&\, r \sim s \Rightarrow s \sim t \,\&\, r \approx s \in \Sigma \overset{\text{(d)}}{\Rightarrow} r \approx t \in \Sigma \Rightarrow r \sim t$.

(f) Let $s \in \mathrm{T}$. We choose a variable $x \notin \mathrm{FV}(s)$. By (a) and L.2.5d we have $\exists x(x \approx s) \in \Sigma$. By (ii) this yields the existence of a $t \in \mathrm{T}_0$ such that $\neg\neg(t \approx s) \in \Sigma$. By (a) and $(\to 4)$ we then obtain $t \approx s \in \Sigma$, i.e. $t \sim s$ and thus $\overline{t} = \overline{s}$, since $\sim$ is an equivalence relation.

HS 2: For each n-ary function symbol f and n-ary relation symbol R we have:

(a) $f^{\mathcal{M}} : |\mathcal{M}|^n \to |\mathcal{M}|$, $f^{\mathcal{M}}(\overline{t_1}, ..., \overline{t_n}) = \overline{ft_1...t_n}$.

(b) $Rt_1...t_n \in \Sigma \Leftrightarrow (\overline{t_1}, ..., \overline{t_n}) \in R^{\mathcal{M}}$.

Proof:

(a) Let $\overline{s_i} = \overline{t_i}$ $(i = 1, ..., n)$. We have to prove: $\overline{fs_1...s_n} = \overline{ft_1...t_n}$, i.e. $(fs_1...s_n \approx ft_1...t_n) \in \Sigma$.

By HS1(a) we have $(fs_1...s_n \approx fs_1...s_n) \in \Sigma$.

From this together with $s_1 \sim t_1, \ldots, s_n \sim t_n$ and HS1(d) we successively obtain $(fs_1...s_n \approx ft_1s_2...s_n) \in \Sigma$, $(fs_1...s_n \approx ft_1t_2s_3...s_n) \in \Sigma$, $\ldots$, $(fs_1...s_n \approx ft_1...t_n) \in \Sigma$, hence $\overline{fs_1...s_n} = \overline{ft_1...t_n}$.

(b) "⇒": By definition.

"⇐": Assume $(\overline{t_1}, ..., \overline{t_n}) \in R^{\mathcal{M}}$. Then there are $s_1, ..., s_n$ with $Rs_1...s_n \in \Sigma$ and $\overline{s_i} = \overline{t_i}$ $(i = 1, ..., n)$. As above we conclude by HS1(d) $Rt_1s_2...s_n \in \Sigma, Rt_1t_2s_3...s_n \in \Sigma, \ldots, Rt_1...t_n \in \Sigma$, i.e., $\overline{Rt_1...t_n} = 1 = \overline{Rs_1...s_n}$.

Proof of 2.6(a) by induction on t:

1. $[\![x]\!]^{\mathcal{M}}_{\xi} = \xi(x) = \overline{x}$.
2. $[\![ft_1...t_n]\!]^{\mathcal{M}}_{\xi} = f^{\mathcal{M}}([\![t_1]\!]^{\mathcal{M}}_{\xi}, \ldots, [\![t_n]\!]^{\mathcal{M}}_{\xi}) \overset{\text{IH}}{=} f^{\mathcal{M}}(\overline{t_1}, \ldots, \overline{t_n}) \overset{\text{HS2a}}{=} \overline{ft_1...t_n}$.

Proof of 2.6(b) by induction on A:

1. $\mathcal{M} \not\models \bot$ and $\bot \notin \Sigma$.
2. $\mathcal{M} \models (s \approx t)[\xi] \Leftrightarrow [\![s]\!]^{\mathcal{M}}_{\xi} = [\![t]\!]^{\mathcal{M}}_{\xi} \overset{\text{(a)}}{\Leftrightarrow} \overline{s} = \overline{t} \Leftrightarrow s \sim t \Leftrightarrow s \approx t \in \Sigma$.
3. $\mathcal{M} \models (Rt_1...t_n)[\xi] \Leftrightarrow ([\![t_1]\!]^{\mathcal{M}}_{\xi}, ..., [\![t_n]\!]^{\mathcal{M}}_{\xi}) \in R^{\mathcal{M}} \overset{\text{(a)}}{\Leftrightarrow} (\overline{t_1}, ..., \overline{t_n}) \in R^{\mathcal{M}} \overset{\text{HS2b}}{\Leftrightarrow} Rt_1...t_n \in \Sigma$.
4. $\mathcal{M} \models (A \to B)[\xi] \Leftrightarrow \mathcal{M} \not\models A[\xi]$ or $\mathcal{M} \models B[\xi] \overset{\text{IH}}{\Leftrightarrow} A \notin \Sigma$ or $B \in \Sigma \overset{\text{HS1b}}{\Leftrightarrow} (A \to B) \in \Sigma$.
5. By (a) and Lemma 1.3 (Corollary) we get:

$\mathcal{M} \models (\forall xA)[\xi] \overset{\text{HS1f}}{\Leftrightarrow} \mathcal{M} \models A[\xi^{\overline{t}}_x]$ for all $t \in \mathrm{T}_0 \overset{\text{(a)+L.1.3}}{\Leftrightarrow} \mathcal{M} \models A_x(t)[\xi]$ for all $t \in \mathrm{T}_0 \overset{\text{IH}}{\Leftrightarrow}$

$\Leftrightarrow A_x(t) \in \Sigma$ for all $t \in \mathrm{T}_0 \overset{\text{HS1c}}{\Leftrightarrow} \forall xA \in \Sigma$.

Definition

A set of formulas Σ which satisfies the assumptions (i),(ii) of Theorem 2.6, is called a *complete Henkin theory*. The interpretation $(\mathcal{M}, \xi)$ defined in 2.6 is called *the canonical term model of* Σ.

Satz 2.7 (Completeness Theorem)

Every consistent set of formulas is satisfiable.

Corollary

(a) Γconsistent $\Leftrightarrow$ Γsatisfiable.

(b) $\Gamma \vdash A \;\Leftrightarrow\; \Gamma \models A$.

(c) $\Gamma \models A \;\Rightarrow$ there exists a *finite* $\Gamma_0 \subseteq \Gamma$ such that $\Gamma_0 \models A$.

Proof of the corollary:

(a) "$\Leftarrow$": Assume $\mathcal{M} \models \Gamma[\eta]$. By the soundness theorem we then have $(\Gamma \vdash A \;\Rightarrow\; \mathcal{M} \models A[\eta])$, for each formula A. Hence $\Gamma \nvdash \bot$.

(b) "$\Leftarrow$": $\Gamma \nvdash A \overset{2.5b}{\Longrightarrow} \Gamma \cup \{\neg A\}$ consistent $\Rightarrow \Gamma \cup \{\neg A\}$ satisfiable $\Rightarrow \Gamma \not\models A$.

(c) This follows from (b) and the fact that every derivation is finite.

Proof of the Theorem:

Let Γ be a consistent set of formulas.

First the proof is carried through under the following *additional assumptions*:

– $\mathcal{L}$ countable,

– $\mathcal{L}$ contains infinitely many constants $c_0, c_1, c_2, \ldots$, which do not occur in Γ.

Let $A_0, A_1, A_2, \ldots$ be an enumeration of all $\mathcal{L}$-formulas.

Definition:

$\Sigma_0 := \Gamma$

$$\Sigma_{n+1} := \begin{cases} \Sigma_n & \text{if } \Sigma_n \cup \{A_n\} \vdash \bot \\ \Sigma_n \cup \{A_n\} & \text{if } \Sigma_n \cup \{A_n\} \nvdash \bot \;\&\; A_n \neq \neg\forall x B \\ \Sigma_n \cup \{A_n, \neg B_x(c_k)\} & \text{if } \Sigma_n \cup \{A_n\} \nvdash \bot \;\&\; A_n = \neg\forall x B \\ \quad \text{where } k := \min\{i : c_i \text{ does not occur in any formula from } \Sigma_n \cup \{B\}\} \end{cases}.$$

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

HS 1: $\Sigma_n \nvdash \bot$.

Proof by induction on n:

1. $n = 0$: trivial.

2. $n \to n+1$:

2.1. $\Sigma_{n+1} = \Sigma_n$: Immediately by IH.

2.2. $\Sigma_{n+1} = \Sigma_n \cup \{A_n\}$ with $\Sigma_n \cup \{A_n\} \nvdash \bot$: trivial.

2.3. $\Sigma_{n+1} = \Sigma_n \cup \{\neg\forall x B, \neg B_x(c_k)\}$ with $\Sigma_n \cup \{\neg\forall x B\} \nvdash \bot$ and c_k not in $\Sigma_n \cup \{B\}$:

$\Sigma_n \cup \{\neg\forall x B, \neg B_x(c_k)\} \vdash \bot \overset{2.5b}{\Longrightarrow} \Sigma_n \cup \{\neg\forall x B\} \vdash B_x(c_k) \overset{2.4a}{\Longrightarrow} \Sigma_n \cup \{\neg\forall x B\} \vdash \forall x B \;\Rightarrow$

$\Rightarrow \Sigma_n \cup \{\neg\forall x B\} \vdash \bot$. *Contradiction.*

Since $\Sigma \vdash \bot$ implies $\Sigma_n \vdash \bot$ for some n, the consistency of Σ follows from HS1.

We now prove that Σ is a complete Henkin theory.

(i) If $A \notin \Sigma$ then, for some n, $A = A_n \notin \Sigma_{n+1}$ and therefore $\Sigma_n \cup \{A_n\} \vdash \bot$; hence $\Sigma \cup \{A\} \vdash \bot$.
(ii) If $\neg\forall x B = A_n \in \Sigma$ then $\Sigma_n \cup \{A_n\} \nvdash \bot$ (since $\Sigma \nvdash \bot$) and thus $\neg B_x(c_k) \in \Sigma$ for some $k \in \mathbb{N}$.

By Theorem 2.6 there exists a model $(\mathcal{M}, \xi)$ of Σ. Since $\Gamma \subseteq \Sigma$, $(\mathcal{M}, \xi)$ is also a model of Γ.

Now we discharge the second additional assumption and prove the theorem for arbitrary countable $\mathcal{L}$. The proof for uncountable $\mathcal{L}$ will be given later.

Let $c_0, c_1, c_2, \ldots$ be a countably infinite sequence of *new* constants, and let $\mathcal{L}' := \mathcal{L} \cup \{c_0, c_1, c_2, \ldots\}$. Then also $\mathcal{L}'$ is countable, and we can carry out the above construction with $\mathcal{L}'$ in place of $\mathcal{L}$. For that we have still to prove that Γ is consistent with respect to the extended language $\mathcal{L}'$, i.e., that there is no $\mathcal{L}'$-derivation of $\bot$ from Γ. To this end we set $\mathcal{L}_k := \mathcal{L} \cup \{c_0, \ldots, c_{k-1}\}$ and $\mathcal{H}_k :=$ set of all $\mathcal{L}_k$-derivations of $\bot$ from Γ.

By 2.4b (with $\mathcal{L}_{k+1}$ in place of $\mathcal{L}$) we obtain: (*) $\mathcal{H}_{k+1} \neq \emptyset \;\Rightarrow\; \mathcal{H}_k \neq \emptyset$.

Now if H were an $\mathcal{L}'$-derivation of $\bot$ from Γ, then we would have $H \in \mathcal{H}_{k_0}$ for some k_0. Together with (*) this would yield $\mathcal{H}_0 \neq \emptyset$, contradicting the assumption that Γ is consistent w.r.t. $\mathcal{L}$.

As we have proved above (with $\mathcal{L}$ in place of $\mathcal{L}'$), there exists an $\mathcal{L}'$-Modell $(\mathcal{M}', \xi)$ of Γ. Let $\mathcal{M}$ be the $\mathcal{L}$-structure which one obtains by restricting $\mathcal{M}'$ to $\mathcal{L}$. Then $(\mathcal{M}, \xi) \models \Gamma$, since Γ is a set of $\mathcal{L}$-formulas.

Theorem 2.8 (Compactness Theorem)
If every finite subset of Γ is satisfiable, then Γ is satisfiable.

Proof:
Γ not satisfiable $\Rightarrow$ $\Gamma \models \bot$ $\overset{\text{Corollary (c)}}{\Rightarrow}$ there exists a finite $\Gamma_0 \subseteq \Gamma$ such that $\Gamma_0 \models \bot$ $\Rightarrow$ Γ_0 not satisfiable.

§3 Elements of model theory

Definition

For each set of formulas Γ let $\mathrm{L}(\Gamma)$ denote the set of all function and relation symbols occurring in Γ.

Definition

Assume that $\mathcal{L} \subseteq \mathcal{L}'$, $\mathcal{M}$ is an $\mathcal{L}$-structure and $\mathcal{M}'$ an $\mathcal{L}'$-structure:
$\mathcal{M}'$ is an *expansion* of $\mathcal{M}$ $:\Leftrightarrow$ $\mathcal{M}$ is a *reduct* of $\mathcal{M}'$ $:\Leftrightarrow$ $|\mathcal{M}| = |\mathcal{M}'|$ and $p^{\mathcal{M}'} = p^{\mathcal{M}}$ for all $p \in \mathcal{L}$.

$\mathcal{M}'{\restriction}\mathcal{L}$ denotes the uniquely determined $\mathcal{L}$-reduct of $\mathcal{M}'$.

Remark. If $\mathcal{M}'$ is an expansion of the $\mathcal{L}$-structure $\mathcal{M}$, and ξ an $\mathcal{M}$-assignment, then $[\![u]\!]^{\mathcal{M}}_{\xi} = [\![u]\!]^{\mathcal{M}'}_{\xi}$ holds for each $\mathcal{L}$-expression u. Therefore the statements "$\Gamma \models A$" and "Γ satisfiable" do not depend on the language $\mathcal{L}$ which appears in the corresponding definitions (provided $\mathcal{L} \supseteq \mathrm{L}(\Gamma \cup \{A\})$). Due to Theorem 2.7 the same holds for "$\Gamma \vdash A$" and "Γ consistent".

Definition. In the sequel by an *axiom system* we always mean a set of closed formulas (sentences).

Definition. A structure $\mathcal{M}$ is called *finite* (*infinite*, *countable*) if its universe $|\mathcal{M}|$ has this property.

Theorem 3.1 (Löwenheim-Skolem)
Every countable, satisfiable set of formulas Γ has a countable model.

Proof:
Let $\mathcal{L}_0 := \mathrm{L}(\Gamma)$ and $\mathcal{L} := \mathcal{L}_0 \cup \{c_i : i \in \mathbb{N}\}$ where $c_0, c_1, \ldots$ are new constants. Then $\mathcal{L}$ is countable and, as shown in the proof of 2.7, Γ can be extended to a complete Henkin theory Σ (in $\mathcal{L}$). In the proof of 2.6 it was shown that Σ has a model $(\mathcal{M}, \xi)$ with $|\mathcal{M}| = \{\overline{t} : t \in T\}$ where T is the set of all $\mathcal{L}$-terms. Since $\mathcal{L}$ is countable, also T and $|\mathcal{M}|$ are countable. Hence Σ (and also $\Gamma \subseteq \Sigma$) has a countable model.

Theorem 3.2
If the axiom system Σ has models of cardinality $\geq n$ for each $n \in \mathbb{N}$ then Σ also has an infinite model.

Proof:
Let $\mathcal{L} := \mathrm{L}(\Sigma)$ and $\Gamma := \Sigma \cup \{\neg(v_i \approx v_j) : i, j \in \mathbb{N} \text{ with } i < j\}$.

HS: Every finite subset of Γ ist satisfiable.

Proof: Let $\Delta \subseteq \Gamma$ be finite. Then there exists an n such that $\Delta \subseteq \Sigma \cup \{\neg(v_i \approx v_j) : i < j \leq n\}$.

By assumption there exists a model $\mathcal{M}$ of Σ with pairwise distinct elements $a_0, \ldots, a_n \in |\mathcal{M}|$. Let η be an $\mathcal{M}$-assignment with $\eta(v_0) = a_0, \ldots, \eta(v_n) = a_n$. Then $(\mathcal{M}, \eta)$ is a model of Δ.

From HS and Theorem 2.8 it follows that Γ is satisfiable. Let $(\mathcal{M}, \xi)$ be a model of Γ. Then $\mathcal{M}$ is a model of Σ, and $i \mapsto \xi(v_i)$ is a one-one mapping from $\mathbb{N}$ into $|\mathcal{M}|$; so $\mathcal{M}$ is infinite.

Definitions
- $\mathrm{For}^0_{\mathcal{L}}$ denotes the set of all $\mathcal{L}$-sentences.
- An axiom system T is called a *Theory*, if it is *deductively closed*, i.e., if $T = \{A \in \mathrm{For}^0_{\mathrm{L}(T)} : T \vdash A\}$.
- If T is a theory and $\Sigma \subseteq T$ with $T = \{A \in \mathrm{For}^0_{\mathrm{L}(T)} : \Sigma \vdash A\}$, then Σ is called an *axiom system of* T.

- An axiom system Σ is *complete*, if $\Sigma \vdash A$ or $\Sigma \vdash \neg A$ holds for each $\mathrm{L}(\Sigma)$-sentence A.
- For each $\mathcal{L}$-structure $\mathcal{M}$ let $\mathrm{Th}(\mathcal{M}) := \{A \in \mathrm{For}^0_{\mathcal{L}} : \mathcal{M} \models A\}$ (*the theory of* $\mathcal{M}$).
- $\mathrm{Mod}_{\mathcal{L}}(\Sigma) := \{\mathcal{M} : \mathcal{M} \text{ is } \mathcal{L}\text{-structure and } \mathcal{M} \models \Sigma\}$ (Σ an axiom system with $\mathrm{L}(\Sigma) \subseteq \mathcal{L}$)
- The $\mathcal{L}$-structures $\mathcal{M}$, $\mathcal{M}'$ are *elementary equivalent* (in symbols $\mathcal{M} \equiv \mathcal{M}'$), if $\mathrm{Th}(\mathcal{M}) = \mathrm{Th}(\mathcal{M}')$.
- The $\mathcal{L}$-structures $\mathcal{M}$, $\mathcal{M}'$ are *isomorphic* (in symbols $\mathcal{M} \cong \mathcal{M}'$),
  if there exists an isomorphism π from $\mathcal{M}$ onto $\mathcal{M}'$.
  An *isomorphism* $\pi : \mathcal{M} \to \mathcal{M}'$ is a bijective mapping $\pi : |\mathcal{M}| \to |\mathcal{M}'|$ such that
  (i) $\pi(f^{\mathcal{M}}(a_1, ..., a_n)) = f^{\mathcal{M}'}(\pi(a_1), ..., \pi(a_n))$ $\quad (f \in \mathcal{L}, \#(f) = n \geq 0, a_1, ..., a_n \in |\mathcal{M}|)$
  (ii) $(a_1, ..., a_n) \in R^{\mathcal{M}} \Leftrightarrow (\pi(a_1), ..., \pi(a_n)) \in R^{\mathcal{M}'}$ $\quad (R \in \mathcal{L}, \#(R) = n \geq 1, a_1, ..., a_n \in |\mathcal{M}|)$

Lemma 3.3

(a) $\mathrm{Th}(\mathcal{M})$ is a complete theory.

(b) If Σ is an axiom system with $\mathrm{L}(\Sigma) \subseteq \mathcal{L}$, then
$\{A \in \mathrm{For}^0_{\mathcal{L}} : \Sigma \vdash A\} = \bigcap\{\mathrm{Th}(\mathcal{M}) : \mathcal{M} \in \mathrm{Mod}_{\mathcal{L}}(\Sigma)\}$ is a theory.

(c) $\mathcal{M} \equiv \mathcal{M}' \iff \mathcal{M} \models \mathrm{Th}(\mathcal{M}')$

(d) If $\mathcal{L}$ is countable, then for every $\mathcal{L}$-structure $\mathcal{M}$ there is countable structure $\mathcal{M}'$ so that $\mathcal{M} \equiv \mathcal{M}'$.

Proof:

(a) $\mathrm{Th}(\mathcal{M}) \vdash A \Rightarrow \mathcal{M} \models A \Rightarrow A \in \mathrm{Th}(\mathcal{M})$. $\mathrm{Th}(\mathcal{M})$ complete: obvious.

(b) $\Sigma \vdash A \Leftrightarrow \Sigma \models A \Leftrightarrow \mathcal{M} \models A$ $(\forall \mathcal{M} \in \mathrm{Mod}_{\mathcal{L}}(\Sigma)) \Leftrightarrow A \in T := \bigcap\{\mathrm{Th}(\mathcal{M}) : \mathcal{M} \in \mathrm{Mod}_{\mathcal{L}}(\Sigma)\}$.

$T \vdash A \Rightarrow \mathrm{Th}(\mathcal{M}) \vdash A$ $(\forall \mathcal{M} \in \mathrm{Mod}_{\mathcal{L}}(\Sigma)) \Rightarrow A \in \mathrm{Th}(\mathcal{M})$ $(\forall \mathcal{M} \in \mathrm{Mod}_{\mathcal{L}}(\Sigma)) \Rightarrow A \in T$.

(c) "$\Rightarrow$": $\mathrm{Th}(\mathcal{M}) = \mathrm{Th}(\mathcal{M}') \Rightarrow \mathcal{M} \models \mathrm{Th}(\mathcal{M}')$.

"$\Leftarrow$": $\mathcal{M} \models \mathrm{Th}(\mathcal{M}') \Rightarrow \mathrm{Th}(\mathcal{M}') \subseteq \mathrm{Th}(\mathcal{M})$ $(*)$.

$A \in \mathrm{Th}(\mathcal{M}) \Rightarrow \neg A \notin \mathrm{Th}(\mathcal{M}) \overset{(*)}{\Rightarrow} \neg A \notin \mathrm{Th}(\mathcal{M}') \Rightarrow A \in \mathrm{Th}(\mathcal{M}')$.

(d) $\mathrm{Th}(\mathcal{M})$ satsifiable and countable $\overset{3.1}{\Rightarrow}$ there is a countable $\mathcal{M}'$ so that $\mathcal{M}' \models \mathrm{Th}(\mathcal{M})$.

Theorem 3.4

Let T be a theory und $\mathcal{L} = \mathrm{L}(T)$. The following propositions are equivalent:

(i) T is complete.

(ii) $\forall \mathcal{M} \in \mathrm{Mod}_{\mathcal{L}}(T) : \mathrm{Th}(\mathcal{M}) = T$.

(iii) $\forall \mathcal{M}, \mathcal{M}' \in \mathrm{Mod}_{\mathcal{L}}(T) : \mathcal{M} \equiv \mathcal{M}'$.

Proof:

(i)$\Rightarrow$(ii): $\mathcal{M} \models T \Rightarrow T \subseteq \mathrm{Th}(\mathcal{M})$. $A \in \mathrm{Th}(\mathcal{M}) \Rightarrow \neg A \notin \mathrm{Th}(\mathcal{M}) \Rightarrow \neg A \notin T \Rightarrow A \in T$.

(ii)$\Rightarrow$(iii): trivial.

(iii)$\Rightarrow$(i): Let $A \in \mathrm{For}^0_{\mathcal{L}}$ and $A \notin T$. Then $T \nvdash A$, i.e., there exists a model $\mathcal{M}_0$ of $T \cup \{\neg A\}$.
From (iii) we obtain $\mathcal{M} \equiv \mathcal{M}_0$ $(\forall \mathcal{M} \in \mathrm{Mod}_{\mathcal{L}}(T))$ and thus $\mathcal{M} \models \neg A$ $(\forall \mathcal{M} \in \mathrm{Mod}_{\mathcal{L}}(T))$, i.e., $T \vdash \neg A$.

Theorem 3.5

Let $\pi : \mathcal{M} \to \mathcal{M}'$ be an isomorphism, and ξ an $\mathcal{M}$-assignment. Then for each $\mathcal{L}$-term t and each $\mathcal{L}$-Formel A the following holds: (a) $\pi([\![t]\!]^{\mathcal{M}}_{\xi}) = [\![t]\!]^{\mathcal{M}'}_{\pi \circ \xi}$, (b) $\mathcal{M} \models A[\xi] \iff \mathcal{M}' \models A[\pi \circ \xi]$.

For closed t, A we have (a) $\pi([\![t]\!]^{\mathcal{M}}) = [\![t]\!]^{\mathcal{M}'}$, (b) $\mathcal{M} \models A \iff \mathcal{M}' \models A$.

Corollary. $\mathcal{M} \cong \mathcal{M}' \Rightarrow \mathcal{M} \equiv \mathcal{M}'$.

Proof by induction on the build up of t, A:

Abb.: $\xi' := \pi \circ \xi$.

(a) 1. $\pi([\![x]\!]^{\mathcal{M}}_{\xi}) = \pi(\xi(x)) = [\![x]\!]^{\mathcal{M}'}_{\xi'}$.

2. $\pi([\![ft]\!]^{\mathcal{M}}_{\xi}) = \pi(f^{\mathcal{M}}([\![t]\!]^{\mathcal{M}}_{\xi})) = f^{\mathcal{M}'}(\pi([\![t]\!]^{\mathcal{M}}_{\xi})) \overset{\text{IH}}{=} f^{\mathcal{M}'}([\![t]\!]^{\mathcal{M}'}_{\xi'}) = [\![ft]\!]^{\mathcal{M}'}_{\xi'}$.

(b) 1. $[\![s \approx t]\!]^{\mathcal{M}}_{\xi} = 1 \Leftrightarrow [\![s]\!]^{\mathcal{M}}_{\xi} = [\![t]\!]^{\mathcal{M}}_{\xi} \Leftrightarrow \pi([\![s]\!]^{\mathcal{M}}_{\xi}) = \pi([\![t]\!]^{\mathcal{M}}_{\xi}) \overset{\text{(a)}}{\Leftrightarrow} [\![s]\!]^{\mathcal{M}'}_{\xi'} = [\![t]\!]^{\mathcal{M}'}_{\xi'} \Leftrightarrow [\![s \approx t]\!]^{\mathcal{M}'}_{\xi'} = 1$.

2. $[\![Rt]\!]^{\mathcal{M}}_{\xi} = R^{\mathcal{M}}([\![t]\!]^{\mathcal{M}}_{\xi}) = R^{\mathcal{M}'}(\pi([\![t]\!]^{\mathcal{M}}_{\xi})) \overset{\text{(a)}}{=} R^{\mathcal{M}'}([\![t]\!]^{\mathcal{M}'}_{\xi'}) = [\![Rt]\!]^{\mathcal{M}'}_{\xi'}$.

3. $[\![A{\to}B]\!]^{\mathcal{M}}_{\xi} = \max\{1{-}[\![A]\!]^{\mathcal{M}}_{\xi}, [\![B]\!]^{\mathcal{M}}_{\xi}\} \overset{\text{IH}}{=} \max\{1{-}[\![A]\!]^{\mathcal{M}'}_{\xi'}, [\![B]\!]^{\mathcal{M}'}_{\xi'}\} = [\![A{\to}B]\!]^{\mathcal{M}'}_{\xi'}$.

4. $\mathcal{M} \models (\forall x A)[\xi] \Leftrightarrow \mathcal{M} \models A[\xi^a_x]$ for all $a \in |\mathcal{M}| \overset{\text{IH}}{\Leftrightarrow} \mathcal{M}' \models A[\pi \circ \xi^a_x]$ for all $a \in |\mathcal{M}| \Leftrightarrow \}$

$\Leftrightarrow \mathcal{M}' \models A[(\pi \circ \xi)^{\pi(a)}_x]$ for all $a \in |\mathcal{M}| \Leftrightarrow \mathcal{M}' \models A[(\pi \circ \xi)^b_x]$ for all $b \in |\mathcal{M}'| \Leftrightarrow \mathcal{M}' \models A[\xi']$.

Notation:

Let $\mathrm{FV}(u) \subseteq \{x_1, ..., x_n\}$ ($x_1, ..., x_n$ pairwise distinct), and $a_1, ..., a_n \in |\mathcal{M}|$:

$u^{\mathcal{M}}[x_1/a_1, ..., x_n/a_n] := [\![u]\!]^{\mathcal{M}}_{\xi}$, where ξ is an $\mathcal{M}$-assignment with $\xi(x_i) = a_i$ for $i = 1, ..., n$.

$\mathcal{M} \models A[x_1/a_1, ..., x_n/a_n] :\Leftrightarrow A^{\mathcal{M}}[x_1/a_1, ..., x_n/a_n] = 1$.

If $x_1, ..., x_n$ are known from the context one briefly writes $u^{\mathcal{M}}[a_1, ..., a_n]$ or $\mathcal{M} \models A[a_1, ..., a_n]$, respectively.

With the just introduced notation Theorem 3.5 can be written as follows:

(a) $\pi(t^{\mathcal{M}}[a_1, ..., a_n]) = t^{\mathcal{M}'}[\pi(a_1), ..., \pi(a_n)]$, (b) $\mathcal{M} \models A[a_1, ..., a_n] \iff \mathcal{M}' \models A[\pi(a_1), ..., \pi(a_n)]$.

Theorem 3.6

For each infinite structure $\mathcal{M}$ there exists an elementary equivalent structure $\mathcal{M}_1$,
which is not isomorphic to $\mathcal{M}$.

Proof:

Let K be the power set of $|\mathcal{M}|$. To each $\alpha \in K$ we assign a new constant c_α. Let $\mathcal{L}' := \mathcal{L} \cup \{c_\alpha : \alpha \in K\}$ and $\Sigma' := \mathrm{Th}(\mathcal{M}) \cup \{\neg(c_\alpha \approx c_\beta) : \alpha, \beta \in K \;\&\; \alpha \neq \beta\}$. Each finite subset $\Delta \subseteq \Sigma'$ is satisfiable: if $\Delta \subseteq \mathrm{Th}(\mathcal{M}) \cup \{\neg(c_\alpha \approx c_\beta) : \alpha, \beta \in \{\alpha_1, ..., \alpha_n\}\}$ choose an $\mathcal{L}'$-expansion $\mathcal{M}'$ of $\mathcal{M}$ so that $c^{\mathcal{M}'}_{\alpha_i} \neq c^{\mathcal{M}'}_{\alpha_j}$ for $1 \leq i, j \leq n$ with $\alpha_i \neq \alpha_j$; then $\mathcal{M}' \models \Delta$. Hence, according to the Compactness Th. there is a model $\mathcal{M}'_1$ of Σ'. Let $\mathcal{M}_1 := \mathcal{M}'_1|\mathcal{L}$ (the $\mathcal{L}$-reduct of $\mathcal{M}'_1$). $\mathcal{M}'_1 \models \Sigma' \Rightarrow \mathcal{M}'_1 \models \mathrm{Th}(\mathcal{M}) \Rightarrow \mathcal{M}_1 \models \mathrm{Th}(\mathcal{M}) \Rightarrow \mathcal{M} \equiv \mathcal{M}_1$. $\mathcal{M}'_1 \models \neg(c_\alpha \approx c_\beta) \Rightarrow c^{\mathcal{M}'_1}_\alpha \neq c^{\mathcal{M}'_1}_\beta$. Therefore the mapping $F : K \to |\mathcal{M}_1|$, $\alpha \mapsto c^{\mathcal{M}'_1}_\alpha$ is one-one. If there were an isomorphism $\pi : \mathcal{M}_1 \to \mathcal{M}$, we so had a one-one mapping from K into $|\mathcal{M}|$ which is impossible.

Remark

According to Theorem 3.6 no infinite structure can be characterized uniquely (up to isomorphism) by an axiom system in 1st order predicate logic. But, for example, the structure $(\mathbb{N}, 0, S)$ (with $S(n) = n{+}1$) of natural numbers is uniquely determined (up to isomorphism) by the Peano axioms. This seems like a contradiction to the previous remark, but it isn't, since the system of Peano axioms does not belong to 1st order predicate logic. The induction axiom $\forall X(0 \in X \wedge \forall x(x \in X \to Sx \in X) \to \forall x(x \in X))$ with the quantifier $\forall X$ ranging over all subsets of $\mathbb{N}$ is a 2nd order formula. A structure, which is elementary equivalent but not isomorphic to $(\mathbb{N}, 0, S)$, is called a *nonstandard model* of the natural numbers. In such a nonstandard model the principle of complete induction does not hold for every set $X \subseteq \mathbb{N}$.

Lemma 3.7

There are countable nonstandard models of the natural numbers.

Proof:

Let $\mathcal{N} := (\mathbb{N}, 0, S)$ und $\Gamma := \mathrm{Th}(\mathcal{N}) \cup \{\neg(v_0 \approx 0), \neg(v_0 \approx S0), \neg(v_0 \approx SS0), \ldots\}$.

Obviously every finite subset of Γ and so Γ itself is satisfiable; hence there exists an interpretation $(\mathcal{N}_1, \xi)$ with $\mathcal{N}_1 \models \Gamma[\xi]$ and $\mathcal{N}_1$ countable. Now $\mathcal{N} \equiv \mathcal{N}_1$, since $\mathcal{N}_1 \models \mathrm{Th}(\mathcal{N})$.

Assumption: There is an isomorphism $\pi : \mathcal{N} \to \mathcal{N}_1$.

Then $\pi(n) = \pi(\mathcal{N}(\underbrace{S...S}_{n}0)) = \mathcal{N}_1(S...S0) \neq \xi(v_0)$, for each $n \in \mathbb{N}$. So π would not be surjective.

Remark

The structure $(\mathbb{R}, 0, 1, +, \cdot, <)$ of real numbers is up to isomorphism the sole complete ordered field. The completeness axiom $\forall X(\emptyset \neq X \text{ bounded } \to \exists y(y = \sup(X)))$ is (as well as the induction axiom) a 2nd order formula. A structure which is elementary equivalent but not isomorphic to $(\mathbb{R}, 0, 1, +, \cdot, <)$, is called a *nonstandard model* of the real numbers. In a nonstandard model of the real numbers not every bounded set $X \neq \emptyset$ has a supremum (but every such set which is definable by a formula of the language $\{0, 1, +, \cdot, <\}$ has a supremum).

Lemma 3.8

For each archimedean ordered field there is an elementary equivalent ordered field which is not archimedean.

Corollary. If a sentence A of the language $\{0, 1, +, \cdot, <\}$ holds in every nonarchimedean field, then A holds in *every* ordered field.

Proof:

Let $\mathcal{K}$ be an archimedean ordered field. $\Gamma := \mathrm{Th}(\mathcal{K}) \cup \{\underline{n} < v_0 : 1 \leq n \in \mathbb{N}\}$, where $(\underline{n} := \overbrace{1 + ... + 1}^{n})$.

Obviously every finite subset of Γ is satisfiable, and so Γ is satisfiable. Assume $\mathcal{M} \models \Gamma[\xi]$. Then $\mathcal{M} \equiv \mathcal{K}$, so $\mathcal{M}$ is an ordered field. From $\mathcal{M} \models (\underline{n} < v_0)[\xi]$ it follows thatt $1^{\mathcal{M}} \cdot n <^{\mathcal{M}} \xi(v_0)$ for all $n \in \mathbb{N}$, hence $\mathcal{M}$ is nonarchimedean.

Proof of the Corollary: Let $\mathcal{K}$ be an archimedean ordered field. Then there exists a nonarchimedean ordered field $\mathcal{M}$ with $\mathcal{K} \equiv \mathcal{M}$. From $\mathcal{M} \models A$ we obtain $\mathcal{K} \models A$.

Definition

A class $\mathcal{S}$ of $\mathcal{L}$-structures is *(finitely) axiomatizable*,
if there is a (finite) axiom system Σ such that $\mathcal{S} = \mathrm{Mod}_{\mathcal{L}}(\Sigma)$.

Proposition

(a) $\mathcal{S}$ finitely axiomatizable $\iff$ There is an $\mathcal{L}$-sentence A such that $\mathcal{S} = \mathrm{Mod}_{\mathcal{L}}(\{A\})$.

(b) If there are structures $\mathcal{M}, \mathcal{M}_0$ with $\mathcal{M} \equiv \mathcal{M}_0$ & $\mathcal{M} \in \mathcal{S}$ & $\mathcal{M}_0 \notin \mathcal{S}$, then $\mathcal{S}$ is not axiomatizable.

Lemma 3.9

Let $\mathcal{S}$ be a class of $\mathcal{L}$-structures and Σ an axiom system.

(a) $\mathcal{S}$ is finitely axiomatizable if, and only if, $\mathcal{S}$ and its complement $\overline{\mathcal{S}}$ are axiomatizable.

(b) If $\mathrm{Mod}_{\mathcal{L}}(\Sigma)$ is finitely axiomatizable,

then there exists a finite subset $\Delta \subseteq \Sigma$ such that $\mathrm{Mod}_{\mathcal{L}}(\Sigma) = \mathrm{Mod}_{\mathcal{L}}(\Delta)$.

Proof:

(a) "$\Rightarrow$": For $\mathcal{S} = \mathrm{Mod}_{\mathcal{L}}(\{A\})$ we have: $\mathcal{M} \in \overline{\mathcal{S}} \;\Leftrightarrow\; \mathcal{M} \not\models A \;\Leftrightarrow\; \mathcal{M} \models \neg A$.

"$\Leftarrow$": Let $\mathcal{S} = \mathrm{Mod}_{\mathcal{L}}(\Sigma_1)$ and $\overline{\mathcal{S}} = \mathrm{Mod}_{\mathcal{L}}(\Sigma_2)$. Then $\Sigma_1 \cup \Sigma_2$ is not satisfiable; hence there is a finite $\Delta \subseteq \Sigma_1$ so that $\Delta \cup \Sigma_2$ is not satisfiable. $\mathcal{M} \in \mathcal{S} \;\Rightarrow\; \mathcal{M} \models \Delta \;\Rightarrow\; \mathcal{M} \not\models \Sigma_2 \;\Rightarrow\; \mathcal{M} \notin \overline{\mathcal{S}} \;\Rightarrow\; \mathcal{M} \in \mathcal{S}$.

(b) Let $\mathrm{Mod}_{\mathcal{L}}(\Sigma) = \mathrm{Mod}_{\mathcal{L}}(\{A\})$. Then $\Sigma \models A$, and consequently there is a finite $\Delta \subseteq \Sigma$ with $\Delta \models A$. Hence: $\mathcal{M} \models \Sigma \;\Rightarrow\; \mathcal{M} \models \Delta \;\Rightarrow\; \mathcal{M} \models A \;\Rightarrow\; \mathcal{M} \models \Sigma$.

Complete theories

I. The theory DO of dense linear orders without endpoints

Let **DO** be the deductive closure of the following axiom system:

(1) $\forall x \neg(x < x) \wedge \forall x \forall y \forall z(x < y \wedge y < z \rightarrow x < z) \wedge \forall x \forall y(x < y \vee x \approx y \vee y < x)$

(2) $\forall x \forall y(x < y \rightarrow \exists z(x < z \wedge z < y))$

(3) $\forall x \exists y(x < y) \wedge \forall x \exists y(y < x)$.

Lemma 3.10

Every countable model of **DO** is isomorphic to the structure $(\mathbb{Q}, <)$ of rational numbers.

Proof:

Let $\mathcal{M} = (M, \prec)$ be a countable model of **DO**, and let $M = \{b_n : n \in \mathbb{N}\}$ and $\mathbb{Q} = \{a_n : n \in \mathbb{N}\}$.

Definition of order preserving functions $F_n \subseteq \mathbb{Q} \times M$ by recursion on n:

1. $F_0 := \{(a_0, b_0)\}$.

2. Definition of F_{n+1}:

2.1. $n = 2m$. If $a_m \in \mathrm{dom}(F_n)$ then $F_{n+1} := F_n$.

Otherwise let $F_{n+1} := F_n \cup \{(a_m, b_k)\}$ with $k := \min\{i : F_n \cup \{(a_m, b_i)\}$ order preserving$\}$.

Since $\mathcal{M}$ is a model of **DO**, such a b_k always exists.

2.2. $n = 2m + 1$. If $b_m \in \mathrm{ran}(F_n)$ then $F_{n+1} := F_n$.

Otherwise let $F_{n+1} := F_n \cup \{(a_k, b_m)\}$ with $k := \min\{i : F_n \cup \{(a_i, b_m)\}$ order preserving$\}$.

Obviously $F := \bigcup_{n \in \mathbb{N}} F_n$ is an isomorphism from $(\mathbb{Q}, <)$ onto $\mathcal{M}$.

Theorem 3.11. The theory **DO** is complete, and **DO**$= \mathrm{Th}(\mathbb{Q}, <)$.

Proof:

By Theorem 3.4 if suffices to prove that every model of **DO** is elementary equivalent to $(\mathbb{Q}, <)$. So let $\mathcal{M}$ be a model of **DO**. By L.3.3d there exists a countable structure $\mathcal{M}_0$ which is elementary equivalent to $\mathcal{M}$. By L.3.10 $\mathcal{M}_0$ isomorphic to $(\mathbb{Q}, <)$; hence $\mathcal{M} \equiv \mathcal{M}_0 \equiv (\mathbb{Q}, <)$.

II. The theory of algebraically closed fields with fixed charakteristic

The class of all algebraically closed fields of charakteristic p (p prime number or $p = 0$) can be axiomatized the axiom system consisting of all field axioms and the sentences $\forall x_0 ... \forall x_n (x_n \neq 0 \rightarrow \exists y (x_n \cdot y^n + x_{n-1} \cdot y^{n-1} + ... + x_0))$ $(n \geq 1)$ and $\underbrace{1 + ... + 1}_{p} = 0$ bzw. (in case $p = 0$) $\underbrace{1 + ... + 1}_{n} \neq 0$ $(1 \leq n \in \mathbb{N})$.

Let ACF_p be the deductive closure of this axiomen system. It can be shown that the theory ACF_p is complete. From this we obtain (e.g.): If a sentence A of the language $\{+, \cdot, 0, 1\}$ holds in the field $\mathbb{C}$ of complex numbers, then it holds in every algebraically closed field of characteristic 0.

On the proof of the completeness of ACF_p.

Theorem 3.6*

If κ is and infinite cardinal $\geq card(\mathcal{L})$, then for every inifinite $\mathcal{L}$-structure $\mathcal{M}$ there exists an elementary equivalent $\mathcal{L}$-structure $\mathcal{M}_1$ of cardinality κ.

Proof will follow in a later chapter.

Definition

Let κ be a cardinal. An axiom system Σ is called *κ-categorical*, if every two models of Σ of cardinality κ are isomorphic.

Theorem 3.12 (Vaught's Test)

Let κ be an infinite cardinal. If Σ is a κ-categorical axiom system with $card(\Sigma) \leq \kappa$ and if Σ has only infinite models, then Σ is complete.

Proof:

Let $\mathcal{L} := L(\Sigma)$ and $\mathcal{M}, \mathcal{N} \in \mathrm{Mod}_{\mathcal{L}}(\Sigma)$. Then $\mathcal{M}, \mathcal{N}$ are infinite and $\kappa \geq card(\mathcal{L})$. By 3.6* there are $\mathcal{L}$-structures $\mathcal{M}_1, \mathcal{N}_1$ of cardinality κ such that $\mathcal{M} \equiv \mathcal{M}_1$ and $\mathcal{N} \equiv \mathcal{N}_1$. Hence $\mathcal{M}_1, \mathcal{N}_1 \models \Sigma$ and thus $\mathcal{M}_1 \cong \mathcal{N}_1$. By Theorem 3.5 we now obtain $\mathcal{M} \equiv \mathcal{N}$.

Satz 3.13. Every two uncountable algebraically closed fields having the same cardinality and the same characteristic are isomorphic. For each prime number p there exists an algebraically closed field with characteristic p.

Lemma 3.14 Every algebraically closed field is infinite.

Proof: If $K = \{a_1, ..., a_n\}$, then the polynomial $(x - a_1) \cdot ... \cdot (x - a_n) + 1$ has no root in K.

From 3.6*, 3.13, 3.14 it follows that ACF_p is $\aleph_1$-categorical. By 3.12 from this it follows that ACF_p is complete.

Remark.

If the axiom system Σ is decidable and complete, then also its deductive closure $T_\Sigma := \{A \in \text{FOR}^0_{\text{L}(\Sigma)} : \Sigma \vdash A\}$ is decidable.

"Proof": Abb.: $\mathcal{L} := \text{L}(\Sigma)$. Assume that Σ is consistent (otherwise $\Sigma = \text{FOR}^0_{\mathcal{L}}$). The set of all derivations H in $\mathcal{L}$ can be effectively enumerated. Let $H_0, H_1, H_2, ...$ be such an enumeration, and let, for each $i \in \mathbb{N}$, A_i be the endformula of H_i. It is decidable whether H_i is a derivation of A_i *from* Σ. Now for an arbitrary $\mathcal{L}$-sentence B by the following procedure it can be decided whether $\Sigma \vdash B$ or not: search through the sequence of derivations $H_0, H_1, ...$ until you find a derivation H_n of A_n from Σ with $A_n \in \{B, \neg B\}$; since Σ is complete, this situation will always be attained, and we have: $\Sigma \vdash B \Leftrightarrow A_n = B$.

Definition (Substructure)

Let $\mathcal{M}_0, \mathcal{M}$ be $\mathcal{L}$-structures.

$\mathcal{M}_0$ is a *substructure* of $\mathcal{M}$ (in symbols $\mathcal{M}_0 \subseteq \mathcal{M}$), if $|\mathcal{M}_0| \subseteq |\mathcal{M}|$ and further

(i) $c^{\mathcal{M}_0} = c^{\mathcal{M}}$ for each constant $c \in \mathcal{L}$,

(ii) $f^{\mathcal{M}_0} = f^{\mathcal{M}} \restriction |\mathcal{M}_0|^n$ for each n-ary function symbol $f \in \mathcal{L}$ $(n \geq 1)$,

(iii) $R^{\mathcal{M}_0} = R^{\mathcal{M}} \cap |\mathcal{M}_0|^n$ for each n-ary relation symbol $R \in \mathcal{L}$.

An *(isomorphic) embedding from* $\mathcal{M}_0$ *into* $\mathcal{M}$ is a one-one mapping $\pi : |\mathcal{M}_0| \to |\mathcal{M}|$ with

(a) $\pi(f^{\mathcal{M}_0}(a_1, ..., a_n)) = f^{\mathcal{M}}(\pi(a_1), ..., \pi(a_n))$ $\qquad (f \in \mathcal{L},\ n \geq 0,\ a_1, ..., a_n \in |\mathcal{M}_0|)$

(b) $(a_1, ..., a_n) \in R^{\mathcal{M}_0} \Leftrightarrow (\pi(a_1), ..., \pi(a_n)) \in R^{\mathcal{M}}$ $\qquad (R \in \mathcal{L},\ n \geq 1,\ a_1, ..., a_n \in |\mathcal{M}|)$

Lemma 3.15

A mapping $\pi : |\mathcal{M}_0| \to |\mathcal{M}|$ is an isomorphic embedding from $\mathcal{M}$ in $\mathcal{M}'$ iff it is an isomorphism from $\mathcal{M}_0$ onto a substructure of $\mathcal{M}$.

§4 Recursive functions

Inductive Definition of sets PR^n of n-ary function symbols

(PR 1) $\mathsf{C}_k^n \in \mathrm{PR}^n$ $(n, k \geq 0)$, $\mathsf{S} \in \mathrm{PR}^1$, $\mathrm{I}_i^n \in \mathrm{PR}^n$ $(1 \leq i \leq n)$.

(PR 2) $h \in \mathrm{PR}^m$ & $g_1, ..., g_m \in \mathrm{PR}^n$ & $m, n \geq 1 \implies (\circ h g_1 ... g_m) \in \mathrm{PR}^n$.

(PR 3) $g \in \mathrm{PR}^n$ & $h \in \mathrm{PR}^{n+2} \implies (\mathsf{R}gh) \in \mathrm{PR}^{n+1}$.

Abbreviation: $\mathrm{PR} := \bigcup_{n \in \mathbb{N}} \mathrm{PR}^n$, $\mathsf{0} := \mathsf{C}_0^0$. Remark: $\mathrm{PR}^0 = \{\mathsf{C}_k^0 : k \in \mathbb{N}\}$.

Definition of the *standard structure* $\mathcal{N}$ for the language PR

$|\mathcal{N}| := \mathbb{N}$,

$(\mathsf{C}_k^n)^{\mathcal{N}}(\vec{a}) := k$,

$\mathsf{S}^{\mathcal{N}}(a) := a{+}1$,

$(\mathrm{I}_i^n)^{\mathcal{N}}(a_1, ..., a_n) := a_i$,

$(\circ h g_1 ... g_m)^{\mathcal{N}}(\vec{a}) := h^{\mathcal{N}}(g_1^{\mathcal{N}}(\vec{a}), \ldots, g_m^{\mathcal{N}}(\vec{a}))$,

$(\mathsf{R}gh)^{\mathcal{N}}(\vec{a}, 0) := g^{\mathcal{N}}(\vec{a})$,

$(\mathsf{R}gh)^{\mathcal{N}}(\vec{a}, b{+}1) := h^{\mathcal{N}}(\vec{a}, b, (\mathsf{R}gh)^{\mathcal{N}}(\vec{a}, b))$.

Definition

A function $F : \mathbb{N}^n \to \mathbb{N}$ is called *primitive recursive*, if $F = f^{\mathcal{N}}$ for some $f \in \mathrm{PR}^n$.

A relation $R \subseteq \mathbb{N}^n$ is called *primitive recursive*, if its characteristic function

$\mathbf{1}_R : \mathbb{N}^n \to \mathbb{N}, \mathbf{1}_R(\vec{a}) := \begin{cases} 1 & \text{if } \vec{a} \in R \\ 0 & \text{otherwise} \end{cases}$ is primitive recursive.

Definition of an n-ary function symbol $\lambda x_1 ... x_n.t \in \mathrm{PR}^n$ for each PR-term t and pairwise distinct variables $x_1, ..., x_n$ $(n \geq 1)$ with $\mathrm{FV}(t) \subseteq \{x_1, ..., x_n\}$:

1. $\lambda x_1 ... x_n.\mathsf{C}_k^0 := \mathsf{C}_k^n$
2. $\lambda x_1 ... x_n.x_i := \mathrm{I}_i^n$
3. $\lambda x_1 ... x_n.h t_1 ... t_m := (\circ h g_1 ... g_m)$ with $g_i := \lambda x_1 ... x_n.t_i$

Lemma 4.1

$(\lambda x_1 ... x_n.t)^{\mathcal{N}}(a_1, ..., a_n) = t^{\mathcal{N}}[x_1/a_1, ..., x_n/a_n]$.

Proof by induction on the build up of t:

Abb.: $t^{\mathcal{N}}[\vec{a}] := t^{\mathcal{N}}[x_1/a_1, ..., x_n/a_n]$.

1. $(\lambda \vec{x}.\mathsf{C}_k^0)^{\mathcal{N}}(\vec{a}) = (\mathsf{C}_k^n)^{\mathcal{N}}(\vec{a}) = k = (\mathsf{C}_k^0)^{\mathcal{N}}[\vec{a}]$.
2. $(\lambda \vec{x}.x_i)^{\mathcal{N}}(\vec{a}) = (\mathrm{I}_i^n)^{\mathcal{N}}(\vec{a}) = a_i = x_i^{\mathcal{N}}[\vec{a}]$
3. Let $t = h t_1 ... t_m$ and $g_i := \lambda x_1 ... x_n.t_i$.

$(\lambda \vec{x}.t)^{\mathcal{N}}(\vec{a}) = (\circ h g_1 ... g_m)^{\mathcal{N}}(\vec{a}) = h^{\mathcal{N}}(g_1^{\mathcal{N}}(\vec{a}), ..., g_m^{\mathcal{N}}(\vec{a})) = h^{\mathcal{N}}(t_1^{\mathcal{N}}[\vec{a}], ..., t_m^{\mathcal{N}}[\vec{a}]) = t^{\mathcal{N}}[\vec{a}]$.

Convention

We use $a, b, c, i, j, k, l, m, n$ as syntactic variables for natural numbers, and s, t for PR-terms.

We will not always distinguish between a function symbol and its interpretation in $\mathcal{N}$: For example, if $f \in \mathrm{PR}^n$ then the function $f^{\mathcal{N}} : \mathbb{N}^n \to \mathbb{N}$ will also be denoted by f. Correspodingly PR (PR^n) also denotes

the set of all (n-ary) primitive recursive functions. So, PR is the least set of functions, which contains the *basic functions* C^n_k, S, I^n_i and is closed under the operations $\circ$ *(composition)* and R *(primitive recursion)*.

Corollary. (PR is closed under explicit definitions)
If t is a PR-term with $\mathrm{FV}(t) \subseteq \{x_1, ..., x_n\}$, and $f : \mathbb{N}^n \to \mathbb{N}$ is defined by $f(a_1, ..., a_n) := t^{\mathcal{N}}[a_1, ..., a_n]$ then f is primitive recursive. (Proof: Lemma 4.1.)

Definition of some special function symbols (primitive recursive functions, resp.):

$\mathsf{plus} := (\mathsf{R}\,\mathrm{I}^1_1(\lambda xyz.\mathsf{S}z))$, $\mathsf{pd} := (\mathsf{R0}\lambda yz.y)$, $\dot{-} := (\mathsf{RI}^1_1\lambda xyz.\mathsf{pd}z)$, $\mathsf{times} := (\mathsf{RC}^1_0\lambda xyz.\mathsf{plus}zx)$.
For $f \in \mathrm{PR}^{n+1}$ we set $(\sum f) := (\mathsf{RC}^n_0\lambda x_1...x_nyz.(\mathsf{plus}\,z f x_1...x_n y))$.

For the just defined functions the following holds:
$\mathsf{plus}(a,b) = a + b$, $\mathsf{times}\,(a,b) = a \cdot b$, $\mathsf{pd}(a) = \begin{cases} 0 & \text{if } a = 0 \\ a-1 & \text{otherwise} \end{cases}$, $\dot{-}(a,b) = \begin{cases} a-b & \text{if } a \geq b \\ 0 & \text{otherwise} \end{cases}$.
$(\sum f)(\vec{a}, b) = \sum_{i<b} f(\vec{a}, i)$.
From now on we write $+$, $\cdot$ for plus, times.

Remark.
A relation $R \subseteq \mathbb{N}^n$ is primitive recursive iff there is an $f \in \mathrm{PR}^n$ with $R = \{\vec{a} \in \mathbb{N}^n : f(\vec{a}) = 0\}$.
(Proof: $1 \dot{-} f(\vec{a}) = \begin{cases} 1 & \text{if } f(\vec{a}) = 0 \\ 0 & \text{otherwise} \end{cases}$)

Abbreviations:
$s < t := (\mathsf{S}s \dot{-} t \approx \mathsf{0})$,
$\forall x < tA := \forall x(x < t \to A)$, $\exists x < tA := \neg\forall x < t\neg A$ (if $x \notin \mathrm{FV}(t)$)
$\forall x \leq tA := \forall x < \mathsf{S}tA$, $\exists x \leq tA := \exists x < \mathsf{S}tA$ (if $x \notin \mathrm{FV}(t)$)
Obviously $\mathcal{N} \models (s < t)[\vec{a}] \Leftrightarrow s^{\mathcal{N}}[\vec{a}] < t^{\mathcal{N}}[\vec{a}]$.

Inductive Definition of Δ_0-formulas (of the language PR)

1. Each atomic PR-formula (i.e. equation $s \approx t$) is a Δ_0-formula.
2. If A, B are Δ_0-formulas, then $\neg A$, $A \wedge B$, $A \vee B$, $A \to B$ are Δ_0-formulas.
3. If A is a Δ_0-formula, and t a PR-term with $x \notin \mathrm{FV}(t)$, then $\forall x < tA$ and $\exists x < tA$ are Δ_0-formulas.

Lemma 4.2
If A is a Δ_0-formula with $\mathrm{FV}(A) \subseteq \{x_1, ..., x_n\}$,
then the relation $\{(a_1, ..., a_n) \in \mathbb{N}^n : \mathcal{N} \models A[x_1/a_1, ..., x_n/a_n]\}$ is primitive recursive.

Proof:
By induction on the build up of A we define a PR-term r_A with $\mathrm{FV}(r_A) = \mathrm{FV}(A)$ and $\mathcal{N} \models r_A \approx 0 \leftrightarrow A$.
$r_{s\approx t} := (s \dot{-} t) + (t \dot{-} s)$, $r_{\neg A} := \mathsf{S0} \dot{-} r_A$, $r_{A\wedge B} := r_A + r_B$, $r_{A\vee B} := r_A \cdot r_B$, $r_{A\to B} := r_{\neg A\vee B}$,
$r_{\forall y<tB} := (\sum \lambda\vec{x}y.r_B)\vec{x}t$, where $\mathrm{FV}(\forall y{<}tB) = \{\vec{x}\}$, $r_{\exists x<tB} := r_{\neg\forall x<t\neg B}$.

Corollary
The set of primitive recursive relations is closed under $\cap$, $\cup$, $\setminus$, bounded quantification, and substitution of prim. rec. functions.

Lemma 4.3

If $f_1, ..., f_{k+1} \in \mathrm{PR}^n$, and if $R_1, ..., R_k \subseteq \mathbb{N}^n$ are pairwise disjoint primitive recursive relations, then also the following function $f : \mathbb{N}^n \to \mathbb{N}$ is primitive recursive:

$$f(\vec{a}) := \begin{cases} f_1(\vec{a}) & \text{if } \vec{a} \in R_1 \\ \dots & \dots \\ f_k(\vec{a}) & \text{if } \vec{a} \in R_k \\ f_{k+1}(\vec{a}) & \text{otherwise} \end{cases} .$$

Proof:

Let $R_{k+1} := \mathbb{N}^n \setminus (R_1 \cup \ldots \cup R_k)$. $\quad f := \lambda\vec{x}.(f_1(\vec{x}) \cdot \mathbf{1}_{R_1}(\vec{x}) + \ldots + f_{k+1}(\vec{x}) \cdot \mathbf{1}_{R_{k+1}}(\vec{x}))$.

Definition (bounded μ-operator)

For $g : \mathbb{N}^{n+1} \to \mathbb{N}$ let $(\overline{\mu}g) : \mathbb{N}^{n+1} \to \mathbb{N}$, $(\overline{\mu}g)(\vec{a}, b) := \begin{cases} \min\{i : g(\vec{a}, i) = 0\} & \text{if } \exists i < b(g(\vec{a}, i) = 0) \\ b & \text{otherwise.} \end{cases}$

Lemma 4.4

Let $g : \mathbb{N}^{n+1} \to \mathbb{N}$ be primitive recursive. Then the following holds:

(a) $(\overline{\mu}g)$ is primitive recursive.

(b) If there exists an $h \in \mathrm{PR}^n$ with $\forall \vec{a} \in \mathbb{N}^n \exists i < h(\vec{a})(\, g(\vec{a}, i) = 0\,)$,

then the function $\vec{a} \mapsto \min\{i : g(\vec{a}, i) = 0\}$ is primitive recursive.

Proof:

(a) By 4.2 and 4.3 the following function $p : \mathbb{N}^{n+1} \to \mathbb{N}$ is primitive recursive:

$p(\vec{a}, c) := \begin{cases} c, & \text{if } g(\vec{a}, c) = 0 \;\&\; \forall i < c(g(\vec{a}, i) \neq 0) \\ 0, & \text{otherwise.} \end{cases}$

Further we have: $(\overline{\mu}g)(\vec{a}, b) = \begin{cases} (\sum p)(\vec{a}, b) & \text{if } \exists i < b(g(\vec{a}, i) = 0) \\ b & \text{otherwise.} \end{cases}$

(b) $\min\{i :\ g(\vec{a}, i) = 0\} = (\overline{\mu}g)(\vec{a}, h(\vec{a}))$.

Definition (The primitive recursive pairing function $\pi : \mathbb{N}^2 \to \mathbb{N}$)

$\pi(a, b) := b + \sum_{i<a+b}(i + 1) = b + (\sum \mathsf{S})(a + b)$

Lemma 4.5

(a) $\pi : \mathbb{N}^2 \to \mathbb{N}$ is bijective. $\quad$ (b) $a, b \leq \pi(a, b)$ and $(0 < a \Rightarrow b < \pi(a, b))$.

Proof:

(a)

$0 = \hat{\pi}(0,0)$	$2 = \hat{\pi}(0,1)$	$5 = \hat{\pi}(0,2)$	$9 = \hat{\pi}(0,3)$	$\dots$
$1 = \hat{\pi}(1,0)$	$4 = \hat{\pi}(1,1)$	$8 = \hat{\pi}(1,2)$	$\dots$	$\dots$
$3 = \hat{\pi}(2,0)$	$7 = \hat{\pi}(2,1)$	$\dots$	$\dots$	$\dots$
$6 = \hat{\pi}(3,0)$	$\vdots\vdots$	$\vdots\vdots$	$\vdots\vdots$	$\vdots\vdots$
$\dots$				

Obviously the above scheme defines a bijection $\hat{\pi}$ from $\mathbb{N}^2$ onto $\mathbb{N}$; and the following holds:

$\hat{\pi}(a, b) = \hat{\pi}(a+b, 0)+b$, and $\hat{\pi}(a+b, 0) =$ number of all equations in previous diagonals $= 1+2+\ldots+(a+b)$.

Consequently $\pi = \hat{\pi}$.

(b) is obvious.

Definition (The inverse functions π_1, π_2 of π)

$\pi_1(a) := \min\{i :\ \exists j \leq a[a = \pi(i, j)]\}$, $\quad \pi_2(a) := \min\{j :\ a = \pi(\pi_1(a), j)\}$.

Lemma 4.6

(a) The functions π_1, π_2 are primitive recursive.

(b) $\pi(\pi_1(a), \pi_2(a)) = a$

(c) $\pi_i(\pi(a_1, a_2)) = a_i \quad (i = 1, 2)$

Proof:

(a) By Lemma 4.5 we have $\exists i \leq a \exists j \leq a(a = \pi(i, j))$ and $\exists j \leq a(a = \pi(\pi_1(a), j))$.

From this we obtain the assertion by Lemma 4.4.

(b) and (c) follow from Lemma 4.5a and the definition of π_1, π_2.

Remark. If $f \in \mathrm{PR}^1$ then also its iteration $(a, k) \mapsto f^k(a)$ defined by $f^0(a) := a$, $f^{k+1}(a) := f(f^k(a))$ is primitive recursive. Note that $f^{k+1}(a) = f^k(f(a))$. We also write $f^{(k)}$ for f^k.

Definition (Coding of finite sequences of natural numbers)

$\mathrm{cons}(a, b) := \pi(a, b) + 1$,

$\mathrm{hd}(a) := \pi_1(a \dot{-} 1)$,

$\mathrm{tl}(a) := \pi_2(a \dot{-} 1)$,

$(a)_i := \mathrm{hd}(\mathrm{tl}^i(a))$ (obviously $(a)_i = 0$ for $\mathrm{lh}(a) \leq i$),

$\mathrm{lh}(a) := \min\{k : \mathrm{tl}^k(a) = 0\} = (\overline{\mu}\tau)(a, a + 1)$, where $\tau(a, k) := \mathrm{tl}^k(a)$.

The functions cons, tl, hd, lh, $(a, i) \mapsto (a)_i$ are primitive recursive.

Lemma 4.7

(a) $\mathrm{lh}(\mathrm{cons}(a, b)) = \mathrm{lh}(b) + 1$

(b) $\big(\mathrm{cons}(a, b)\big)_0 = a \;\&\; \forall i < \mathrm{lh}(b)[\,\big(\mathrm{cons}(a, b)\big)_{i+1} = (b)_i\,]$

(c) $i < \mathrm{lh}(a) \;\Rightarrow\; (a)_i < a$

Proof:

Let $n := \mathrm{lh}(b)$ and $c := \mathrm{cons}(a, b)$. Then $a = \mathrm{hd}(c)$ and $\mathrm{tl}^{i+1}(c) = \mathrm{tl}^i(b)$, especially $b = \mathrm{tl}(c)$.

(a) We have $\mathrm{tl}^{n+1}(c) = \mathrm{tl}^n(b) = 0 \;\&\; \forall i < n(\mathrm{tl}^{i+1}(c) = \mathrm{tl}^i(b) \neq 0) \;\&\; \mathrm{tl}^0(c) = c \neq 0$; hence $\mathrm{lh}(c) = n + 1$.

(b) $(c)_0 = \mathrm{hd}(\mathrm{tl}^0(c)) = a$ and, for $i < n$, $(c)_{i+1} = \mathrm{hd}(\mathrm{tl}^{i+1}(c)) = \mathrm{hd}(\mathrm{tl}^i(b)) = (b)_i$.

(c) $i < \mathrm{lh}(a) \;\Rightarrow\; (a)_i = \pi_1(\mathrm{tl}^i(a) \dot{-} 1) \leq \mathrm{tl}^i(a) \dot{-} 1 < \mathrm{tl}^i(a) \leq a$.

Lemma 4.8

$\mathrm{lh}(c) = \mathrm{lh}(c') \;\&\; \forall i < \mathrm{lh}(c)((c)_i = (c')_i) \;\Longrightarrow\; c = c'$.

Proof:

Fall 1: $\mathrm{lh}(c) = 0$. Then $c = 0 = c'$.

Fall 2: $\mathrm{lh}(c) = k + 1$. Then $c, c' > 0$ and there are a, b, a', b' with $c = \mathrm{cons}(a, b)$, $c' = \mathrm{cons}(a', b')$. By Lemma 4.7 and the assumption we now obtain $\mathrm{lh}(b) = k = \mathrm{lh}(b') \;\&\; a = (c)_0 = (c')_0 = a' \;\&\; \forall i < k((b)_i = (c)_{i+1} = (c')_{i+1} = (b')_i)$. By I.H. this yields $b = b'$, and thus $c = c'$.

Definition of $\langle a_0, \ldots, a_{n-1} \rangle$

$\langle\rangle := 0$, $\langle a_0, \ldots, a_n \rangle := \mathrm{cons}(a_0, \langle a_1, \ldots, a_n \rangle)$

Obviously for each fixed $n \in \mathbb{N}$ the function $(a_0, \ldots, a_n) \mapsto \langle a_0, \ldots, a_n \rangle$ is primitive recursive.

Lemma 4.9

(a) $\mathrm{lh}(\langle a_0, ..., a_{n-1}\rangle) = n$

(b) $i < n \Rightarrow \big(\langle a_0, ..., a_{n-1}\rangle\big)_i = a_i$

(c) $a = \langle (a)_0, ..., (a)_{n-1}\rangle$ mit $n := \mathrm{lh}(a)$.

Proof: Lemmata 4.7, 4.8.

Lemma 4.10

The function $a * b := \langle (a)_0, ..., (a)_{\mathrm{lh}(a)-1}, (b)_0, ..., (b)_{\mathrm{lh}(b)-1}\rangle$ is primitive recursive.

Proof: Let $h(a,b,0) := b$, $h(a,b,i{+}1) := \mathrm{cons}((a)_{\mathrm{lh}(a) \dot{-} (i+1)}, h(a,b,i))$.

For $i \leq \mathrm{lh}(a)$ we then have $h(a,b,i) = \langle (a)_{\mathrm{lh}(a)-i}, ..., (a)_{\mathrm{lh}(a)-1}, (b)_0, ..., (b)_{\mathrm{lh}(b)-1}\rangle$

and thus $h(a,b,\mathrm{lh}(a)) = a * b$.

$[h(a,b,0) = b = \langle (b)_0, ..., (b)_{\mathrm{lh}(b)-1}\rangle.\ h(a,b,i{+}1) = \mathrm{cons}((a)_{\mathrm{lh}(a) \dot{-} (i+1)}, h(a,b,i)) \overset{\text{I.H.}}{=}$

$= \mathrm{cons}((a)_{\mathrm{lh}(a) \dot{-} (i+1)}, \langle (a)_{\mathrm{lh}(a)-i}, ..., (a)_{\mathrm{lh}(a)-1}, (b)_0, ..., (b)_{\mathrm{lh}(b)-1}\rangle)]$

Definition

For $f : \mathbb{N}^{n+1} \to \mathbb{N}$ let $\overline{f} : \mathbb{N}^{n+1} \to \mathbb{N}$, $\overline{f}(\vec{a}, b) := \langle f(\vec{a}, 0), \ldots, f(\vec{a}, b-1)\rangle$.

Remark: $f \in \mathrm{PR}^{n+1} \implies \overline{f} \in \mathrm{PR}^{n+1}$.

Proof: $\overline{f}(\vec{a}, 0) = 0$, $\overline{f}(\vec{a}, b+1) = \overline{f}(\vec{a}, b) * \langle f(\vec{a}, b)\rangle$.

Lemma 4.11 (Course-of-value recursion)

For each $h \in \mathrm{PR}^{n+2}$ the function f defined recursively by $f(\vec{a}, b) := h(\vec{a}, b, \overline{f}(\vec{a}, b))$

is primitive recursive too.

Proof:

$\overline{f}(\vec{a}, 0) = 0$, $\overline{f}(\vec{a}, b+1) = \overline{f}(\vec{a}, b) * \langle h(\vec{a}, b, \overline{f}(\vec{a}, b))\rangle$, $f(\vec{a}, b) = \big(\overline{f}(\vec{a}, b+1)\big)_b$.

Corollary 4.12

Let $H \subseteq \mathbb{N}^{n+2}$ be primitive recursive, and let $Q \subseteq \mathbb{N}^{n+1}$ so that:

$\forall (\vec{a}, b) \in \mathbb{N}^{n+1} [\, Q(\vec{a}, b) \ \Leftrightarrow\ H(\vec{a}, b, \overline{\mathbf{1}_Q}(\vec{a}, b))\,]$.

Then Q is primitive recursive.

Proof: Obviously $\mathbf{1}_Q(\vec{a}, b) = \mathbf{1}_H(\vec{a}, b, \overline{\mathbf{1}_Q}(\vec{a}, b))$.

Remark . For $Q \subseteq \mathbb{N}^{n+1}$ the following holds: $\forall i < \mathrm{lh}(b) [\, Q(\vec{a}, (b)_i) \ \Leftrightarrow\ \big(\overline{\mathbf{1}_Q}(\vec{a}, b)\big)_{(b)_i} = 1\,]$.

Recursively enumerable relations

Definition

A relation $Q \subseteq \mathbb{N}^n$ is *recursively enumerable* (*r.e.* for short), if there is primitive recursive relation $R \subseteq \mathbb{N}^{n+1}$ such that $Q = \{\vec{a} \in \mathbb{N}^n : \exists b R(\vec{a}, b)\}$.

Remark. Q primitive recursive $\Longrightarrow$ Q recursively enumerable.

Lemma 4.12

A nonempty set $Q \subseteq \mathbb{N}$ is recursively enumerable iff there exists an $f \in \mathrm{PR}^1$ with $Q = f(\mathbb{N})$.

Proof:

"$\Leftarrow$": If $f \in \mathrm{PR}^1$ then $R := \{(a, i) \in \mathbb{N}^2 : f(i) = a\}$ is prim. rec., and $f(\mathbb{N}) = \{a : \exists i R(a, i)\}$.

"$\Rightarrow$": Let $a_0 \in Q = \{a \in \mathbb{N} : \exists b R(a, b)\}$ with primitive recursive R.

Definition: $f : \mathbb{N} \to \mathbb{N}$, $f(c) := \begin{cases} \pi_1(c) & \text{if } R(\pi_1(c), \pi_2(c)) \\ a_0 & \text{otherwise} \end{cases}$.

Then $f \in \mathrm{PR}^1$ and $f(\mathbb{N}) \subseteq Q$. Remains to prove $Q \subseteq f(\mathbb{N})$:

$a \in Q \;\Rightarrow\; R(a, b)$ for some $b \in \mathbb{N} \;\Rightarrow\; R(\pi_1(c), \pi_2(c))$ with $c := \pi(a, b) \;\Rightarrow\; a = \pi_1(c) = f(c)$.

Inductive Definition of Σ-formulas

1. Each Δ_0-formula is a Σ-formula.
2. If A, B are Σ-formulas, then also $A \wedge B$, $A \vee B$, $\exists x A$ are Σ-formulas.
3. If A is a Σ-formula, and t a PR-term with $x \notin \mathrm{FV}(t)$, then $\forall x < t A$ is a Σ-formula.

A Σ-formula is called a Σ_1-*formula* if it is of the shape $\exists x C$ with $C \in \Delta_0$.

Lemma 4.13

A relation $Q \subseteq \mathbb{N}^n$ is recursive enumerable if, and only if, it can be defined by a Σ-formula, i.e., if there exists a Σ-formula A such that $\mathrm{FV}(A) \subseteq \{v_1, ..., v_n\}$ and $Q = \{(a_1, ..., a_n) \in \mathbb{N}^n : \mathcal{N} \models A[a_1, ..., a_n]\}$.

Proof:

From Lemma 4.2 it follows that a relation is r.e. iff it can be defined by a Σ_1-formula. By induction on the build up we now define for each Σ-formula A a Σ_1-Formel A' so that $\mathrm{FV}(A) = \mathrm{FV}(A')$ and $\mathcal{N} \models A \leftrightarrow A'$. This proves the Lemma.

1. For $A \in \Delta_0$ let $A' := \exists z A$ with $z \notin \mathrm{FV}(A)$.
2. Let $A' = \exists x \widetilde{A}$ and $B' = \exists y \widetilde{B}$. Then
   $(A \vee B)' := \exists z(\widetilde{A}_x(z) \vee \widetilde{B}_y(z))$, $(A \wedge B)' := \exists z(\widetilde{A}_x(\pi_1 z) \wedge \widetilde{B}_y(\pi_2 z))$, $(\exists v A)' := \exists z\, \widetilde{A}_{x,v}(\pi_1 z, \pi_2 z)$,
   $(\forall v < t\, A)' := \exists z \forall v < t \exists x < z\, \widetilde{A}$, where $z \notin \mathrm{vars}(\widetilde{A}) \cup \mathrm{vars}(\widetilde{B})$, $z \notin \mathrm{vars}(\widetilde{A})(\cup \mathrm{vars}(t))$, respectively.

Definition

1. A function $f : \mathbb{N}^n \to \mathbb{N}$ is *recursive* iff the relation
   $\mathsf{Graph}(f) := \{(\vec{a}, b) \in \mathbb{N}^{n+1} : f(\vec{a}) = b\}$ is recursively enumerable.
2. A relation $R \subseteq \mathbb{N}^n$ is *recursive*, if its characteristic function $\mathbf{1}_R$ is recursive.

Lemma 4.14

A relation $Q \subseteq \mathbb{N}^n$ is recursive iff Q and its complement $\mathbb{N}^n \setminus Q$ are recursively enumerable.

Proof:

1. Assume that Q and $\mathbb{N}^n \setminus Q$ are r.e.

$[\,\mathbf{1}_Q(\vec{a}) = b \;\Leftrightarrow\; (\vec{a} \in Q \wedge b = 1) \vee (\vec{a} \in \mathbb{N}^n \setminus Q \wedge b = 0)\,] \overset{4.13}{\Longrightarrow} \mathsf{Graph}(\mathbf{1}_Q)$ r.e. $\Longrightarrow \mathbf{1}_Q$ recursive.

2. Assume that Q is recursive. Then $G := \mathsf{Graph}(\mathbf{1}_Q)$ is r.e., and we have $[\,\vec{a} \in Q \;\Leftrightarrow\; (\vec{a}, 1) \in G\,]$ and $[\,\vec{a} \in \mathbb{N}^n \setminus Q \;\Leftrightarrow\; (\vec{a}, 0) \in G\,]$. Hence Q and $\mathbb{N}^n \setminus Q$ are r.e.

Corollary

$f : \mathbb{N}^n \to \mathbb{N}$ recursive $\Longrightarrow \mathsf{Graph}(f)$ recursive.

Proof: Let $G := \mathsf{Graph}(f)$.

f recursive $\Rightarrow$ G r.e. & $\mathbb{N}^{n+1} \setminus G = \{(\vec{a}, b) : \exists i((\vec{a}, i) \in G \wedge i \neq b)\} \Rightarrow$ G, $\mathbb{N}^{n+1} \setminus G$ r.e..

Lemma 4.15

The set of all recursive functions is the least set of functions $\mathcal{R} = \bigcup_{n \in \mathbb{N}} \mathcal{R}^n$ such that:

(R1) $\mathsf{C}^n_k, \mathsf{S}, \mathrm{I}^n_i \in \mathcal{R}$,

(R2) $h \in \mathcal{R}^m$ & $g_1, ..., g_m \in \mathcal{R}^n$ & $m, n \geq 1 \;\Rightarrow\; (\circ h g_1 ... g_m) \in \mathcal{R}^n$,

(R3) $g \in \mathcal{R}^n$ & $h \in \mathcal{R}^{n+1} \;\Rightarrow\; (\mathsf{R} g h) \in \mathcal{R}^{n+1}$,

(R4) $g \in \mathcal{R}^{n+1}$ & $\forall \vec{a} \in \mathbb{N}^n \exists i[g(\vec{a}, i) = 0]$ & $n \geq 1 \;\Rightarrow\; (\mu g) \in \mathcal{R}^n$,
where $(\mu g)(\vec{a}) := \min\{i \in \mathbb{N} : \; g(\vec{a}, i) = 0\}$.

Especially every primitive recursive function is recursive.

Proof:

Let $\mathcal{R}$ be the least set of functions satisfying (R1)–(R4), and let $\mathcal{R}'$ be the set of all recursive functions.

I. $\mathcal{R}' \subseteq \mathcal{R}$: Let $f : \mathbb{N}^n \to \mathbb{N}$ and $\mathsf{Graph}(f)$ r.e.

Then there exists an $h \in \mathrm{PR}^{n+1}$ with $(f(\vec{a}) = b \Leftrightarrow \exists i\, h(\vec{a}, b, i) = 0)$. Hence

$f(\vec{a}) = \pi_1(\min\{j : h(\vec{a}, \pi_1(j), \pi_2(j)) = 0\})$ and thus $f = (\circ \pi_1(\mu g))$, where $g(\vec{a}, j) := h(\vec{a}, \pi_1(j), \pi_2(j))$.

II. $\mathcal{R} \subseteq \mathcal{R}'$:

Here it suffices to prove that the closure conditions (R1)–(R4) hold for $\mathcal{R}'$ (in place of $\mathcal{R}$).

But this follows by Lemma 4.13 from the following equivalences:

$(\circ h g_1 ... g_m)(\vec{a}) = b \;\Leftrightarrow\; \exists b_1 \ldots \exists b_m [h(b_1, ..., b_m) = b \wedge g_1(\vec{a}) = b_1 \wedge \ldots \wedge g_m(\vec{a}) = b_m]$.

$(\mathsf{R} g h)(\vec{a}, k) = b \;\Leftrightarrow\; \exists c[g(\vec{a}) = (c)_0 \wedge b = (c)_k \wedge \forall i < k(\, h(\vec{a}, i, (c)_i) = (c)_{i+1}\,)]$.

$(\mu g)(\vec{a}) = b \;\Leftrightarrow\; g(\vec{a}, b) = 0 \wedge \forall i < b \exists c[c \neq 0 \wedge g(\vec{a}, i) = c]$.

Church's Thesis

A function $f : \mathbb{N}^n \to \mathbb{N}$ (relation $Q \subseteq \mathbb{N}^n$) is computable (decidable) in the intuitive sense iff it is recursive.

5 Gödel's 1st Incompleteness Theorem

We start with an *arithmetization of syntax.*

Definition

A *primitive recursively represented language* is a pair $(\mathcal{L}, SN)$
consisting of a 1st order language $\mathcal{L}$ and a one-one mapping $SN : \mathcal{L} \to \mathbb{N}$ such that:
– For each n-ary functions symbol $f \in \mathcal{L}$ we have $SN(f) = \langle n, 1, i\rangle$ with $i \in \mathbb{N}$.
– For each n-ary relation symbol $p \in \mathcal{L}$ we have $SN(p) = \langle n, 2, i\rangle$ with $i \geq 1$.
– The set $SN(\mathcal{L}) := \{SN(p) : p \in \mathcal{L}\}$ is primitive recursive.

($SN(p)$ is called the *symbol number* of p.

Convention

In the sequel $(\mathcal{L}, SN)$ denotes a fixed primitive recursively represented language.
If nothing else is said, function- and relation symbols are elements of $\mathcal{L}$.
The notions "term", "formula" etc. always refer to $\mathcal{L}$. – Abb.: $(a)_{i,j} := ((a)_i)_j$.

Definition

$SN(v_i) := \langle 0, 0, i\rangle$, $SN(\to) := \langle 2, 0, 0\rangle$, $SN(\forall) := \langle 2, 0, 1\rangle$, $SN(\bot) := \langle 0, 2, 0\rangle$, $SN(\approx) := \langle 2, 2, 0\rangle$.

Definition of the *Gödel number* $\ulcorner u \urcorner$ of an expression u

$\ulcorner v_i \urcorner := \langle SN(v_i)\rangle$, $\ulcorner p u_1 ... u_n \urcorner := \langle SN(p), \ulcorner u_1 \urcorner, ..., \ulcorner u_n \urcorner\rangle$, $\ulcorner \forall x A \urcorner := \langle SN(\forall), \ulcorner x \urcorner, \ulcorner A \urcorner\rangle$

Remark $\ulcorner u \urcorner = \ulcorner u' \urcorner \Rightarrow u = u'$.

Definition

$A_{x_1,...,x_n}(t_1, ..., t_n)$ denotes the formula resulting from A when every free occurrence of x_i is replaced by the term t_i, simultaneously for $i = 1, ..., n$. If $x_1, ..., x_n$ are known from the context we briefly write $A(t_1, ..., t_n)$.

A formula A is called *n-ary*, if $\mathrm{FV}(A) \subseteq \{v_1, ..., v_n\}$. In this case $A(t_1, ..., t_n) := A_{v_1,...,v_n}(t_1, ..., t_n)$.

PRIM := set of all atomic (or prime) formulas
FOR := set of all formulas
FOR^n := set of all n-ary formulas.

Definition

If X is a set of expressions then $\ulcorner X \urcorner := \{\ulcorner u \urcorner : u \in X\}$.
A set X of expressions is called *primitive recursive* (*recursive*, *recursive enumerable*, resp.)
if the set $\ulcorner X \urcorner$ has this property.

$\ulcorner \mathrm{FV} \urcorner := \{(\ulcorner x \urcorner, \ulcorner u \urcorner) : u \text{ is an expression and } x \in \mathrm{FV}(u)\}$
$\ulcorner \mathsf{subst} \urcorner := \{(\ulcorner A \urcorner, \ulcorner x \urcorner, \ulcorner t \urcorner) : t \in \mathrm{TER}, x \in \mathrm{VARS}, A \in \mathrm{FOR}, \mathsf{subst}(A, x, t)\}$

Lemma 5.1

The relations $\ulcorner \mathrm{VARS} \urcorner$, $\ulcorner \mathrm{TER} \urcorner$, $\ulcorner \mathrm{PRIM} \urcorner$, $\ulcorner \mathrm{FOR} \urcorner$, $\ulcorner \mathrm{FV} \urcorner$, $\ulcorner \mathrm{FOR}^n \urcorner$, $\ulcorner \mathsf{subst} \urcorner$ are primitive recursive.

Proof:

The assertion follows by Lemmata 4.2, 4.11 from the following equivalences:

1. $a \in \ulcorner \textsc{Vars} \urcorner \;\Leftrightarrow\; a = \langle\langle 0, 0, (a)_{0,2}\rangle\rangle$

2. $a \in \ulcorner \textsc{Ter} \urcorner \;\Leftrightarrow\; a \in \ulcorner \textsc{Vars} \urcorner \vee [\,(a)_0 = \langle \mathrm{lh}(a) \dot{-} 1, 1, (a)_{0,2}\rangle \in SN(\mathcal{L}) \wedge \forall i < \mathrm{lh}(a) \dot{-} 1((a)_{i+1} \in \ulcorner \textsc{Ter} \urcorner)\,]$
$\Leftrightarrow\; a \in \ulcorner \textsc{Vars} \urcorner \vee [\,(a)_0 = \langle \mathrm{lh}(a) \dot{-} 1, 1, (a)_{0,2}\rangle \in SN(\mathcal{L}) \wedge \forall i < \mathrm{lh}(a) \dot{-} 1(\,(\overline{\mathbf{1}_{\ulcorner \textsc{Ter} \urcorner}}(a))_{(a)_{i+1}} = 1\,)\,]$

3. $a \in \ulcorner \textsc{Prim} \urcorner \;\Leftrightarrow$
$\mathrm{lh}(a) > 0 \wedge \forall i < \mathrm{lh}(a) \dot{-} 1((a)_{i+1} \in \ulcorner \textsc{Ter} \urcorner) \wedge (a)_0 = \langle \mathrm{lh}(a) \dot{-} 1, 2, (a)_{0,2}\rangle \in SN(\mathcal{L}) \cup \{SN(\approx), SN(\bot)\}$

4. $a \in \ulcorner \textsc{For} \urcorner \;\Leftrightarrow$
$a \in \ulcorner \textsc{Prim} \urcorner \vee$
$[\,\mathrm{lh}(a) = 3 \wedge (a)_0 = SN(\to) \wedge (a)_1 \in \ulcorner \textsc{For} \urcorner \wedge (a)_2 \in \ulcorner \textsc{For} \urcorner\,] \vee$
$[\,\mathrm{lh}(a) = 3 \wedge (a)_0 = SN(\forall) \wedge (a)_1 \in \ulcorner \textsc{Vars} \urcorner \wedge (a)_2 \in \ulcorner \textsc{For} \urcorner\,]$

5. $(j, a) \in \ulcorner \mathrm{FV} \urcorner \;\Leftrightarrow$
$(a \in \ulcorner \textsc{Vars} \urcorner \wedge j = a) \vee$
$[\,a \in \ulcorner \textsc{For} \urcorner \cup \ulcorner \textsc{Ter} \urcorner \wedge \exists i < \mathrm{lh}(a) \dot{-} 1((j, (a)_{i+1}) \in \ulcorner \mathrm{FV} \urcorner) \wedge ((a)_0 = SN(\forall) \to j \neq (a)_1)\,]$

6. $a \in \ulcorner \textsc{For}^n \urcorner \;\Leftrightarrow\; a \in \ulcorner \textsc{For} \urcorner \wedge \forall i < a((i, a) \in \ulcorner \mathrm{FV} \urcorner \to 1 \leq (i)_{0,2} \leq n)$

7. $(a, j, c) \in \ulcorner \mathsf{subst} \urcorner \;\Leftrightarrow$
$a \in \ulcorner \textsc{For} \urcorner \wedge j \in \ulcorner \textsc{Vars} \urcorner \wedge c \in \ulcorner \textsc{Ter} \urcorner \wedge$
$(a \in \ulcorner \textsc{Prim} \urcorner \vee$
$[\,(a)_0 = SN(\to) \wedge ((a)_1, j, c) \in \ulcorner \mathsf{subst} \urcorner \wedge ((a)_2, j, c) \in \ulcorner \mathsf{subst} \urcorner\,] \vee$
$[\,(a)_0 = SN(\forall) \wedge (\,(j, a) \in \ulcorner \mathrm{FV} \urcorner \to ((a)_1, c) \notin \ulcorner \mathrm{FV} \urcorner \wedge ((a)_2, j, c) \in \ulcorner \mathsf{subst} \urcorner)]).$

Definition ($\mathsf{Sub} : \mathbb{N}^3 \to \mathbb{N}$)

$$\mathsf{Sub}(a, c_1, c_2) := \begin{cases} c_2 & \text{if } a = c_1 \\ a & \text{if } a = \langle SN(\forall), c_1, (a)_2\rangle \\ \langle (a)_0, \mathsf{Sub}((a)_1, c_1, c_2), \ldots, \mathsf{Sub}((a)_{\mathrm{lh}(a)-1}, c_1, c_2)\rangle & \text{otherwise.} \end{cases}$$

Lemma 5.2

(a) If $u \in \textsc{Ter} \cup \textsc{For}$, $x \in \textsc{Vars}$, $t \in \textsc{Ter}$, then $\mathsf{Sub}(\ulcorner u \urcorner, \ulcorner x \urcorner, \ulcorner t \urcorner) = \ulcorner u_x(t) \urcorner$.

(b) The function Sub is primitive recursive.

Proof:

(a) Induction on the build up of u: Abb.: $\phi(u) := \mathsf{Sub}(\ulcorner u \urcorner, \ulcorner x \urcorner, \ulcorner t \urcorner)$

1. $u = x$: $\phi(u) = \ulcorner t \urcorner = \ulcorner u_x(t) \urcorner$.

2. $u = y \neq x$: $\phi(u) = \langle SN(y)\rangle = \ulcorner y \urcorner = \ulcorner u_x(t) \urcorner$.

3. $u = pu_1 \ldots u_n$ with $p \in \mathcal{L} \cup \{\approx, \bot, \to\}$: $\phi(u) = \langle SN(p), \phi(u_1), \ldots, \phi(u_n)\rangle \stackrel{\text{I.H.}}{=}$
$\langle SN(p), \ulcorner u_1(x/t) \urcorner, \ldots, \ulcorner u_n(x/t) \urcorner\rangle = \ulcorner pu_1(x/t) \ldots u_n(x/t) \urcorner = \ulcorner u(x/t) \urcorner$.

4. $u = \forall y A$ with $y \neq x$: $\phi(u) = \langle SN(\forall), \phi(y), \phi(A)\rangle \stackrel{\text{I.H.}}{=} \langle SN(\forall), \ulcorner y \urcorner, \ulcorner A_x(t) \urcorner\rangle = \ulcorner \forall y\, A_x(t) \urcorner = \ulcorner u_x(t) \urcorner$.

5. $u = \forall x A$: $\phi(u) = \ulcorner u \urcorner = \ulcorner u_x(t) \urcorner$.

(b) Define a prim. rec. function g by $g(a, d, 0) := \langle (a)_0\rangle$, $g(a, d, k+1) := g(a, d, m) * \langle (d)_{(a)_{k+1}}\rangle$.

Then $g(a, d, k) = \langle (a)_0, (d)_{(a)_1}, \ldots, (d)_{(a)_k}\rangle$ and $\mathsf{Sub}(a, c_1, c_2) = f(c_1, c_2, a)$ where

$$f(c_1, c_2, a) := h(c_1, c_2, a, \overline{f}(c_1, c_2, a)) \text{ with } h(c_1, c_2, a, d) := \begin{cases} c_2 & \text{if } a = c_1 \\ a & \text{if } a = \langle SN(\forall), c_1, (a)_2\rangle \\ g(a, d, \mathrm{lh}(a) \dot{-} 1) & \text{otherwise.} \end{cases}$$

[$\mathsf{Sub}(a,c_1,c_2) = f(c_1,c_2,a)$ is proved by induction on a. In the case "otherwise" we compute:
$\mathsf{Sub}(c_1,c_2,a) \stackrel{\text{IH}}{=} \langle (a)_0, f(c_1,c_2,(a)_1), \ldots, f(c_1,c_2,(a)_{\mathrm{lh}(a)-1})\rangle =$
$\langle (a)_0, \big(\overline{f}(c_1,c_2,a)\big)_{(a)_1}, \ldots, \big(\overline{f}(c_1,c_2,a)\big)_{(a)_{\mathrm{lh}(a)-1}}\rangle = g(a,\overline{f}(c_1,c_2,a),\mathrm{lh}(a) \dot{-} 1) = h(c_1,c_2,a,\overline{f}(a,c_1,c_2)).$]

Definition

For each axiom system Φ let $\mathsf{Prf}_\Phi := \{(\ulcorner A \urcorner, \langle \ulcorner A_0 \urcorner, ..., \ulcorner A_n \urcorner\rangle) : (A_0, ..., A_n)$ is a derivation of A from $\Phi\}$.

Theorem 5.3

If the axiom system Φ is primitive recursive (recursive, r.e., resp.) then Prf_Φ has this property too.

Proof:

$\mathsf{Prf}_\Phi(a,c) \Leftrightarrow$
$\mathrm{lh}(c) \geq 1 \wedge a = (c)_{\mathrm{lh}(c)\dot{-}1} \wedge \forall k < \mathrm{lh}(c)[(c)_k \in \ulcorner \Phi \urcorner \cup \ulcorner \mathsf{AX} \urcorner \vee \exists i,j < k((c)_j = \langle SN(\rightarrow),(c)_i,(c)_k\rangle)]$

It remains to prove that AX is primitive recursive.

HS 1: The set AX_0 of all logical axioms not starting with $\forall$ is primitive recursive.

Proof: Abb.: $(a \dot{\rightarrow} b) := \langle SN(\rightarrow), a, b\rangle$.

Obviously the following holds: $k \in \mathsf{AX}_0 \Leftrightarrow k \in \ulcorner \textsc{For} \urcorner$ and one of the cases 1.–9. applies.

1. $\exists a < k[\, k = (a \dot{\rightarrow} a)\,]$
2. $\exists a,b < k[\, k = (a \dot{\rightarrow} (b \dot{\rightarrow} a))\,]$
3. $\exists a,b,c < k[\, k = ((c \dot{\rightarrow} (a \dot{\rightarrow} b)) \dot{\rightarrow} ((c \dot{\rightarrow} a) \dot{\rightarrow} (c \dot{\rightarrow} b)))\,]$
4. $\exists a < k[\, k = (((a \dot{\rightarrow} \ulcorner \bot \urcorner) \dot{\rightarrow} \ulcorner \bot \urcorner) \dot{\rightarrow} a) \wedge a \in \ulcorner \textsc{Prim} \urcorner\,]$
5. $\exists a,i,c < k[\, k = (\langle SN(\forall), i, a\rangle \dot{\rightarrow} \mathsf{Sub}(a,i,c)) \wedge (a,i,c) \in \ulcorner \mathsf{subst} \urcorner)\,]$
6. $\exists a,b,i < k[\, k = (\langle SN(\forall), i, (a \dot{\rightarrow} b)\rangle \dot{\rightarrow} (\langle SN(\forall), i, a\rangle \dot{\rightarrow} \langle SN(\forall), i, b\rangle))\,]$
7. $\exists a,i < k[\, k = (a \dot{\rightarrow} \langle SN(\forall), i, a\rangle) \wedge (i,a) \notin \ulcorner \mathrm{FV} \urcorner\,]$
8. $\exists c < k[\, k = \langle SN(\approx), c, c\rangle\,]$
9. $\exists a,c,d < k[\, a \in \ulcorner \textsc{Prim} \urcorner \wedge c,d \in \ulcorner \textsc{Vars} \urcorner \wedge k = (\langle SN(\approx), c, d\rangle \dot{\rightarrow} (a \dot{\rightarrow} \mathsf{Sub}(a,c,d)))\,]$.

HS 2: AX is primitive recursive.

Proof: $a \in \ulcorner \mathsf{AX} \urcorner \Leftrightarrow a \in \ulcorner \mathsf{AX}_0 \urcorner \vee (\mathrm{lh}(a) = 3 \wedge (a)_0 = SN(\forall) \wedge (a)_1 \in \ulcorner \textsc{Vars} \urcorner \wedge (a)_2 \in \ulcorner \mathsf{AX} \urcorner)$.

Theorem 5.4

If Φ is a recursively enumerable axiom system,
then the set $\{A : \Phi \vdash A\}$ of all logical consequences from Φ is also recursively enumerable.

Proof:

$a \in \{\ulcorner A \urcorner : \Phi \vdash A\} \Leftrightarrow$ there exists a derivation $(A_0, ..., A_n)$ from Φ with $a = \ulcorner A_n \urcorner \Leftrightarrow \exists b\, \mathsf{Prf}_\Phi(a,b)$.

Lemma 5.5

Let T be a theory with primitive recursively represented language $\mathrm{L}(T)$.

Then the following statements are equivalent:

(i) T is recursively enumerable.

(ii) T has a primitive recursive axiom system.

(iii) T has a recursive axiom system.

(iv) T has a recursively enumerable axiom system.

Definition: Such a theory is called *recursively axiomatizable.*

Proof: w.l.o.g. $\mathrm{L}(T) = \mathcal{L}$.

(i)⇒(ii): Let $f \in \mathrm{PR}^1$ with $\ulcorner T \urcorner = f(\mathbb{N})$.

Def.: $A_n :=$ the formula with $\ulcorner A_n \urcorner = f(n)$, $B_0 := A_0$, $B_{n+1} := B_n \wedge A_{n+1}$; $\Phi := \{B_n : n \in \mathbb{N}\}$

1. Φ is an axiom system for T: obvious.

2. Φ ist prim. rec.: As one easily sees, the function $n \mapsto \ulcorner B_n \urcorner$ is prim. rec., and $n < \ulcorner B_n \urcorner$ holds for all $n \in \mathbb{N}$. From the latter we conclude: $a \in \ulcorner \Phi \urcorner \Leftrightarrow \exists n < a (a = \ulcorner B_n \urcorner)$.

(ii)⇒(iii) and (iii)⇒(iv): trivial.

(iv)⇒(i): Theorem 5.4 + "For^0 is prim. rec." + $T = \{A \in \mathrm{For}^0 : \Phi \vdash A\}$.

Satz 5.6

Each recursively axiomatizable, complete theory T is recursive.

Proof: w.l.o.g. $\mathrm{L}(T) = \mathcal{L}$.

If T is inconsistent, then $T = \mathrm{For}^0$. If T is consistent then $a \notin \ulcorner T \urcorner \Leftrightarrow (a \notin \ulcorner \mathrm{For}^0 \urcorner \vee \langle SN(\rightarrow), a, \ulcorner \bot \urcorner \rangle \in \ulcorner T \urcorner)$, and the assertion follows by Lemmata 5.5, 4.14.

Definition

Let $\mathcal{M}$ be an $\mathcal{L}$-structure. A relation $R \subseteq |\mathcal{M}|^n$ is *definable in* $\mathcal{M}$, if there is an n-ary formula A, such that $R = \{(a_1, .., a_n) \in |\mathcal{M}|^n : \mathcal{M} \models A[a_1, ..., a_n]\}$.

Convention. In the following we assume that $\mathcal{L}$ contains at least the function symbols $\mathsf{0}$ and S.

Definition of the terms $\underline{n}$ ($n \in \mathbb{N}$): $\underline{0} := \mathsf{0}$, $\underline{n+1} := \mathsf{S}\underline{n}$. (These terms are called *numerals.*)

Remark

If $\mathcal{M}$ is an $\mathcal{L}$-structure with $|\mathcal{M}| = \mathbb{N}$, $\mathsf{0}^{\mathcal{M}} = 0$, $\mathsf{S}^{\mathcal{M}}(a) = a + 1$, then for each n-ary formula A and all $a_1, ..., a_n \in \mathbb{N}$ the following holds: $\mathcal{M} \models A[a_1, ..., a_n] \Leftrightarrow \mathcal{M} \models A(\underline{a_1}, ..., \underline{a_n})$.

Definition of the function $\mathbf{s} \in \mathrm{PR}^2$: $\mathbf{s}(a, k) := \mathsf{Sub}(a, \ulcorner v_1 \urcorner, \ulcorner \underline{k} \urcorner)$.

Theorem 5.7 (Tarski)

Let $\mathcal{M}$ be an $\mathcal{L}$-Struktur with $|\mathcal{M}| = \mathbb{N}$, $\mathsf{0}^{\mathcal{M}} = 0$, $\mathsf{S}^{\mathcal{M}}(a) = a + 1$.

Assume that every primitive recursive relation (thence also every r.e. relation) is definable in $\mathcal{M}$.

Then $\ulcorner \mathrm{Th}(\mathcal{M}) \urcorner$ is **not** definable in $\mathcal{M}$ and therefore $\mathrm{Th}(\mathcal{M})$ is not r.e. and not recursively axiomatizable (cf. Lemma 5.5).

Proof:

Preliminary remark:

If X is a set and H a function with $\mathrm{dom}(H) = X$, then $Q := \{x \in X : x \notin H(x)\} \notin H(X)$.

(Proof: If $Q = H(x_0)$ then $x_0 \in H(x_0) \Leftrightarrow x_0 \in Q \Leftrightarrow x_0 \notin H(x_0)$.)

Now let $H : \ulcorner \mathrm{For}^1 \urcorner \to \mathcal{P}(\mathbb{N})$, $H(\ulcorner A \urcorner) := \{k \in \mathbb{N} : \mathcal{M} \models A(\underline{k})\}$ and $Q := \{k \in \ulcorner \mathrm{For}^1 \urcorner : k \notin H(k)\}$.
Then we have $Q \notin H(\ulcorner \mathrm{For}^1 \urcorner)$, i.e., Q is not definable in $\mathcal{M}$.

HS 1: $k \in Q \Leftrightarrow k \in \ulcorner \mathrm{For}^1 \urcorner \wedge \mathsf{s}(k,k) \notin \ulcorner \mathrm{Th}(\mathcal{M}) \urcorner$.
Proof: Assume $k = \ulcorner A \urcorner \in \ulcorner \mathrm{For}^1 \urcorner$ (otherwise the claim is trivial).
$k \in Q \Leftrightarrow k \notin H(\ulcorner A \urcorner) \Leftrightarrow \mathcal{M} \not\models A(\underline{k}) \Leftrightarrow A(\underline{k}) \notin \mathrm{Th}(\mathcal{M}) \Leftrightarrow \mathsf{Sub}(\ulcorner A \urcorner, \ulcorner v_1 \urcorner, \ulcorner \underline{k} \urcorner) \notin \ulcorner \mathrm{Th}(\mathcal{M}) \urcorner$

HS 2: $\ulcorner \mathrm{Th}(\mathcal{M}) \urcorner$ definable in $\mathcal{M} \implies Q$ definable in $\mathcal{M}$.
Proof: *Assumption:* D is a 1-ary formula with $\ulcorner \mathrm{Th}(\mathcal{M}) \urcorner = \{k \in \mathbb{N} : \mathcal{M} \models D(\underline{k})\}$.

From the assumption that every primitive recursive relation is definable in $\mathcal{M}$ it follows that there exists a 2-ary formula A with $\mathcal{M} \models A(\underline{k},\underline{b}) \Leftrightarrow k \in \ulcorner \mathrm{For}^1 \urcorner \,\&\, \mathsf{s}(k,k) = b$.

Hence: $k \in Q \overset{\mathrm{HS1}}{\Leftrightarrow} k \in \ulcorner \mathrm{For}^1 \urcorner \,\&\, \mathsf{s}(k,k) \notin \ulcorner \mathrm{Th}(\mathcal{M}) \urcorner$
$\Leftrightarrow$ there exists a b with $k \in \ulcorner \mathrm{For}^1 \urcorner \,\&\, \mathsf{s}(k,k) = b \,\&\, b \notin \ulcorner \mathrm{Th}(\mathcal{M}) \urcorner$
$\Leftrightarrow$ there exists a b with $\mathcal{M} \models A(\underline{k},\underline{b})$ and $b \notin \ulcorner \mathrm{Th}(\mathcal{M}) \urcorner$
$\Leftrightarrow \mathcal{M} \models C(\underline{k})$, where $C := \exists y(A(v_1,y) \wedge \neg D(y))$.

Since Q is not definable in $\mathcal{M}$, HS2 yields that $\ulcorner \mathrm{Th}(\mathcal{M}) \urcorner$ is not definable in $\mathcal{M}$.

Remark.
Let A, C be as in the proof of HS 2, but with D now an *arbitrary* 1-ary formula.
Let further $k := \ulcorner C \urcorner$. Then $\ulcorner C(\underline{k}) \urcorner = \mathsf{Sub}(k, \ulcorner v_1 \urcorner, \ulcorner \underline{k} \urcorner) = \mathsf{s}(k,k)$ and
$\mathcal{M} \models C(\underline{k}) \Leftrightarrow$ there exists a b with $\mathcal{M} \models A(\underline{k},\underline{b}) \wedge \neg D(\underline{b}) \Leftrightarrow \mathcal{M} \models \neg D(\underline{\mathsf{s}(k,k)})$.

For $G := C(\underline{k})$ we therefore have: $\mathcal{M} \models G \leftrightarrow \neg D(\underline{\ulcorner G \urcorner})$.

Definition (Representability)
Let Φ be an axiom system with $\mathsf{0}, \mathsf{S} \in \mathrm{L}(\Phi)$. An $(n{+}1)$-ary formula A *represents* the function $f : \mathbb{N}^n \to \mathbb{N}$ in Φ, if the following holds for all $a_1, \ldots, a_n, b \in \mathbb{N}$:
(1) $f(a_1,\ldots,a_n) = b \implies \Phi \vdash A(\underline{a_1},\ldots,\underline{a_n},\underline{b}) \wedge \forall y(A(\underline{a_1},\ldots,\underline{a_n},y) \to \underline{b} \approx y)$.
(2) $f(a_1,\ldots,a_n) \neq b \implies \Phi \vdash \neg A(\underline{a_1},\ldots,\underline{a_n},\underline{b})$.

(*Remark:* If $\Phi \vdash \neg(\underline{b} \approx \underline{c})$ holds for all $b, c \in \mathbb{N}$ with $b \neq c$, then (2) is a consequence of (1).)

An n-ary formula A *represents* the relation $R \subseteq \mathbb{N}^n$ in Φ, if the following holds for all $a_1, \ldots, a_n \in \mathbb{N}$:
(1) $(a_1,\ldots,a_n) \in R \implies \Phi \vdash A(\underline{a_1},\ldots,\underline{a_n})$.
(2) $(a_1,\ldots,a_n) \notin R \implies \Phi \vdash \neg A(\underline{a_1},\ldots,\underline{a_n})$.

$f : \mathbb{N}^n \to \mathbb{N}$ ($R \subseteq \mathbb{N}^n$, resp.) is *representable in* Φ, if there is an $\mathrm{L}(\Phi)$-formula A such that
A represents f (R, resp.) in Φ.

$R \subseteq \mathbb{N}^n$ is *weakly representable in* Φ, if there is an n-ary $\mathrm{L}(\Phi)$-formula A such that
$R = \{(a_1,\ldots,a_n) \in \mathbb{N}^n : \Phi \vdash A(\underline{a_1},\ldots,\underline{a_n})\}$.

Theorem 5.8 (Fixpoint Lemma)
If the primitive recursive function $a \mapsto \mathsf{s}(a,a)$ is representable Φ,
then for every 1-ary $\mathrm{L}(\Phi)$-formula D one can construct an $\mathrm{L}(\Phi)$-sentence G so that $\Phi \vdash G \leftrightarrow D(\underline{\ulcorner G \urcorner})$.

Proof:

Let A be a 2-ary $\mathrm{L}(\Phi)$-formula which represents the function $a \mapsto \mathbf{s}(a,a)$ in Φ, and let

$C := \exists y(A(v_1, y) \wedge D(y))$. – Then for all $k \in \mathbb{N}$ we have $\Phi \vdash C(\underline{k}) \leftrightarrow D(\underline{\mathbf{s}(k,k)})$.

[Proof: Let $b := \mathbf{s}(k,k)$.

Then $\Phi \vdash A(\underline{k}, \underline{b}) \wedge \forall y(A(\underline{k}, y) \rightarrow \underline{b} = y)$, and from this we obtain $\Phi \vdash \exists y(A(\underline{k}, y) \wedge D(y)) \leftrightarrow D(\underline{b})$.]

Morover, for $k := \ulcorner C \urcorner$, $G := C(\underline{k})$ we have $\mathbf{s}(k,k) = \ulcorner C(\underline{k}) \urcorner = \ulcorner G \urcorner$. Hence $\Phi \vdash G \leftrightarrow D(\underline{\ulcorner G \urcorner})$.

Theorem 5.9

Every consistent theory T, in which all recursive functions are representable,
is undecidable (i.e., not recursive).

Proof:

Assumption: T recursive.

Then the characteristic function $\mathbf{1}_{\ulcorner T \urcorner}$ and thus the 1-ary relation $\ulcorner T \urcorner$ are representable in T.

Therefor there is a 1-ary $\mathrm{L}(T)$-formula D so that the following holds for each sentence A:

(1) $A \in T \;\Rightarrow\; T \vdash D(\underline{\ulcorner A \urcorner})$, (2) $A \notin T \;\Rightarrow\; T \vdash \neg D(\underline{\ulcorner A \urcorner})$.

By the Fixpoint Lemma there is an $\mathrm{L}(T)$-sentence G with (3) $T \vdash G \leftrightarrow \neg D(\underline{\ulcorner G \urcorner})$.

Now, from (1),(2),(3) we obtain ($G \in T \;\Rightarrow\; T \vdash \bot$) and ($G \notin T \;\Rightarrow\; G \in T$); hence $T \vdash \bot$. *Contradiction.*

Definition

An axiom system Φ with $\mathsf{0}, \mathsf{S} \in \mathrm{L}(\Phi)$ is called *ω-consistent*, if the following holds for each 1-ary $\mathrm{L}(\Phi)$-formula A:

$\Phi \vdash \neg A(\underline{n})$, for all $n \in \mathbb{N} \;\Rightarrow\; \Phi \nvdash \exists x A(x)$.

Remark: "ω-consistent" implies "consistent".

Satz 5.10 (Gödel's 1. Incompleteness Theorem)

(Version A) Every consistent, recursively axiomatizable theory T, in which all recursive functions are representable, is incomplete.

(Version B) For each ω-consistent, recursive axiom system Φ, in which all recursive functions are representable, one can construct an $\mathrm{L}(\Phi)$-sentence G, so that $\Phi \nvdash G$ and $\Phi \nvdash \neg G$.

Proof:

(A) The claim follows from theorems 5.6, 5.9.

(B) w.l.o.g.: $\mathcal{L} = \mathrm{L}(\Phi)$. According to Theorem 5.3 and the assumption that all recursive functions (thence all recursive relations) are representable in Φ, we can construct a 2-ary $\mathcal{L}$-formula P such that:

(1) $(a,b) \in \mathsf{Prf}_\Phi \;\Rightarrow\; \Phi \vdash P(\underline{a}, \underline{b})$,

(2) $(a,b) \notin \mathsf{Prf}_\Phi \;\Rightarrow\; \Phi \vdash \neg P(\underline{a}, \underline{b})$.

From the Fixpoint Lemma we obtain an $\mathcal{L}$-sentence G such that:

(3) $\Phi \vdash G \leftrightarrow \neg \exists z P(\underline{\ulcorner G \urcorner}, y)$.

Assumption: $\Phi \vdash G$.

Now by (1) there exists a b with $\Phi \vdash P(\underline{\ulcorner G \urcorner}, \underline{b})$. From this we get $\Phi \vdash \exists z P(\underline{\ulcorner G \urcorner}, z)$ and then, by (3), $\Phi \vdash \neg G$. Contradiction.

Therefore we have $\Phi \nvdash G$. By (2) from this we get $\Phi \vdash \neg P(\underline{\ulcorner G \urcorner}, \underline{b})$ for all $b \in \mathbb{N}$.

From (3) and the ω-consistency of Φ we further obtain $\Phi \nvdash \neg G$.

6 Representability

Inductive Definition of sets REK^n of n-ary functions

1. $\mathsf{C}_k^n, \mathrm{I}_i^n \in \text{REK}^n$, $\mathsf{S} \in \text{REK}^1$, $+, \cdot, \mathbf{1}_< \in \text{REK}^2$.
2. $h \in \text{REK}^m$ & $g_1, ..., g_m \in \text{REK}^m$ & $m, n \geq 1 \implies (\circ h g_1 ... g_m) \in \text{REK}^n$.
3. $g \in \text{REK}^{n+1}$ & $\forall \vec{a} \in \mathbb{N}^n \exists i[g(\vec{a}, i) = 0]$ & $n \geq 1 \implies (\mu g) \in \text{REK}^n$.

Abbreviation: $\text{REK} := \bigcup_{n \in \mathbb{N}} \text{REK}^n$

We will now prove that REK is closed under primitive recursion and thus (by Lemma 4.15) coincides with the set of all recursive functions.

Abbreviation: $\text{Rest}_b^a(i) :=$ remainder of a divided by $b(i+1)+1$

Lemma 6.1

The functions $\pi, \pi_1, \pi_2, \dot{-}$ and $(a, b, i) \mapsto \text{Rest}_b^a(i)$ are in REK.

Proof:

1. $\pi(a, b) = b + \sum_{i<a+b}(i+1) = b + \frac{1}{2}(a+b)(a+b+1) = b + H((a+b) \cdot (a+b+1))$,
where $H(c) := \min\{i : c \leq 2i\} = \min\{i : \mathbf{1}_<(2i, c) = 0\}$.

From this we get $\pi \in \text{REK}$, since REK is closed under explicit definitions and the μ-operator.

2. Let $J(c) := \min\{i : 2c < (i+1) \cdot (i+2)\}$. Then $J \in \text{REK}$ and
(1) $J(c) \cdot (J(c)+1) \leq 2c < (J(c)+1) \cdot (J(c)+2)$.

On the other side:
(2) $(a+b) \cdot (a+b+1) \leq (a+b) \cdot (a+b+1) + 2b < (a+b+1) \cdot (a+b+2)$.

Assume now that $c = \pi(a, b)$, hence $2c = (a+b) \cdot (a+b+1) + 2b$. From (1),(2) we get $J(c) = a+b$ and thus $\pi_2(c) = b = c \dot{-} H(J(c)(J(c)+1))$ and $\pi_1(c) = a = J(c) \dot{-} \pi_2(c)$.

3. $a \dot{-} b = \min\{c : a \leq b + c\}$.

4. $\text{Rest}_b^a(i) = a \dot{-} f(a, b, i) \cdot (b(i+1)+1)$, where $f(a, b, i) := \min\{k : a < (k+1)(b(i+1)+1)\}$.

Lemma 6.2

For arbitrary $k, m_0, ..., m_k \in \mathbb{N}$ there exist $a, b \in \mathbb{N}$ such that $\text{Rest}_b^a(i) = m_i$ for $i = 0, ..., k$.

Proof:

Let $s := \max\{k, m_0, ..., m_k\} + 1$, $b := s!$, $b_i := b(i+1)+1$. Then $m_i < b_i$, for $i = 0, ..., k$.

HS 1: $i < j \leq k \Rightarrow b_i, b_j$ are relatively prime.

Proof: Let p be a prime number with $p | b_i$ and $p | b_j$. Then $p | (b_j - b_i)$, i.e., $p | (j-i)b$. From $j - i \leq s$ and $b = s!$ we get $(j-i) | b$, hence $p | b$. Together with $p | (b(j+1)+1)$ this yields $p = 1$.

HS 2: $0 \leq a < a' < b_0 \cdot ... \cdot b_k \Rightarrow \exists i \leq k (\text{Rest}_b^a(i) \neq \text{Rest}_b^{a'}(i))$.

Proof by contradiction: Assume $r_i := \text{Rest}_b^a(i) = \text{Rest}_b^{a'}(i)$ for $i = 0, ..., k$.
Then $a = b_i n_i + r_i$ & $a' = b_i n_i' + r_i$ & $r_i < b_i$ with $n_i < n_i'$. Hence $a' - a = (n_i - n_i')b_i$, i.e., $b_i | (a' - a)$ for $i = 0, ..., k$. By HS 1 from this we get $b_0 \cdot ... \cdot b_k | (a' - a)$ which contradicts $a' - a < b_0 \cdot ... \cdot b_k$.

By HS 2 the mapping $a \mapsto (\mathrm{Rest}_b^a(0), ..., \mathrm{Rest}_b^a(k))$ is a one-one mapping from $\{a \in \mathbb{N} : a < b_0 \cdot ... \cdot b_k\}$ into $\{j : j < b_0\} \times ... \times \{j : j < b_k\}$. Since both sets have the same (finite) cardinality, this mapping is also onto (surjective). Therefore there exists an $a \in \mathbb{N}$ with $(\mathrm{Rest}_b^a(0), ..., \mathrm{Rest}_b^a(k)) = (m_0, ..., m_k)$.

Satz 6.3

The set REK is closed under primitive recursion and therefore coincides with the set of all recursive functions (cf. Lemma 4.15).

Proof:

Let $g \in \mathrm{REK}^n$, $h \in \mathrm{REK}^{n+2}$ and $f = (\mathsf{R}gh)$.

Definitions

$\beta(c,i) := \mathrm{Rest}_{\pi_2(c)}^{\pi_1(c)}(i)$,

$|a-b| := (a \dot{-} b) + (b \dot{-} a)$,

$G_0(\vec{a}, b, c) := \min\{i : (b \dot{-} i) \cdot (1 \dot{-} |h(\vec{a}, i, \beta(c,i)) - \beta(c, i+1)|) = 0\}$,

$G(\vec{a}, b, c) := |g(\vec{a}) - \beta(c,0)| + (b \dot{-} G_0(\vec{a}, b, c))$.

Then the following holds:

(1) $G \in \mathrm{REK}$

(2) $b \dot{-} G_0(\vec{a}, b, c) = 0 \;\Leftrightarrow\; \forall i < b(h(\vec{a}, i, \beta(c,i)) = \beta(c, i+1))$

(3) $\forall \vec{a} \in \mathbb{N}^n \forall b \exists c(G(\vec{a}, b, c) = 0)$

Proof of (3): Let $\vec{a}$ and b be given. By 6.2 there is a $c \in \mathbb{N}$ with $\beta(c,i) = \mathrm{Rest}_{\pi_2(c)}^{\pi_1(c)}(i) = f(\vec{a}, i)$ für $i = 0, ..., b$. By (2) we obtain $G(\vec{a}, b, c) = 0$.

(4) $f(\vec{a}, b) = \beta(\min\{i : G(\vec{a}, b, i) = 0\}, b)$

Proof: Let $c := \min\{i : G(\vec{a}, b, i) = 0\}$. Then $G(\vec{a}, b, c) = 0$ and thus (by (2))
$g(\vec{a}) = \beta(c,0) \wedge \forall i < b(h(\vec{a}, i, \beta(c,i)) = \beta(c, i+1))$, i.e., $f(\vec{a}, b) = \beta(c, b)$.

The axiom system Q

(Q1) $\forall x \neg(\mathsf{S}x = \mathsf{0}) \wedge \forall x \forall y(\mathsf{S}x = \mathsf{S}y \to x = y)$

(Q2) $\forall x(x + \mathsf{0} = x) \wedge \forall x \forall y(x + \mathsf{S}y = \mathsf{S}(x + y))$

(Q3) $\forall x(x \cdot \mathsf{0} = \mathsf{0}) \wedge \forall x \forall y(x \cdot \mathsf{S}y = (x \cdot y) + x))$

(Q4) $\forall x(x \approx \mathsf{0} \vee \exists y(x \approx \mathsf{S}y))$.

Abbreviation: $s < t := \exists z(\mathsf{S}z + s \approx t)$.

Obviously Q is consistent.

Lemma 6.4 For all $a, b, k \in \mathbb{N}$:

(a) $\mathsf{Q} \vdash \underline{a} + \underline{b} \approx \underline{a+b}$

(b) $\mathsf{Q} \vdash \underline{a} \cdot \underline{b} \approx \underline{a \cdot b}$

(c) $a < b \;\Rightarrow\; \mathsf{Q} \vdash \underline{a} < \underline{b}$,

(d) $a < b \;\Rightarrow\; \mathsf{Q} \vdash \neg(\underline{a} = \underline{b})$,

(e) $a \leq b \;\Rightarrow\; \mathsf{Q} \vdash \neg(\underline{b} < \underline{a})$,

(f) $\mathsf{Q} \vdash \mathsf{S}z + \underline{a} \approx z + \mathsf{S}\underline{a}$,

(g) $\mathsf{Q} \vdash x < \underline{k} \to\ x \approx \underline{0} \vee ... \vee x \approx \underline{k{-}1}$,

(h) $\mathsf{Q} \vdash x \approx \underline{0} \vee ... \vee x \approx \underline{k} \vee \underline{k} < x$.

Proof:

Abbreviation: $\vdash B \ :\Leftrightarrow\ \mathsf{Q} \vdash B\ ,\quad A \vdash B \ :\Leftrightarrow\ \mathsf{Q} \cup \{A\} \vdash B$.

Preliminary remark: Due to (Q4) we have $(*) \vdash A_x(\mathbf{0})\ \&\ \vdash A_x(\mathsf{S}y)\ \&\ y \notin \mathrm{FV}(A)\ \Rightarrow \vdash A$.

(a) Induction on b:

1. $b = 0$: $\vdash \underline{a} + \mathbf{0} \overset{(\mathsf{Q}2)}{\approx} \underline{a}$.

2. $b = m{+}1$: $\vdash \underline{a} + \underline{b} \approx \underline{a} + \mathsf{S}\underline{m} \overset{(\mathsf{Q}2)}{\approx} \mathsf{S}(\underline{a} + \underline{m}) \overset{\mathrm{IV}}{\approx} \mathsf{S}\underline{a{+}m} \approx \underline{a{+}b}$.

(b) is proved in the same way as (a).

(c) $a < b\ \Rightarrow\ (c{+}1) + a = b \overset{\mathrm{a)}}{\Rightarrow} \vdash \mathsf{S}\underline{c} + \underline{a} \approx \underline{b}\ \Rightarrow \vdash \exists z(\mathsf{S}z + \underline{a} \approx \underline{b})$.

(d) Induction on a. Let $b = m + 1$.

1. $a = 0$: $(\mathsf{Q}1) \Rightarrow \vdash \neg(\mathbf{0} \approx \mathsf{S}\underline{m})\ \Rightarrow \vdash \neg(\underline{a} \approx \underline{b})$.

2. $a = k + 1$: $k < m \overset{\mathrm{IV}}{\Rightarrow} \vdash \neg(\underline{k} \approx \underline{m}) \overset{(\mathsf{Q}1)}{\Rightarrow} \vdash \neg(\mathsf{S}\underline{k} \approx \mathsf{S}\underline{m})$.

(e) Induction on b: 1. $b = 0$: Then also $a = 0$. $\mathsf{S}z + \mathbf{0} \approx \mathbf{0} \vdash \mathsf{S}z \approx \mathbf{0} \vdash \bot$.

Hence $\vdash \forall z(\mathsf{S}z + \mathbf{0} \approx \mathbf{0} \to \bot)$, i.e., $\vdash \neg(\mathbf{0} < \mathbf{0})$.

2. $b = n + 1$ and $a = 0$: $\mathsf{S}z + \underline{b} \approx \mathbf{0} \vdash \mathsf{S}(\mathsf{S}z + \underline{n}) \approx \mathbf{0} \vdash \bot$.

3. $b = n + 1$ and $a = m + 1$: Then $m \leq n$ and by IH $\vdash \mathsf{S}z + \underline{n} \approx \underline{m} \to \bot$.

Hence $\mathsf{S}z + \underline{b} \approx \underline{a} \vdash \mathsf{S}(\mathsf{S}z + \underline{n}) \approx \mathsf{S}\underline{m} \vdash \mathsf{S}z + \underline{n} \approx \underline{m} \vdash \bot$.

(f) Induction on a: 1. $\vdash \mathsf{S}z + \mathbf{0} \approx \mathsf{S}z \approx \mathsf{S}(z + \mathbf{0}) \approx z + \mathsf{S}\mathbf{0}$.

2. $\vdash \mathsf{S}z + \underline{n{+}1} \approx \mathsf{S}z + \mathsf{S}\underline{n} \approx \mathsf{S}(\mathsf{S}z + \underline{n}) \overset{\mathrm{IV}}{\approx} \mathsf{S}(z + \underline{n{+}1}) \approx z + \mathsf{S}\underline{n{+}1}$.

(g) Induction on k:

1. $k = 0$: $\left.\begin{array}{l} \vdash \mathsf{S}z + \mathbf{0} \approx \mathsf{S}z \neq \mathbf{0}\ \Rightarrow \vdash \mathbf{0} < \mathbf{0} \to \bot \\ \vdash \mathsf{S}z + \mathsf{S}y \approx \mathsf{S}(\mathsf{S}z + y) \neq \mathbf{0}\ \Rightarrow \vdash \mathsf{S}y < \mathbf{0} \to \bot \end{array}\right\} \Longrightarrow \vdash x < \underline{0} \to \bot$.

2. $k > 0$:

(1) $\vdash \mathsf{S}z \approx \underline{k} \wedge z \approx \mathbf{0} \to \mathbf{0} \approx \underline{k{-}1}$,

(2) $\vdash \mathsf{S}z \approx \underline{k} \wedge z \approx \mathsf{S}y \to \mathsf{S}y + \mathbf{0} \approx \underline{k-1}$,

(3) $\vdash \mathsf{S}z + \mathbf{0} \approx \underline{k} \wedge (z \approx \mathbf{0} \vee \exists y(z \approx \mathsf{S}y)) \to \mathbf{0} < \underline{k{-}1} \vee \mathbf{0} \approx \underline{k{-}1}$, [by (1),(2)]

(4) $\vdash \mathsf{S}z + \mathbf{0} \approx \underline{k} \to \mathbf{0} < \underline{k{-}1} \vee \mathbf{0} \approx \underline{k{-}1}$ [by (3) and (Q1)]

(5) $\vdash \mathbf{0} < \underline{k} \to \mathbf{0} = \underline{0} \vee ... \vee \mathbf{0} \approx \underline{k{-}1}$, [by (4) and IH]

(6) $\mathsf{S}z + \mathsf{S}y \approx \underline{k} \vdash \mathsf{S}z + y \approx \underline{k{-}1} \overset{\mathrm{IV}}{\vdash}\ y \approx \mathbf{0} \vee ... \vee y \approx \underline{k{-}2} \vdash \mathsf{S}y \approx \underline{1} \vee ... \vee \mathsf{S}y \approx \underline{k{-}1}$,

(7) $\vdash \mathsf{S}y < \underline{k} \to \mathsf{S}y \approx \underline{0} \vee ... \vee \mathsf{S}y \approx \underline{k{-}1}$, [by (6)]

(8) $\vdash x < \underline{k} \to x \approx \underline{0} \vee ... \vee x \approx \underline{k{-}1}$, [by (7),(5),(*)].

(h) Induction on k:

1. By (Q4) and (Q2) we have $\vdash x \approx \mathbf{0} \vee \exists z(\mathsf{S}z + \mathbf{0} \approx x)$, i.e., $\vdash x \approx \underline{0} \vee \underline{0} < x$.

2. $k \to k{+}1$:

(1) $\mathsf{S}\mathbf{0} + \underline{k} \approx x \vdash x \approx \underline{k{+}1}$, [by (a)]

(2) $\mathsf{SS}y + \underline{k} \approx x \overset{\text{f)}}{\vdash} \mathsf{S}y + \underline{k{+}1} \approx x \vdash \underline{k{+}1} < x$,

(3) $\vdash \mathsf{S}z + \underline{k} \approx x \to x \approx \underline{k{+}1} \vee \underline{k{+}1} < x$, [by (2) and $(*)$]

(4) $\vdash \underline{k} < x \to x \approx \underline{k{+}1} \vee \underline{k{+}1} < x$, [by (3)]

(5) $\vdash x \approx \underline{0} \vee ... \vee x \approx \underline{k} \vee x \approx \underline{k{+}1} \vee \underline{k{+}1} < x$, [by IH and (4)].

Theorem 6.5

Every recursive function (and thus every recursive relation) is representable in Q.

Proof:

Abb.: $\vdash A \ :\Leftrightarrow\ \mathsf{Q} \vdash A$.

By induction on the definition of REK we prove that every function $f \in \text{REK}$ is representable in Q. (Note that by 6.4d we have $\vdash \neg(\underline{a} \approx \underline{b})$ for $a \neq b$, and need not to show condition (2) in the definition of "representable".)

1. The functions C^n_k, S, I^n_i are represented by the formulas $v_{n+1} \approx \mathbf{0}$, $v_2 \approx \mathsf{S}v_1$, $v_{n+1} \approx v_i$.

2. The function $\mathbf{1}_<$ is represented by the formula $A := (v_1 < v_2 \wedge \underline{1} \approx v_3) \vee (\neg(v_1 < v_2) \wedge \underline{0} \approx v_3)$.

Proof: $\mathbf{1}_<(a_1, a_2) = 1 \ \Rightarrow\ a_1 < a_2$ [6.4c]$\Rightarrow \vdash \underline{a_1} < \underline{a_2} \ \Rightarrow \vdash A(\underline{a_1}, \underline{a_2}, \underline{1}) \wedge (A(\underline{a_1}, \underline{a_2}, y) \to \underline{1} \approx y)$.

$\mathbf{1}_<(a_1, a_2) = 0 \ \Rightarrow\ a_2 \leq a_1$ [6.4e]$\Rightarrow \vdash \neg(\underline{a_1} < \underline{a_2}) \ \Rightarrow \vdash A(\underline{a_1}, \underline{a_2}, \underline{0}) \wedge (A(\underline{a_1}, \underline{a_2}, y) \to \underline{0} \approx y)$.

3. Let $f = (\circ h g_1 ... g_m)$ with $g_1, ..., g_m \in \text{REK}^n$ and $h \in \text{REK}^m$. By I.H. there are $(n{+}1)$-ary formulas $B_1, ..., B_m$ and an $(m+1)$-ary formula C such that:

(i) $h(b_1, ..., b_m) = b \ \Rightarrow \vdash C(\underline{b_1}, ..., \underline{b_m}, \underline{b}) \wedge \forall y(C(\underline{b_1}, ..., \underline{b_m}, y) \to \underline{b} \approx y)$,

(ii) $g_i(\vec{a}) = b_i \ \Rightarrow \vdash B_i(\underline{\vec{a}}, \underline{b_i}) \wedge \forall y(B_i(\underline{\vec{a}}, y) \to \underline{b_i} \approx y)$ $(i = 1, ..., m)$.

Def.: $A := \exists y_1 ... \exists y_m [\, C(y_1, ..., y_m, v_{n+1}) \wedge B_1(v_1, ..., v_n, y_1) \wedge ... \wedge B_m(v_1, ..., v_n, y_m) \,]$.

Now let $f(\vec{a}) = b$. Then $h(b_1, ..., b_m) = b$ with $b_i := g_i(\vec{a})$ $(i = 1, ..., m)$.

By (i),(ii) we obtain $\vdash C(\underline{b_1}, ..., \underline{b_m}, \underline{b}) \wedge B_1(\underline{\vec{a}}, \underline{b_1}) \wedge ... \wedge B_m(\underline{\vec{a}}, \underline{b_m})$, and then $\vdash A(\underline{\vec{a}}, \underline{b})$.

By (i),(ii) we also have:

$\vdash (B_1(\underline{\vec{a}}, y_1) \to \underline{b_1} \approx y_1) \wedge ... \wedge (B_m(\underline{\vec{a}}, y_m) \to \underline{b_m} \approx y_m) \wedge (C(\underline{b_1}, ..., \underline{b_m}, y) \to \underline{b} \approx y)$.

From this together with $\vdash \underline{b_1} \approx y_1 \wedge ... \wedge \underline{b_m} \approx y_m \wedge C(y_1, ..., y_m, y) \to C(\underline{b_1}, ..., \underline{b_m}, y)$, we obtain

$\vdash B_1(\underline{\vec{a}}, y_1) \wedge ... \wedge B_m(\underline{\vec{a}}, y_m) \wedge C(y_1, ..., y_m, y) \to \underline{b} \approx y$, and then

$\vdash \exists y_1 ... \exists y_m [\, B_1(\underline{\vec{a}}, y_1) \wedge ... \wedge B_m(\underline{\vec{a}}, y_m) \wedge C(y_1, ..., y_m, y) \,] \to \underline{b} \approx y$.

4. Let $f = (\mu g)$ with $g \in \text{REK}^{n+1}$ and $\forall \vec{a} \exists i [g(\vec{a}, i) = 0]$. By I.H. there is an $(n{+}2)$-ary formula B such that: $g(\vec{a}, i) = c \ \Rightarrow \vdash B(\underline{\vec{a}}, \underline{i}, \underline{c}) \wedge \forall y(B(\underline{\vec{a}}, \underline{i}, y) \to \underline{c} \approx y)$.

Definition: $A := B(v_1, ..., v_n, u, \mathbf{0}) \wedge \forall x < u \exists y(\neg(y \approx \mathbf{0}) \wedge B(v_1, ..., v_n, x, y))$, where $u := v_{n+1}$.

Assume now $f(\vec{a}) = k$. Then $g(\vec{a}, k) = 0$ und $c_i := g(\vec{a}, i) > 0$ for all $i < k$.

We obtain:

(1) $\vdash B(\underline{\vec{a}}, \underline{k}, y) \to y \approx \mathbf{0}$

(2) $\vdash B(\underline{\vec{a}}, \underline{i}, y) \to \neg(y \approx \mathbf{0})$, for $i = 0, ..., k-1$ [(Q1)]

(3) $\vdash \neg B(\underline{\vec{a}}, \underline{i}, \mathbf{0})$, for $i = 0, ..., k-1$ [(2)]

(4) $\vdash B(\underline{\vec{a}}, u, \mathbf{0}) \to \neg(u \approx \underline{i})$, für $i = 0, ..., k-1$ [(3)]

(5) $\vdash \neg\exists y(\neg(\mathbf{0} \approx y) \wedge B(\vec{\underline{a}}, \underline{k}, y))$ [(1)]

(6) $\vdash \forall x < u \exists y(\neg(y \approx \mathbf{0}) \wedge B(\vec{\underline{a}}, x, y)) \rightarrow \neg(\underline{k} < u)$ [(5)]

(7) $\vdash A(\vec{\underline{a}}, u) \rightarrow \neg(u \approx \mathbf{0}) \wedge ... \wedge \neg(u \approx \underline{k-1}) \wedge \neg(\underline{k} < u)$ [(4),(6)]

(8) $\vdash A(\vec{\underline{a}}, u) \rightarrow \underline{k} \approx u$ [(7), 6.4h]

Further:

(9) $\vdash B(\vec{\underline{a}}, \underline{k}, \mathbf{0})$

(10) $\vdash \neg(\underline{c_i} \approx \mathbf{0}) \wedge B(\vec{\underline{a}}, \underline{i}, \underline{c_i})$, for $i = 0, ..., k-1$

(11) $\vdash \underline{i} \approx x \rightarrow \exists y(\neg(y \approx \mathbf{0}) \wedge B(\vec{\underline{a}}, x, y))$, for $i = 0, ..., k-1$

(12) $\vdash \mathbf{\underline{0}} \approx x \vee ... \vee \underline{k-1} \approx x \rightarrow \exists y(\neg(y \approx \mathbf{0}) \wedge B(\vec{\underline{a}}, x, y))$,

(13) $\forall x < \underline{k} \exists y(\neg(y \approx \mathbf{0}) \wedge B(\vec{\underline{a}}, x, y))$ [(12), 6.4g]

Definition

Let $\mathcal{L}_{ar} := \mathrm{L}(\mathsf{Q}) = \{\mathbf{0}, \mathsf{S}, +, \cdot\}$. The formulas of the language $\mathcal{L}_{ar}$ are called *arithmetical formulas.*

Let $\mathcal{N}$ be the standard model of Q.

A relation $R \subseteq \mathbb{N}^n$ is called *arithmetical*, if it is definable in $\mathcal{N}$.

Theorem 6.6

Every recursively enumerable relation is arithmetical.

Proof:

Let $Q \subseteq \mathbb{N}$ be recursively enumerable. Then there is a (primitive) recursive relation $R \subseteq \mathbb{N}^2$ with $Q = \{a \in \mathbb{N} : \exists b\,(a, b) \in R\}$. By 6.5 there is an arithmetical formula A with $[\,(a, b) \in R \Rightarrow \mathsf{Q} \vdash A(\underline{a}, \underline{b})\,]$ and $[\,(a, b) \notin R \Rightarrow \mathsf{Q} \vdash \neg A(\underline{a}, \underline{b})\,]$. Since $\mathcal{N} \models \mathsf{Q}$, from this we get $R = \{(a, b) : \mathcal{N} \models A(\underline{a}, \underline{b})\}$ and then $Q = \{a : \mathcal{N} \models \exists y A(\underline{a}, y)\}$.

Theorem 6.7 (Tarski)

$\ulcorner \mathrm{Th}(\mathcal{N}) \urcorner$ is not arithmetical.

Corollary (Gödel)

$\mathrm{Th}(\mathcal{N})$ ist not recursively enumerable and consequently not recursively axiomatizable.

Proof: 5.7 and 6.6.

Theorem 6.8 Every consistent theory T with $\mathsf{Q} \subseteq T$ is undecidable.

Proof: 5.9 and 6.6.

Remark (Gödel's 1st Incompleteness Theorem)

In Theorem 5.10 the premise "all recursive functions are representable in T (Φ, resp.)" can be replaced by "$\mathsf{Q} \subseteq T$ ($\Phi \vdash \mathsf{Q}$, resp.)". From 5.10(A) we obtain

Every consistent, recursively enumerable theory T with $\mathsf{Q} \subseteq T$ is incomplete.

Theorem 6.10 (Undecidability of 1st-order predicate logic)

The set of all logically valid $\mathcal{L}_{ar}$-sentences is undecidable (not recursive).

Proof:

Let P be the set of all logically valid $\mathcal{L}_{ar}$-sentences and $\mathrm{Cn}(\mathsf{Q})$ the set of all $\mathcal{L}_{ar}$-sentences which are logically consequences of Q. By 6.8 $\mathrm{Cn}(\mathsf{Q})$ is not recursive. On the other side we have: $a \in \ulcorner \mathrm{Cn}(\mathsf{Q}) \urcorner \Leftrightarrow \langle SN(\rightarrow), \ulcorner \bigwedge \mathsf{Q} \urcorner, a \rangle \in \ulcorner P \urcorner$, where $\bigwedge \mathsf{Q} := (\mathsf{Q}1) \wedge \ldots \wedge (\mathsf{Q}4)$. Hence P cannot be recursive either.

Theorem 6.11 (Rossers refinement of Gödel's 1st incompleteness theorem)

Let Φ be a consistent, recursive axiom system such that $\Phi \vdash \mathsf{Q}$. Then one can define a true arithmetic sentence G such that $\Phi \nvdash G$ and $\Phi \nvdash \neg G$.

Proof:

By 5.3 Prf_Φ is recursive, and by 6.6 we have a 2-ary arithmetical formula B satisfying:

(1) $(a,b) \in \mathsf{Prf}_\Phi \;\Rightarrow\; \mathsf{Q} \vdash B(\underline{a}, \underline{b})$,

(2) $(a,b) \notin \mathsf{Prf}_\Phi \;\Rightarrow\; \mathsf{Q} \vdash \neg B(\underline{a}, \underline{b})$.

Def.: $(a,b) \in \mathsf{Rft}_\Phi \;:\Leftrightarrow\; (\langle SN(\rightarrow), a, \ulcorner \bot \urcorner \rangle, b) \in \mathsf{Prf}_\Phi$.

With Prf_Φ also Rft_Φ is recursive. Therefore we have a 2-ary arithmetical formula B^* such that:

(1') $(a,b) \in \mathsf{Rft}_\Phi \;\Rightarrow\; \mathsf{Q} \vdash B^*(\underline{a}, \underline{b})$,

(2') $(a,b) \notin \mathsf{Rft}_\Phi \;\Rightarrow\; \mathsf{Q} \vdash \neg B^*(\underline{a}, \underline{b})$.

By the Fixpointlemma we obtain an arithmetical sentence G such that

(3') $\mathsf{Q} \vdash G \leftrightarrow \neg \exists y (B(\underline{\ulcorner G \urcorner}, y) \wedge \forall z{<}y \neg B^*(\underline{\ulcorner G \urcorner}, z))$.

Assumption: $\Phi \vdash G$.

Then there exists a $b \in \mathbb{N}$ with $(\ulcorner G \urcorner, b) \in \mathsf{Prf}_\Phi$. Since Φ is consistent, $(\ulcorner G \urcorner, n) \notin \mathsf{Rft}_\Phi$ holds for all $n \in \mathbb{N}$.

By (1),(2') from this we obtain $\mathsf{Q} \vdash B(\underline{\ulcorner G \urcorner}, \underline{b})$ and $\mathsf{Q} \vdash \neg B^*(\underline{\ulcorner G \urcorner}, \underline{n})$ for all n.

Further we get:

$\mathsf{Q} \vdash B(\underline{\ulcorner G \urcorner}, \underline{b}) \wedge \forall z{<}\underline{b} \neg B^*(\underline{\ulcorner G \urcorner}, z)$,

$\mathsf{Q} \vdash \exists y (B(\underline{\ulcorner G \urcorner}, y) \wedge \forall z{<}y \neg B^*(\underline{\ulcorner G \urcorner}, z))$,

$\mathsf{Q} \vdash \neg G$,

$\Phi \vdash \neg G$. Contradiction.

Hence $\Phi \nvdash G$. From this we get $\mathcal{N} \models \neg \exists y (B(\underline{\ulcorner G \urcorner}, y) \wedge \forall z < y \neg B^*(\underline{\ulcorner G \urcorner}, z))$ and then by (3') $\mathcal{N} \models G$.

Assumption: $\Phi \vdash \neg G$.

Then there exists a $b \in \mathbb{N}$ with $(\ulcorner G \urcorner, b) \in \mathsf{Rft}_\Phi$. Since Φ is consistent, we have $(\ulcorner G \urcorner, n) \notin \mathsf{Prf}_\Phi$ for all $n \in \mathbb{N}$.

By (2),(1') from this we get $\mathsf{Q} \vdash B^*(\underline{\ulcorner G \urcorner}, \underline{b})$ and $\mathsf{Q} \vdash \neg B(\underline{\ulcorner G \urcorner}, \underline{n})$ for all n.

Further we obtain:

$\mathsf{Q} \vdash [\underline{b} < y \rightarrow \exists z < y B^*(\underline{\ulcorner G \urcorner}, z)] \wedge [y \leq \underline{b} \rightarrow \neg B(\underline{\ulcorner G \urcorner}, y)]$,

$\mathsf{Q} \vdash \forall y (\neg B(\underline{\ulcorner G \urcorner}, y) \vee \exists z < y B^*(\underline{\ulcorner G \urcorner}, z))$,

$\mathsf{Q} \vdash \neg \exists y (B(\underline{\ulcorner G \urcorner}, y) \wedge \forall z < y \neg B^*(\underline{\ulcorner G \urcorner}, z))$,

$\mathsf{Q} \vdash G$. Contradiction.

7 The axiom system ZF of Zermelo-Fraenkel set theory

The language of ZF consists of a single binary relation symbol ϵ. The *axioms of* ZF are

(Extensionality) $\forall x_0 \forall x_1 (\forall y (y \,\epsilon\, x_0 \leftrightarrow y \,\epsilon\, x_1) \rightarrow x_0 = x_1)$

(Pair) $\forall x_0 \forall x_1 \exists y \forall z (z \,\epsilon\, y \leftrightarrow z = x_0 \vee z = x_1)$

(Union) $\forall x \exists y \forall z (z \,\epsilon\, y \leftrightarrow \exists v (v \,\epsilon\, x \wedge z \,\epsilon\, v))$

(Powerset) $\forall x \exists y \forall z (z \,\epsilon\, y \leftrightarrow \forall v (v \,\epsilon\, z \rightarrow v \,\epsilon\, x))$

(Infinity) $\exists w (\exists x [x \,\epsilon\, w \wedge \neg \exists y (y \,\epsilon\, x)] \wedge \forall x [x \,\epsilon\, w \rightarrow \exists y (y \,\epsilon\, w \wedge \forall z (z \,\epsilon\, y \leftrightarrow z \,\epsilon\, x \vee z = x))])$

(Foundation) $\forall x (\exists y (y \,\epsilon\, x) \rightarrow \exists y [y \,\epsilon\, x \wedge \neg \exists z (z \,\epsilon\, x \wedge z \,\epsilon\, y)])$

(Replacement) $\forall z_1 ... \forall z_n (\forall x \forall y_0 \forall y_1 [\varphi(x, y_0, \vec{z}) \wedge \varphi(x, y_1, \vec{z}) \rightarrow y_0 = y_1] \rightarrow$
$\rightarrow \forall u \exists w \forall y [y \,\epsilon\, w \leftrightarrow \exists x (x \,\epsilon\, u \wedge \varphi(x, y, \vec{z}))])$, for each $(n{+}2)$-ary $\{\epsilon\}$-formula φ $(n \geq 0)$.

We are now going do develop a considerable amount of set theory on the basis of ZF, in other words we will derive logical consequences from ZF. But we will do this semantically by working in an arbitrary ZF-model.

Assumption. Let (V, ϵ^V) be a model of ZF.

In the sequel we use $a, b, c, d, e, f, g, h, u, w, x, y, z$ as syntactic variables for objects from V.

Definition

A collection X of objects from V is called a *class*, if there are an $(n{+}1)$-ary $\{\epsilon\}$-formula φ and objects $a_1, ..., a_n \in V$ such that $X = \{x : (V, \epsilon^V) \models \varphi[a_1, ..., a_n, x]\}$. (Since x ranges over objects from V, "$\{x : ....\}$" has the same meaning as "$\{x \in V : ....\}$".)

We use $A, B, C, D, E, F, G, H, R$ as syntactic variables for classes.

For each $a \in V$ we define the class $E(a) := \{x : x \,\epsilon^V a\}$.
Since $(V, \epsilon^V) \models$ (Extensionality), we have: $E(a) = E(b) \Rightarrow a = b$.
Therefore in the sequel every $a \in V$ will be identified with its "extension" $E(a)$.
Then for $a, b \in V$ we have: $b \in a \Leftrightarrow b \,\epsilon^V a$. Hence we can write $(V, \in)$ for (V, ϵ^V).

CONVENTION:
From now on, by a *set* we always mean an object from V.

Remark:
1. Since a is identified with $E(a)$, every set is a class.
2. Not every class is a set.

Proof of 2.: $R := \{x : x \notin x\}$ is a class, and $\forall a \in V (a \in R \leftrightarrow a \notin a)$ holds.
Now if R would be a set, then we would have $R \in R \leftrightarrow R \notin R$.

Remark
$\emptyset := \{x : x \neq x\}$ is a class. If A, B are classes then $A \cup B$, $A \cap B$, $A \setminus B$,
$\bigcup A := \{x : \exists y \in A (x \in y)\}$, $\bigcap A := \{x : \forall y \in A (x \in y)\}$, $\mathcal{P}(A) := \{x : x \subseteq A\}$ are classes.

From the assumption $(V, \in) \models$ (Pair) $\wedge$ (Union) $\wedge$ (Powerset) we get:

Lemma 7.1

(a) $\forall a \forall b(\{a, b\} \in V)$,

(b) $\forall a(\bigcup a \in V)$,

(c) $\forall a(\mathcal{P}(a) \in V)$,

(d) $\forall a \forall b(a \cup b = \bigcup\{a, b\} \in V)$,

(e) $\forall a_1 \ldots \forall a_n(\{a_1, ..., a_n\} \in V)$.

Definition. $(a, b) := \{\{a\}, \{a, b\}\}$ (ordered pair of a and b)

Lemma 7.2

(a) $(a, b) \in V$,

(b) $(a_1, b_1) = (a_2, b_2) \implies a_1 = a_2 \wedge b_1 = b_2$.

Proof:

(a) follows from 7.1.

(b) Let $c := (a_1, b_1) = (a_2, b_2)$.

Case 1: $a_1 = b_1$. $\{\{a_1\}\} = c = \{\{a_2\}, \{a_2, b_2\}\} \Rightarrow \{a_1\} = \{a_2\} = \{a_2, b_2\} \Rightarrow a_1 = a_2 = b_2$.

Case 2: $a_2 = b_2$. Analoguous to case 1.

Case 3: $a_1 \neq b_1$ & $a_2 \neq b_2$. $\{a_1\}, \{a_1, b_1\} \in c = \{\{a_2\}, \{a_2, b_2\}\}$ & $\{a_1\} \neq \{a_2, b_2\}$ & $\{a_1, b_1\} \neq \{a_2\} \implies$ $\{a_1\} = \{a_2\}$ & $\{a_1, b_1\} = \{a_2, b_2\} \Rightarrow a_1 = a_2$ & $b_1 = b_2$.

Definition

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

A *relation* is a class of ordered pairs, i.e., a subclass of $V \times V$.

A *function* is a relation F such that $\forall x, y, z((x, y) \in F \wedge (x, z) \in F \to y = z)$ holds.

Abbreviations

Let R be a class.

$\mathrm{Rel}(R) :\Leftrightarrow R$ is a relation (i.e. $R \subseteq V \times V$)

$\mathrm{Fun}(R) :\Leftrightarrow R$ is a function

$Rxy :\Leftrightarrow xRy :\Leftrightarrow (x, y) \in R$

$\mathrm{dom}(R) := \{x : \exists y\, Rxy\}$ *(domain of R)*

$\mathrm{ran}(R) := \{y : \exists x\, Rxy\}$ *(range of R)*

$R|A := \{(x, y) \in R : x \in A\}$ *(restriction of R to A)*

$R[A] := \{y : \exists x \in A\, Rxy\}$ *(image of A under R)*

$Q \circ R := \{(x, y) : \exists z(Rxz \wedge Qzy)\}$ *(composition of Q and R)*

$R^{-1} := \{(y, x) : Rxy\}$ *(inverse of R)*

If F is a function and $x \in \mathrm{dom}(F)$, then $F(x)$ denotes the uniquely determined y such that $(x, y) \in F$, (i.e. the *value of F at x*).

$F : A \to B :\Leftrightarrow \mathrm{Fun}(F) \wedge \mathrm{dom}(F) = A \wedge \mathrm{ran}(F) \subseteq B$ *(F is a function from A to B)*

$b^a := \{f \in V : f : a \to b\}$

A function F is called *injective*, if $\forall x_0, x_1 \in \mathrm{dom}(F)(F(x_0) = F(x_1) \to x_0 = x_1)$.

Corollary

1. If A, B, R, Q are classes then $A \times B, \mathrm{dom}(R), \mathrm{ran}(R), R|_A, R[A], Q \circ R, R^{-1}$ are classes.

2. If F, G are functions, and $A \subseteq \mathrm{dom}(F) \cap \mathrm{dom}(G)$, then:

(a) $\mathrm{Fun}(F|A)$ & $\mathrm{dom}(F|A) = A$ & $\mathrm{ran}(F|A) = F[A]$ & $\forall x \in A(F(x) = (F|A)(x))$,

(b) $\mathrm{ran}(F) = F[\mathrm{dom}(F)]$ & $F = F|\mathrm{dom}(F)$

(c) $\forall x \in A(F(x) = G(x)) \;\Rightarrow\; F|A = G|A$

(d) $\mathrm{dom}(F) = \mathrm{dom}(G) = A$ & $\forall x \in A(F(x) = G(x)) \;\Rightarrow\; F = G$

Lemma 7.3 (Consequences from $(V, \in) \models$ (Replacement))

(a) If F is a function and $a \in V$, then $F[a] \in V$.

(b) If C is a class and $a \in V$, then $\{x \in a : x \in C\} \in V$.

(c) $C \subseteq A$ & $A \in V \;\Rightarrow\; C \in V$.

(d) $A \neq \emptyset \;\Rightarrow\; \bigcap A \in V$.

Proof:

(a) Obviously $(x_0, x_1) = y \;\Leftrightarrow\; \exists u, v \in y(\forall z(z \in y \leftrightarrow z = u \vee z = v) \wedge \forall z(z \in u \leftrightarrow z = x_0) \forall z(z = x_0 \vee z = x_1))$, and $(x, y) \in F \leftrightarrow \exists z(z \in F \wedge z = (x, y))$. Therefore, since F is a class, there are an $(n{+}2)$-ary formula φ and sets $c_1, ..., c_n$ such that $\forall x, y(\,(x, y) \in F \;\Leftrightarrow\; (V, \in) \models \varphi[x, y, c_1, ..., c_n]\,)$. Since F is a function, we also have $(V, \in) \models \forall x \forall y_0 \forall y_1(\varphi(x, y_0, \vec{z}) \wedge \varphi(x, y_1, \vec{z}) \to y_0 = y_1)[\vec{z}/\vec{c}]$. Now by (Replacement) we obtain $(V, \in) \models \exists w \forall y(y \,\epsilon\, w \leftrightarrow \exists x(x \,\epsilon\, u \wedge \varphi(x, y, \vec{z})))[u/a, \vec{z}/\vec{c}]$. Hence $\{y : \exists x(x \in a \wedge (x, y) \in F)\} \in V$, i.e. $F[a] \in V$.

(b) Obviously, $F := \{(x, x) : x \in C\}$ is a function. Therefore (by 7.3a) $F[a] \in V$. Further we have
$y \in F[a] \;\Leftrightarrow\; \exists x \in a((x, y) \in F) \;\Leftrightarrow\; \exists x \in a \exists x_0(x_0 \in C \wedge (x, y) = (x_0, x_0)) \;\Leftrightarrow\; y \in a \wedge y \in C$.

(c) $C \subseteq A \;\Rightarrow\; C = \{x \in A : x \in C\}$. (d) Let $a \in A$. Then $\bigcap A \subseteq a \in V$.

Lemma 7.4

(a) $a, b, r \in V \;\Rightarrow\; a \cup b, a \cup b, a \setminus b, a \times b, b^a, \mathrm{dom}(r), \mathrm{ran}(r), r[a], r|a, r^{-1}, a \circ r \in V$.

(b) $\mathrm{Rel}(R)$ & $\mathrm{dom}(R) \in V$ & $\mathrm{ran}(R) \in V \;\Rightarrow\; R \in V$.

(c) $\mathrm{Fun}(F) \;\Rightarrow\; F|a \in V$.

(d) $\mathrm{Fun}(F)$ & $\mathrm{dom}(F) \in V \;\Rightarrow\; F \in V$.

Beweis:

a) 1. $a \cup b = \bigcup\{a, b\}$, $a \cap b \subseteq a$, $a \setminus b \subseteq a$.

2. $x \in a$ & $y \in b \;\Rightarrow\; \{x\}, \{x, y\} \in \mathcal{P}(a \cup b) \;\Rightarrow\; (x, y) = \{\{x\}, \{x, y\}\} \in \mathcal{P}(\mathcal{P}(a \cup b))$.
Hence $a \times b \subseteq \mathcal{P}(\mathcal{P}(a \cup b))$.

3. $b^a \subseteq \mathcal{P}(a \times b)$.

4. $(x, y) \in r \;\Rightarrow\; \{x, y\} \in \bigcup r \;\Rightarrow\; x, y \in \bigcup\bigcup r$. Hence $\mathrm{dom}(r), \mathrm{ran}(r), r[a] \subseteq \bigcup\bigcup r$.

5. $r|a \subseteq r$, $r^{-1} \subseteq \mathrm{ran}(r) \times \mathrm{dom}(r)$, $a \circ r \subseteq \mathrm{dom}(r) \times \mathrm{ran}(a)$.

(b),(c),(d) $\mathrm{Rel}(R) \;\Rightarrow\; R \subseteq \mathrm{dom}(R) \times \mathrm{ran}(R)$. $F|a \subseteq a \times F[a]$. $F = F|\mathrm{dom}(F)$.

8 Ordinal numbers

Definition

1. A relation R is *transitive* iff $\forall x, y, z(Rxy \wedge Ryz \to Rxz)$.

2. $x_R := \{y : Ryx\} \qquad (x \in V,\ R \text{ relation})$

3. $\forall yRx(....) := \forall y(yRx \to ....)$, $\exists yRx(....) := \exists y(yRx \wedge ....)$

Lemma 8.1

For each relation R the following are equivalent:

(I) *Existence of a minimal element*:

$C \neq \emptyset \to \exists x \in C \forall yRx(y \notin C)$, for each class C.

(II) *Induction over R*:

$\forall x(\forall yRx(y \in C) \to x \in C) \to \forall x(x \in C)$, for each class C.

Definition

Let R be a relation.

R *wellfounded* $:\Leftrightarrow \forall u(u \neq \emptyset \to \exists x \in u \forall yRx(y \notin u)) \ \& \ \forall x(x_R \in V)$.

Remark

Every wellfounded relation ist irreflexive. [$Rzz \Rightarrow \{z\} \neq \emptyset \wedge \neg\exists x \in \{z\} \forall yRx(y \notin \{z\})$]

Lemma 8.2

If R is a wellfounded relation then (I) and (II) hold for R.

Proof of (I): Let us assume that R is transitive. (The proof for arbitrary R will be given later.)

Assume $u \in C$. Case 1: $\forall yRu(y \notin C)$. Then we are done.

Case 2: $\exists yRu(y \in C)$. Dann $a := u_R \cap C \neq \emptyset$. Since R is wellfounded, there is a $c \in a$ such that $\forall yRc(y \notin a)$. From cRu and the transitivity of R we obtain $\forall y(yRc \to yRu)$. Hence $c \in C$ and $\forall yRc(y \notin C)$.

Theorem 8.3 (Recursion along wellfounded relations)

Let R be wellfounded and $G : A \times V \to V$. Then there exists a unique function $F : A \to V$ such that $F(x) = G(x, F|x_R)$ for all $x \in A$.

Proof: Let us assume that R is transitive. (The proof for arbitrary R will be given later.)

Let $C := \{f : \text{Fun}(f) \wedge \text{dom}(f) \subseteq A \wedge \forall x \in \text{dom}(f)(A \cap x_R \subseteq \text{dom}(f) \wedge f(x) = G(x, f|x_R))\}$ and $F := \bigcup C$.

(1) $f_1, f_2 \in C \ \& \ x \in \text{dom}(f_1) \cap \text{dom}(f_2) \Rightarrow f_1(x) = f_2(x)$.

Proof by induction over R: Let $x \in \text{dom}(f_1) \cap \text{dom}(f_2)$. Then $A \cap x_R \subseteq \text{dom}(f_1) \cap \text{dom}(f_2)$ and by I.H. $\forall y \in A \cap x_R(f_1(y) = f_2(y))$, i.e. $f_1|x_R = f_2|x_R$. Since $f_1, f_2 \in C$, we obtain $f_1(x) = f_2(x)$.

(2) $\text{Fun}(F) \ \& \ \text{dom}(F) = \bigcup\{\text{dom}(f) : f \in C\} \ \& \ \forall f \in C(f = F|\text{dom}(f))$.

Proof by (1). (obvious)

(3) $\forall x \in \text{dom}(F)(A \cap x_R \subseteq \text{dom}(F) \ \& \ F(x) = G(x, F|x_R))$

Proof: Let $x \in \mathrm{dom}(F)$. Then $x \in \mathrm{dom}(f) \subseteq \mathrm{dom}(F)$ for some $f \in C$. Hence $A \cap x_R \subseteq \mathrm{dom}(F)$.
By (2) we have $F(x) = f(x) = G(x, f|x_R) = G(x, F|x_R)$.

(4) $A = \mathrm{dom}(F)$

Proof: The inclusion $\mathrm{dom}(F) \subseteq A$ is trivial. $A \subseteq \mathrm{dom}(F)$ is proved by induction over R.
Let $x \in A$, and $a := A \cap x_R \subseteq \mathrm{dom}(F)$ (I.V.). To be shown: $x \in \mathrm{dom}(F)$.
Since R is wellfounded, we have $\neg Rxx$, thence $x \notin a$. Let $f := F|a \cup \{(x, G(x, F|x_R))\}$.
We prove $f \in C$, from which we get $x \in \mathrm{dom}(F)$.
Since $x \notin a$, we have $\mathrm{Fun}(f)$. Obviously $\mathrm{dom}(f) \subseteq A$. Since $a \subseteq \mathrm{dom}(F)$, we also have $a \subseteq \mathrm{dom}(f)$.
Assume now $y \in \mathrm{dom}(f)$:
1. $y \in a$: Then $A \cap y_R \subseteq a$ (since R transitive) and $f(y) = F(y) = G(y, F|y_R) = G(y, f|y_R)$.
2. $y = x$: In this case $A \cap x_R \subseteq a$ and $f(x) = G(x, F|x_R) = G(x, f|x_R)$.

Uniqueness: Assume $F_i : A \to V$ with $F_i(x) = G(x, F_i|x_R)$ for all $x \in A$ $(i = 0, 1)$.
Then by induction over R one obtains $\forall x \in A(F_0(x) = F_1(x))$.

Definition

1. $R \subseteq A \times A$ is called a *linear ordering on* A, if the following holds:
(i) $\forall x \neg Rxx$, (ii) $\forall x, y, z(Rzy \wedge Ryx \to Rzx)$, (iii) $\forall x, y \in A(Ryx \vee x = y \vee Rxy)$.

2. A linear ordering R on A is called a *wellordering on* A, if R is wellfounded.

3. A is called *wellordered (linearly ordered) by* R, if $R \cap (A \times A)$ is a wellordering (linear ordering) on A.

4. If $A \in V$ and R is a wellordering on A, then (A, R) is called a *wellordered set* oder briefly a *wellordering*.

Remark: If A is wellordered by R, then:
(a) Every nonempty class $C \subseteq A$ has a (unique) *least element* $\min_R(C)$.
(b) Every subclass $B \subseteq A$ is also wellordered by R.

Proof of (a): Let $Q := R \cap (A \times A)$. By Lemma 8.2 there exists an $a \in C$ such that $C \cap \{x : Qxa\} = \emptyset$.
Hence $\forall y \in C(\neg Rya)$ and further $\forall y \in C(y = a \vee Ray)$, i.e. $a = \min_R(C)$.

Definition

1. A class A is called *transitive* iff $\forall x \in A(x \subseteq A)$.
2. An *ordinal (number)* is a transitive set, which is wellordered by the relation $\in$.
3. $On := \{x : x \text{ is an ordinal number}\}$

In the following the letters $\alpha, \beta, \gamma, \delta, \xi, \eta, \zeta$ always denote ordinals.
We also write $\alpha < \beta$ for $\alpha \in \beta$, and $\alpha \leq \beta$ for $\alpha \in \beta \vee \alpha = \beta$.

Lemma 8.4

(a) $\alpha \in On \Rightarrow \alpha \subseteq On$ (d.h. On is transitive).
(b) $\alpha \leq \beta \Leftrightarrow \alpha \subseteq \beta$.

Proof:

(a) Let $x \in \alpha$. Then $x \subseteq \alpha$ and thus x is wellordered by $\in$. It remains to prove that x is transitive. Let $y \in x \;\&\; z \in y$. Since $x \in \alpha$ und α transitive, we obtain $x, y, z \in \alpha$. Hence $z \in x$, since α is linearly ordered by $\in$.

(b) "$\Rightarrow$" From $\alpha \in \beta \in On$ it follows that $\alpha \subseteq \beta$.

"$\Leftarrow$" Let $\alpha \subseteq \beta$ and $\alpha \neq \beta$. Then $\emptyset \neq \beta \setminus \alpha \subseteq \beta$. Consequently there exists a $c \in \beta \setminus \alpha$ with $c \cap (\beta \setminus \alpha) = \emptyset$ (i.e. c is $\in$-minimal). We now prove $\alpha = c$ and thus $\alpha \in \beta$.

1. Let $x \in \alpha$. Since $c \notin \alpha$, we cannot have $c = x \vee c \in x$. Therefore from $c, x \in \beta$ it follows that $x \in c$.

2. Let $x \in c$. Then $x \in \beta$, and together with $c \cap (\beta \setminus \alpha) = \emptyset$ we obtain $x \in \alpha$.

Theorem 8.5

(a) On is wellordered by $\in$.

(b) $On \notin V$.

Proof:

(a) 1. Let $\alpha \in On$. If $\alpha \in \alpha$, then $\exists x \in \alpha (x \in x)$ and α would not be wellordered by $\in$. Hence $\alpha \notin \alpha$.

2. Since every $\gamma \in On$ is transitive, we have: $\alpha, \beta, \gamma \in On \;\&\; \alpha \in \beta \;\&\; \beta \in \gamma \;\Rightarrow\; \alpha \in \gamma$.

3. Let $\alpha, \beta \in On$. Obviously $\gamma := \alpha \cap \beta \in On$. By 1. we have $\gamma \notin \gamma$, and therefore $\gamma \notin \alpha \vee \gamma \notin \beta$. By 8.4 from $\gamma \subseteq \alpha \wedge \gamma \subseteq \beta$ we obtain $(\gamma \in \alpha \vee \gamma = \alpha) \wedge (\gamma \in \beta \vee \gamma = \beta)$. Hence $\gamma = \alpha \vee \gamma = \beta$ and therefore $\alpha \subseteq \beta \vee \beta \subseteq \alpha$. By 8.4 this means $\alpha \leq \beta \vee \beta \leq \alpha$.

4. Let $\emptyset \neq u \subseteq On$, and take some $\alpha \in u$. To prove: $\exists \gamma \in u (u \cap \gamma = \emptyset)$. If $u \cap \alpha = \emptyset$, we are done. Otherwise there exists a $\gamma \in u \cap \alpha$ with $u \cap \alpha \cap \gamma = \emptyset$. From $\gamma \in \alpha$ we get $\gamma \subseteq \alpha$ and then $u \cap \gamma = u \cap \alpha \cap \gamma = \emptyset$.

(b) By 8.4a and 8.5a On is transitive and wellordered by $\in$. If $On \in V$, then On would be an ordinal, i.e. $On \in On$ contradicting (a).

Corollary

(a) Every nonempty class C of ordinals contains a least element $\min(C)$.

(b) $\forall \alpha (\forall \xi \in \alpha (\xi \in C) \;\rightarrow \alpha \in C) \;\rightarrow\; \forall \alpha (\alpha \in C)$, for each class C.

(c) For each $G : On \times V \rightarrow V$ there exists a unique $F : On \rightarrow V$ with $F(\alpha) = G(\alpha, F|\alpha)$, for all $\alpha \in On$.

Proof of (b):

(b) Let $C_1 := \{x : \; x \in On \rightarrow x \in C\}$ and $R := \{\langle \xi, \eta \rangle \in On \times On : \eta \in \xi\}$. By 8.5 and 8.2 we then have $\forall x (\forall y R x (y \in C_1) \rightarrow x \in C_1) \rightarrow \forall x (x \in C_1)$, i.e. $\forall x \in On (\forall y \in x (y \in C) \rightarrow x \in C) \rightarrow \forall x \in On (x \in C)$.

Lemma 8.6

If $A \subseteq On$ is transitive, then $(A \in V \;\Rightarrow\; A \in On)$ and $(A \notin V \;\Rightarrow\; A = On)$.

Proof:

Let $A \subseteq On$ be transitive. By 8.5a A is wellordered by $\in$; hence $(A \in V \;\Rightarrow\; A \in On)$.

If $A \notin V$, then we have: $\alpha \in On \;\Rightarrow\; A \not\subseteq \alpha \;\Rightarrow\; \exists \beta \in A (\beta \notin \alpha) \;\Rightarrow\; \exists \beta \in A (\alpha \leq \beta) \;\Rightarrow\; \alpha \in A$.

Lemma 8.7

(a) For each nonempty class $A \subseteq On$ we have $\bigcap A = \min(A)$, i.e. $\bigcap A \in A \;\&\; \forall \beta \in A(\bigcap A \leq \beta)$.

(b) For each set $A \subseteq On$ we have $\bigcup A \in On$ and $\bigcup A = \sup(A)$ $(:= \min\{x \in On : \forall y \in A(y \leq x)\})$.

(c) For each class $A \subseteq On$ the following holds: $A \in V \;\Leftrightarrow\; \exists \alpha \in On \forall x \in A(x \leq \alpha)$.

Proof:

(a) Let $\alpha := \min(A)$. We then have $\forall \beta \in A(\alpha \subseteq \beta)$ and $\alpha \in A$. This yields $\alpha \subseteq \bigcap A \subseteq \alpha$.

(b) Let $A \subseteq On$ and $A \in V$. Then $\bigcup A$ is a transitive set of ordinals; hence $\bigcup A \in On$ by Lemma 8.6. By definition of $\bigcup A$ and 8.4b we also have $\forall x \in On(\bigcup A \leq x \;\Leftrightarrow\; \forall y \in A(y \leq x))$, i.e. $\bigcup A = \min\{x \in On : \forall y \in A(y \leq x)\}$.

(c) "$\Rightarrow$": $A \in V \overset{(b)}{\Rightarrow} \bigcup A \in On$ and $\forall x \in A(x \leq \bigcup A)$.

"$\Leftarrow$": $\forall x \in A(x \leq \alpha) \;\Rightarrow\; A \subseteq \alpha \cup \{\alpha\} \;\Rightarrow\; A \in V$.

Definition

Let R be a wellordering on A.

An *ordering functions of* (A, R) is a function F, such that:

(i) $\mathrm{dom}(F) \in On$ or $\mathrm{dom}(F) = On$,

(ii) $\mathrm{ran}(F) = A$,

(iii) $\forall \alpha, \beta \in \mathrm{dom}(F)(\beta < \alpha \rightarrow F(\beta) R F(\alpha))$.

Theorem 8.8

If R is a wellordering on A, then there exists a unique ordering function F of (A, R).

It is $F(\alpha) = \min_R\{x \in A : F[\alpha] \subseteq x_R\}$ for all $\alpha \in \mathrm{dom}(F)$.

Further we have $(A \in V \Rightarrow \mathrm{dom}(F) \in On)$ und $(A \notin V \Rightarrow \mathrm{dom}(F) = On)$.

$\mathrm{dom}(F)$ is called the *ordertype* of (A, R).

Proof:

Uniqueness: Let F be an ordering function of (A, R).

(1) $\alpha \in \mathrm{dom}(F) \Rightarrow F(\alpha) = \min_R\{x \in A : F[\alpha] \subseteq x_R\}$.

Proof by transfinite induction on α: Let $\alpha \in \mathrm{dom}(F)$. By (iii) in the above definition we have $F[\alpha] \subseteq F(\alpha)_R$. Now let $x \in A$ such that $F[\alpha] \subseteq x_R$. Since $\mathrm{ran}(F) = A$, there exists a $\beta \in \mathrm{dom}(F)$ such that $F(\beta) = x$. Then $\forall \xi \in \alpha(F(\xi) R F(\beta))$ which by (iii) yields $\forall \xi \in \alpha(\xi < \beta)$. Hence $\alpha \leq \beta$ and $F(\alpha) \leq F(\beta) = x$.

Now assume that F_1, F_2 are ordering functions of A. Then w.l.o.g. $\mathrm{dom}(F_1) \subseteq \mathrm{dom}(F_2)$, using (1) by induction on α we obtain $F_1(\alpha) = F_2(\alpha)$ for all $\alpha \in \mathrm{dom}(F_1)$. It remains to prove $\mathrm{dom}(F_2) \subseteq \mathrm{dom}(F_1)$: $\alpha \in \mathrm{dom}(F_2) \;\Rightarrow\; F_2(\alpha) \in A \;\Rightarrow\; F_2(\alpha) = F_1(\beta) = F_2(\beta)$ for some $\beta \in \mathrm{dom}(F_1) \;\Rightarrow\; \alpha = \beta \in \mathrm{dom}(F_1)$.

Existence:

Definition (by transfinite recursion):

$F_1 : On \rightarrow V,\; F_1(\alpha) = \begin{cases} \min_R\{x \in A : F_1[\alpha] \subseteq x_R\} & \text{if } \exists x \in A(F_1[\alpha] \subseteq x_R) \\ 0 & \text{otherwise.} \end{cases}$

$D := \{\alpha \in On : \exists x \in A(F_1[\alpha] \subseteq x_R\}$ is a transitive class of ordinals $[\beta \in \alpha \in D \Rightarrow F_1[\beta] \subseteq F_1[\alpha] \subseteq x_R]$; hence $D \in On$ or $D = On$.

Claim: $F := F_1|D$ is ordering function of (A, R).
Obviously $\mathrm{dom}(F) = D$, $\mathrm{ran}(F) \subseteq A$ and $\forall \alpha, \beta \in D(\beta < \alpha \to F(\beta)RF(\alpha))$.

Remains to prove: $x \in A \Rightarrow x \in \mathrm{ran}(F)$.

Proof by R-induction: Let $x \in A$. Then by I.H. $x_R \subseteq \mathrm{ran}(F)$. Now $F^{-1}[x_R] = \{\beta \in D : F(\beta) \in x_R\}$ is transitive [$\gamma \in \beta$ & $F(\beta)Rx \Rightarrow F(\gamma)RF(\beta)Rx$]. Since F is injective, we have $F^{-1}[x_R] \in V$; hence $\alpha := F^{-1}[x_R] \in On$. Since $x_R \subseteq \mathrm{ran}(F)$, it follows that $F[\alpha] = x_R$. From this we obtain $\forall y \in A(yRx \Rightarrow F[\alpha] \not\subseteq y_R)$ and so $x = \min_R\{y \in A : F[\alpha] \subseteq y_R\} = F(\alpha)$.

Corollary

If A is a proper class (i.e., $A \notin V$) and R a wellordering on A, then there is a unique bijection $F : On \to A$ with $\forall \alpha, \beta(\beta < \alpha \leftrightarrow F(\beta)RF(\alpha))$.

Definitions

$0 := \emptyset$, $\quad x' := x \cup \{x\}$, $\quad \omega := \bigcap\{u : 0 \in u \wedge \forall x \in u(x' \in u)\}$

α is a *successor ordinal*, if $\exists \beta(\alpha = \beta')$.

α is a *limit ordinal*, if $\alpha \neq 0$ and $\neg\exists\beta(\alpha = \beta')$.

$Lim :=$ class of limit ordinals.

Theorem 8.9

(a) 0 is the least ordinal and $\forall \alpha(\alpha' = \min\{\beta : \alpha < \beta\} \in On)$.

(b) $\alpha \in Lim \Leftrightarrow 0 < \alpha \wedge \forall \beta(\beta < \alpha \to \beta' < \alpha)$.

(c) ω is the least limit ordinal.

(d) For every class C: $0 \in C \wedge \forall x \in \omega(x \in C \to x' \in C) \to \forall x \in \omega(x \in C)$. *(complete induction)*

(e) For $a_0 \in V$ and $G : B \to B$ there exists a unique function $F : \omega \to B$
such that $F(0) = a_0$ and $F(x') = G(F(x))$, for all $x \in \omega$.

Proof:

From the axiom of infinity it follows that the class $J := \{u : 0 \in u \wedge \forall x \in u(x' \in u)\}$ is nonempty. Hence:
$(*)$ $\omega = \bigcap J \in V$ & $0 \in \omega$ & $\forall x \in \omega(x' \in \omega)$.

(a) By $(*)$ we have $0 \in V$. Moreover $0 \subseteq On$ and 0 transitive; Hence $0 \in On$ by 8.6. That 0 is the *least* ordinal, follows by 8.4b. Obviously $\alpha' = \alpha \cup \{\alpha\}$ is a transitive subset of On; hence $\alpha' \in On$ by 8.6. By 8.4 we also get $\forall \beta(\alpha < \beta \Leftrightarrow \alpha' \leq \beta)$, i.e. $\alpha' = \min\{\beta : \alpha < \beta\}$.

(b) follows from $\forall \beta(\beta < \alpha \leftrightarrow \beta' \leq \alpha)$.

(c) Let $a := \{x \in \omega : x \in On \wedge x \subseteq \omega\}$. From $0 \in On \wedge \forall \alpha(\alpha' \in On)$ and $(*)$ it follows that $a \in J$ and so $\omega \subseteq a$, i.e. $\forall x \in \omega(x \in On \wedge x \subseteq \omega)$. Hence ω is a transitive set of ordinals, which by 8.6a implies $\omega \in On$. From $\omega \in On$, (b), $(*)$ we obtain $\omega \in Lim$. From (b) we also get $Lim \subseteq J$ and thus $\forall x \in Lim(\omega \subseteq x)$.

(d) From the premise and from $\omega \in V$ it follows that $u := \omega \cap C \in J$ and then $\omega \subseteq u \subseteq C$.

(e) Let $G_1 : \omega \times V \to V$, $G_1(x, f) := \begin{cases} a_0 & \text{if } x = 0 \vee \neg\mathrm{Fun}(f) \vee \cup x \notin \mathrm{dom}(f) \\ G(f(\cup x)) & \text{otherwise} \end{cases}$. By Theorem 8.3 there is a unique function $F : \omega \to V$ such that $\forall x \in \omega(F(x) = G_1(x, F|x))$. It follows that $F(0) = a_0$ and $F(x') = G_1(x', F|x') = G((F|x')(\cup x')) = G(F(x))$ for all $x \in \omega$. (Note that $\forall x \in On(\cup(x') = x)$.)

Definition

1. F is called a *choice function for* A $:\Leftrightarrow$ $\mathrm{Fun}(F)$ & $\forall x \in A(x \neq \emptyset \Rightarrow x \in \mathrm{dom}(F)$ & $F(x) \in x)$
2. A relation R is called a *partial ordering*, if R is irreflexive and transitive.
3. A set c is called an *R-chain*, if $\forall x, y \in c(Rxy \vee x = y \vee Ryx)$.

Theorem 8.10

The following statements are equivalent:

(AC) $\forall a \exists f($ f is a choice function for $a)$ *(Axiom of Choice)*

(WO) $\forall a \exists r($ r is a wellordering on $a)$ *(Wellordering Theorem)*

(ZL) For each nonempty, partially ordered set (a, r) the following holds: *(Zorn's Lemma)*
If every r-chain $c \subseteq a$ has an upper bound in a, then a has a maximal element.
[I.e., $\forall c \subseteq a(c$ is an r-chain $\rightarrow \exists b \in a \forall x \in c((x, b) \in r \vee x = b)) \Rightarrow \exists p \in a \neg \exists x \in a((p, x) \in r)$]

Proof:

(AC) $\Rightarrow$ (ZL): Let (a, r) be a nonempty, partially ordered set such that every r-chain $c \subseteq a$ has an upper bound in a. By (AC) there is a function $G : V \rightarrow V$ such that $\forall x \in \mathcal{P}(a)(x \neq \emptyset \rightarrow G(x) \in x)$. By transfinite recursion we define $F : On \rightarrow V$ with $F(\xi) = G(\{x \in a : F[\xi] \subseteq x_r\})$ for all $\xi \in On$. (Remember the abbreviation $x_r := \{y : (y, x) \in r\}$.) Let $D := \{\xi : \{x \in a : F[\xi] \subseteq x_r\} \neq \emptyset\}$.

Then we have:

(1) D transitive,

(2) $F[D] \subseteq a$,

(3) $\forall \eta, \xi \in D(\eta < \xi \rightarrow F(\eta) r F(\xi))$.

By (2) and (3), $F[D] \in V$ and $F|D$ injective, hence $D \in V$. By (1) it follows that $D \in On$. From $D \in On \setminus D$ we conclude $\{x \in a : F[D] \subseteq x_r\} = \emptyset$. By (3) $F[D]$ is an r-chain. Let $p \in a$ be an upper bound of $F[D]$. Since r is transitive, we then have $\neg \exists x \in a((p, x) \in r)$.

(ZL) $\Rightarrow$ (AC): Let $a \in V$. We set $W := \{f : \mathrm{Fun}(f) \wedge \mathrm{dom}(f) \subseteq a \wedge \forall x \in \mathrm{dom}(f)(f(x) \in x)\}$. W is a set, since $W \subseteq \mathcal{P}(a \times \bigcup a)$. Therefore we can apply (ZL) to $(W, \subsetneq)$. Obviously every $\subsetneq$-chain $c \subseteq W$ has an upper bound in W, namely $\bigcup c$. Hence there exists a $\subsetneq$-maximal element $f_0 \in W$.
Assumption: There is an $x \in a \setminus \{\emptyset\}$ with $x \notin \mathrm{dom}(f_0)$. Let $y \in x$. Then $f := f_0 \cup \{(x, y)\} \in W$ and $f_0 \subsetneq f$. *Contradiction.* Therefore $a \setminus \{\emptyset\} \subseteq \mathrm{dom}(f_0)$, and f_0 is a choice function for a.

(AC) $\Rightarrow$ (WO): Let $a \in V$. By (AC) there exists a $G : V \rightarrow V$ such that $\forall x \in \mathcal{P}(a)(x \neq \emptyset \rightarrow G(x) \in x)$. By transfinite recursion we define $F : On \rightarrow V$ such that $F(\xi) = G(a \setminus F[\xi])$. Let $D := \{\xi : a \setminus F[\xi] \neq \emptyset\}$.
Then the following holds:
(1) D transitive,
Proof: $\xi \in D \wedge \beta \in \xi \Rightarrow \emptyset \neq a \setminus F[\xi] \subseteq a \setminus F[\beta]$.

(2) $F[D] \subseteq a$ and $F|D$ injective,
Proof: $\xi \in D \Rightarrow a \setminus F[\xi] \in \mathcal{P}(a) \setminus \{\emptyset\} \Rightarrow F(\xi) = G(a \setminus F[\xi]) \in a \setminus F[\xi]$.

(3) $D \in On$,

Proof: $D \subseteq On = \mathrm{dom}(F)$ & $F[D] \subseteq a \;\Rightarrow\; D \subseteq On$ & $D \in V \;\Rightarrow\; D \in On$.

(4) $a \subseteq F[D]$.

Proof: $D \in On \Rightarrow D \in On \setminus D = \{\xi : a \setminus F[\xi] = \emptyset\} \Rightarrow a \setminus F[D] = \emptyset \Rightarrow a \subseteq F[D]$.

So, $F|D$ is a bijection from the ordinal D onto a. Hence $r := \{(F(\eta), F(\xi)) : \xi, \eta \in D \wedge \eta < \xi\}$ is a wellordering on a.

(WO) $\Rightarrow$ (AC): Let $a \in V$. By (WO) there exists a wellordering r on $\bigcup a$.
Then $f : a \setminus \{\emptyset\} \to \cup a,\ f(x) := \min_r(x)$ is a choice function for a.

We are now going to show that in the proofs of Lemma 8.2 and Theorem 8.3 the assumption "R transitive" is not needed.

Definition (transitive closure of a relation R)

$R^+ := \{(y, x) : \exists f \exists n > 0\, R(f, n, y, x)\}$ where

$R(f, n, y, x) \;:\Leftrightarrow\; \mathrm{Fun}(f) \wedge \mathrm{dom}(f) = n' \in \omega \wedge f(0) = x \wedge f(n) = y \wedge \forall i < n (f(i') R f(i))$

Lemma 8.11

For every relation R we have:

(a) $R \subseteq R^+$ and R^+ is transitive.

(b) $R \subseteq Q \subseteq V \times V$ & Q transitive $\Rightarrow\; R^+ \subseteq Q$.

(c) $R \in V \;\Rightarrow\; R^+ \in V$.

(d) $\forall x (x_R \in V) \;\Rightarrow\; \forall x (x_{R^+} \in V)$.

Proof:

(a) $R \subseteq R^+$ is trivial. — Assume now $y R^+ x$. By induction on n we prove: $R(f, n, z, y) \Rightarrow z R^+ x$.

1. $n = 0$: Then $z = y$.

2. $n \to n'$: Assume $R(f, n', z, y)$ and let $u := f(n)$. Then $R(f|n', n, u, y)$ and by I.H. $u R^+ x$. So we have $R(g, m, u, x)$ for some g, m. Together with $z R u$ this yields $R(g_1, m', z, x)$ for $g_1 := g \cup \{(m', z)\}$.

(b) By induction on n one proves: $R(f, n, y, x)$ & $n > 0 \;\Rightarrow\; (y, x) \in Q$.

(c) Let $R \in V$. Then also $a := \mathrm{dom}(R) \cup \mathrm{ran}(R)$ and $a \times a$ are sets. Further $a \times a$ is a transitive relation with $R \subseteq a \times a$. Therefore by (b) $R^+ \subseteq a \times a$ which yields $R^+ \in V$.

(d) Let $a \in V$. Definition of $h : \omega \to V$: $h(0) := \{a\},\ h(n') := \bigcup \{x_R : x \in h(n)\}$.
By induction on n we obtain $R(f, n, y, a) \Rightarrow y \in h(n)$, and thus $a_{R^+} \subseteq \bigcup_{n \in \omega} h(n) = \bigcup \mathrm{ran}(h) \in V$.

Proof of 8.2 for arbitrary wellfounded R.

Claim: Every nonempty class C has an R-minimal element.

Assume $u \in C$. Case 1: $u_R \cap C = \emptyset$. Then we are done.
Case 2: $u_R \cap C \neq \emptyset$. Then $a := u_{R^+} \cap C \neq \emptyset$. Since R is wellfounded, there exists a $c \in a$ such that $c_R \cap a = \emptyset$. From $c \in u_{R^+}$ we get $c_R \subseteq u_{R^+}$. Hence $c \in C$ and $c_R \cap C \subseteq c_R \cap a = \emptyset$.

Lemma 8.12

If R is wellfounded then R^+ is wellfounded.

Proof:

By 8.1 (and 8.11d) it suffices to prove the induction principle for R^+.

Assume (1) $\forall x(x_{R^+} \subseteq C \to x \in C)$. To prove: $\forall x(x \in C)$.

Let $\overline{C} := \{x : x_{R^+} \subseteq C\}$. Then by (1) we have (2) $\overline{C} \subseteq C$. Further: (3) $\forall x(x_R \subseteq \overline{C} \to x \in \overline{C})$.

[Proof of (3): $x_R \subseteq \overline{C} \overset{(2)}{\Rightarrow} x_R \subseteq C \wedge \forall y \in x_R(y_{R^+} \subseteq C) \Rightarrow x_{R^+} \subseteq C \Rightarrow x \in \overline{C}$.]

R wellfounded $\overset{8.1,8.2,(3)}{\Longrightarrow} \forall x(x \in \overline{C}) \overset{(2)}{\Longrightarrow} \forall x(x \in C)$.

Proof of 8.3 for arbitrary wellfounded R.

Claim: For $G : A \times V \to V$ there exists a unique $F : A \to V$ with $F(x) = G(x, F|x_R)$ for all $x \in A$.

We define $G' : A \times V \to V$, $G'(x, f) := G(x, f|x_R)$. By 8.12 and 8.11 R^+ is wellfounded and transitive. Therefore, by 8.3 (for transitive relations), there exists a unique $F : A \to V$ with $F(x) = G'(x, F|x_{R^+})$ for all $x \in A$. But $G'(x, F|x_{R^+}) = G(x, F|x_R)$.

Since the relation $\in$ is wellfounded (due to the axiom of foundation) we obtain from 8.2 and 8.3:

($\in$-induction) $\forall x(x \subseteq C \to x \in C) \to \forall x(x \in C)$

($\in$-recursion) For $G : V \to V$ there exists a unique $F : V \to V$ such that $F(x) = G(F|x)$ for all x.

9 Cardinal numbers

Abbreviations:

$F : A \longleftrightarrow B :\Leftrightarrow F : A \to B$ injective $\& F[A] = B$,

$a \sim b :\Leftrightarrow \exists f(f : a \longleftrightarrow b)$,

$a \preceq b :\Leftrightarrow \exists f(f : a \to b$ injective).

Lemma 9.1

For each set a the following statements are equivalent:

(i) $\exists r(r$ is wellordering of $a)$,

(ii) $\exists \alpha \in On(\alpha \sim a)$,

(iii) $\exists \alpha \in On \exists f(f : \alpha \to a \,\&\, f[\alpha] = a)$,

(iv) $\exists \alpha \in On(a \preceq \alpha)$.

Proof:

(i)⇒(ii): 8.8. (ii)⇒(iii): trivial. (iii)⇒(iv): $g(x) := \min\{\xi \in \alpha : f(\xi) = x\}$.

(iv)⇒(i): $r := \{(y, x) \in a \times a : g(y) < g(x)\}$ where $g : a \to \alpha$ injective.

Lemma 9.2

(a) $a \preceq a$,

(b) $a \preceq b \,\&\, b \preceq c \Rightarrow a \preceq c$,

(c) $a \subseteq b \Rightarrow a \preceq b$,

(d) $\sim$ is an equivalence relation on V.

Theorem 9.3 (Cantor-Bernstein)

$a \preceq b \,\&\, b \preceq a \;\Rightarrow\; a \sim b.$

Proof:

W.l.o.g.: $b \subseteq a$. Let $f : a \to b$ be injective. We define:

$h : \omega \to \mathcal{P}(a),\ h(0) := a \setminus b,\ h(n') := f[h(n)], \quad \hat{a} := \bigcup \mathrm{ran}(h) = \bigcup_{n\in\omega} h(n).$

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

(1) $\forall x \in \hat{a}(g(x) \in \hat{a})$

Proof: $x \in h(n) \Rightarrow g(x) = f(x) \in h(n') \subseteq \hat{a}.$

(2) $g[a] \subseteq b$

Proof: $x \in \hat{a} \Rightarrow g(x) = f(x) \in b. \quad x \in a \setminus \hat{a} \Rightarrow g(x) = x \in a \setminus h(0) = b.$

(3) g injective

Proof by case distinction: If $x \in \hat{a}$ and $y \notin \hat{a}$, then $g(x) \in \hat{a}$ and $g(y) = y \notin \hat{a}$, hence $g(x) \neq g(y)$. The other cases are trivial.

(4) $g[a] = b$

Proof: Let $y \in b$. If $y \notin \hat{a}$, then $g(y) = y$. Assume now that $y \in \hat{a}$, i.e. $y \in h(n)$ for some $n \in \omega$. Then $n > 0$, since $y \in b$. Hence $n = k'$ and $y = f(x)$ with $x \in h(k) \subseteq \hat{a}$. This yields $y = g(x)$.

Theorem 9.4 $\mathcal{P}(a) \not\preceq a.$

Proof:

Obviously $a \preceq \mathcal{P}(a)$. Therefore by 9.3 ($\mathcal{P}(a) \preceq a \;\Rightarrow\; \mathcal{P}(a) \sim a$). But $\mathcal{P}(a) \sim a$ cannot hold, since if $f : a \to \mathcal{P}(a)$, then $u := \{x \in a : x \notin f(x)\} \in \mathcal{P}(a) \setminus f[a]$.

Definition

An ordinal number α is called a *cardinal (number)*, if $\neg\exists\beta < \alpha(\alpha \sim \beta)$.

$Card :=$ class of all cardinal numbers.

$|a| := \begin{cases} \min\{\xi \in On : a \sim \xi\} & \text{if } \exists\xi(a \sim \xi) \\ undefined & \text{otherwise} \end{cases}$

Remark

If $|a|$ is defined (i.e. $\exists\xi(a \sim \xi)$), then $|a|$ is a cardinal number, *the cardinality of* a. In the following we write $|a| \in Card$ to express that $|a|$ is defined. Obviously $\forall\alpha(|\alpha| \in Card \,\&\, |\alpha| \leq \alpha)$.

Lemma 9.5.

(a) $\alpha \in Card \;\Leftrightarrow\; \neg\exists\xi < \alpha(\alpha \preceq \xi).$

If $|a| \in Card$ then

(b) $|a| = \min\{\xi \in On : a \preceq \xi\},$

(c) $a \sim b \;\Leftrightarrow\; |a| = |b|,$

(d) $b \preceq a \;\Leftrightarrow\; |b| \leq |a|,$

(e) $|F[a]| \leq |a|$, for each function F.

Proof:

(a) "$\Leftarrow$" trivial. "$\Rightarrow$" $\xi < \alpha \,\&\, \alpha \preceq \xi \overset{9.3}{\Rightarrow} \alpha \sim \xi \Rightarrow \alpha \leq \xi$. *Contradiction.*

(b) $\xi < |a| \overset{(a)}{\Rightarrow} |a| \not\preceq \xi \Rightarrow a \not\preceq \xi$.

(c) trivial.

(d) "$\Leftarrow$" trivial. "$\Rightarrow$" $a \sim |a| \,\&\, b \preceq a \Rightarrow b \preceq |a| \overset{9.1}{\Rightarrow} |b| \in Card \,\&\, b \preceq |a| \overset{(b)}{\Rightarrow} |b| \leq |a|$.

(e) By 9.1 there exists a wellordering r of a. We define: $h : F[a] \to a$, $h(y) := \min_r\{x \in a : F(x) = y\}$. Obviously h is injective and so $F[a] \preceq a$, which yields $|F[a]| \leq |a|$ by (d).

Lemma 9.6

If $\kappa \in Card$ and $\kappa^+ = \min\{\mu \in Card : \kappa < \mu\}$, then the following holds for all $\alpha \in On$:

(a) $|\alpha| < \kappa \Rightarrow \alpha < \kappa$,

(b) $|\alpha| > \kappa \Rightarrow \alpha \geq \kappa^+$,

(c) $|\alpha| = \kappa \Rightarrow \kappa \leq \alpha < \kappa^+$.

Proof:

(a) $|\alpha| < \kappa \Rightarrow \neg(\kappa \preceq \alpha) \Rightarrow \alpha < \kappa$.

(b) $|\alpha| > \kappa \Rightarrow \alpha \geq |\alpha| \geq \kappa^+$.

(c) $\kappa = |\alpha| \leq \alpha$. $|\alpha| < \kappa^+ \overset{(a)}{\Rightarrow} \alpha < \kappa^+$.

Definition $\alpha^+ := \{\xi \in On : \xi \preceq \alpha\}$

Theorem 9.7 $\alpha^+ = \min\{\kappa \in Card : \alpha < \kappa\}$.

Proof:

Let $W := \{(b, r) : b \subseteq \alpha \wedge r \text{ is wellordering of } b\}$. Obviously $W \subseteq \mathcal{P}(\alpha) \times \mathcal{P}(\alpha \times \alpha)$ and therefore $W \in V$.

Definition: $\mathbf{ot} : W \to On$, $\mathbf{ot}((b, r)) :=$ the order type of (b, r) (cf. Theorem 8.8).

Then $\alpha^+ = \{\mathbf{ot}(x) : x \in W\}$ and thus $\alpha^+ \in V$. [Proof of $\alpha^+ \subseteq \{\mathbf{ot}(x) : x \in W\}$: $f : \xi \overset{1\text{-}1}{\longrightarrow} \alpha \,\&\, b := f[\xi] \,\&\, r := \{(f(x), f(y)) : x < y < \xi\} \Rightarrow (b, r) \in W \,\&\, \xi = \mathbf{ot}((b, r))$.] Since $\alpha^+ \subseteq On$ is transitive, it follows that $\alpha^+ \in On$. Further we have $\alpha^+ \not\preceq \alpha$ and $\forall \xi \in \alpha^+(\xi \preceq \alpha)$; hence $\forall \xi \in \alpha^+(\alpha^+ \not\preceq \xi)$, i.e. $\alpha^+ \in Card$. From $\alpha \preceq \alpha$ it follows that $\alpha < \alpha^+$. If $\alpha < \kappa \in Card$, then $\forall \xi \geq \kappa(\xi \not\preceq \alpha)$ and thus $\forall \xi < \alpha^+(\xi < \kappa)$, i.e. $\alpha^+ \leq \kappa$.

Definition

By 9.7, $Card \setminus \omega$ is unbounded in On and therefore $Card \setminus \omega \notin V$.

Let $\alpha \mapsto \aleph_\alpha$ ($\alpha \in On$) be the ordering function of the class $Card \setminus \omega$.

(Especially $\aleph_0 = \omega$.)

Notation. $\alpha{+}1 := \alpha'$, $1 := 0'$, $2 := 1'$, ...

Lemma 9.8

(a) The class $Card$ is closed, i.e. $\forall u(\emptyset \neq u \subseteq Card \Rightarrow \sup(u) \in Card)$.

(b) $\omega \leq \kappa \in Card \Rightarrow \kappa \in Lim$.

(c) $\forall \alpha(\alpha \leq \aleph_\alpha) \,\&\, \forall \alpha \in Lim(\aleph_\alpha = \sup_{\xi < \alpha} \aleph_\xi)$.

Proof:

(a) Let $\gamma := \sup(u)$ and $f : \gamma \to \alpha$ injective. Then $\forall \xi \in u(\xi \preceq \alpha)$. Since $u \subseteq Card$, we obtain $\forall \xi \in u(\xi \leq \alpha)$ and then $\gamma = \sup(u) \leq \alpha$.

(b) Let $\omega \leq \beta{+}1$. Definition: $f : \beta{+}1 \to \beta$, $f(x) := \begin{cases} 0 & \text{if } x = \beta \\ x{+}1 & \text{if } x \in \omega \\ x & \text{otherwise.} \end{cases}$

Obviously f is injective, and therefore $\beta{+}1$ is not a cardinal.

(c) cf. L11.1d and L.11.2.

Definition

a is *finite* $:\Leftrightarrow$ $|a| < \omega$ (i.e. $\exists k \in \omega(a \sim k)$);

a is *D-finite* (D for Dedekind) $:\Leftrightarrow$ $\neg\exists f(f : a \to a$ injective and $f[a] \neq a)$;

a (D-)infinite $:\Leftrightarrow$ a not (D-)finite.

Remark. a, b finite $\Longrightarrow$ $a \cup b$ and $a \times b$ finite. [Proof by induction on $|b|$.]

Lemma 9.9

(a) $b \preceq a$ & a finite $\Rightarrow$ b finite

(b) $a \subseteq \omega \Rightarrow (a$ finite $\Leftrightarrow \exists x \in \omega(a \subseteq x))$

(c) ω is the least infinite cardinal number.

(d) a D-infinite $\Leftrightarrow$ $\omega \preceq a$

(e) a finite $\Rightarrow$ a D-finite

(f) $\omega \subseteq Card$

(g) (AC) $\Rightarrow$ (a D-finite $\Leftrightarrow$ a finite)

Proof:

(a) follows from 9.5d.

(b) "$\Rightarrow$" Induction on $|a|$. "$\Leftarrow$" follows from (a).

(c) *Assumption*: ω finite. By (b) we then have $\omega \subseteq k$ for some $k \in \omega$; hence $k \in k$. *Contradiction.*
So ω is infinite, and therefore $\neg\exists \alpha \in \omega(\alpha \sim \omega)$, i.e. $\omega \in Card$.

(d) "$\Rightarrow$" Let $f : a \to a$ be injective with $f[a] \neq a$. We choose an $x_0 \in a \setminus f[a]$ and define $g : \omega \to a$, $g(0) := x_0$, $g(n{+}1) := f(g(n))$. By complete induction we obtain $\forall n \in \omega \forall i \in n(g(n) \neq g(i))$, i.e. g is injective.

"$\Leftarrow$" Let $g : \omega \to a$ be injective. Then $f := \{(g(i), g(i{+}1)) : i \in \omega\} \cup \{(x, x) : x \in a \setminus g[\omega]\}$ is an injective function from a into a with $g(0) \notin f[a]$. Hence a is D-infinite.

(e) If a is D-infinite, then by (d) $\omega \preceq a$, and by (c) and (a) it follows, that a is infinite.

(f) Let $n \in \omega$, $m \leq n$ and $f : n \to m$ injective. To prove: $m = n$. By (e) n is D-finite. Since $f[n] \subseteq m \subseteq n$, we therefore have $n = f[n] \subseteq m$ and thus $m = n$.

(g) "$\Rightarrow$" We prove: "a infinite $\Rightarrow$ a D-infinite".
Let h be a choice function for $\mathcal{P}(a)$ with $\emptyset \in \mathrm{dom}(h)$. By recursion over ω we define $g\colon \omega \to V$,
$g(n) := h(a \setminus g[n])$. Then $\forall n \in \omega(g[n]$ finite) by 9.5e. Since by assumption a is infinite, we obtain $\forall n \in \omega(a \setminus g[n] \neq \emptyset)$. Hence g is an injective function from ω into a, and by (d) a is D-infinite.