Virgil Drăghici

MATHEMATICAL LOGIC

Presa Universitară Clujeană

Virgil Drăghici

•

MATHEMATICAL LOGIC

Virgil Drăghici

MATHEMATICAL LOGIC

PRESA UNIVERSITARĂ CLUJEANĂ
2023

Referenți științifici:

Prof. univ. dr. Ştefan Cobzaş

Prof. univ. dr. Ionel Narița

ISBN 978-606-37-1938-7

Universitatea Babeş-Bolyai
Presa Universitară Clujeană
Director: Codruța Săcelean
Str. Hasdeu nr. 51
400371 Cluj-Napoca, România
Tel./fax: (+40)-264-597.401
E-mail: editura@editura.ubbcluj.ro
http://www.editura.ubbcluj.ro/

Preface

The present book is an introduction to the main subjects of what is usually called mathematical logic. It is structured in the following chapters: *Propositional Logic*, *First-order Logic*, *Formal Number Theory* and *Modal Logic of Provability*.[1] The leading aim of the whole construction is to get a clear idea of what a First-order Logic (FOL) is (Chapters 1 and 2), a proper extension of the First-order Logic axiomatized (FOL^{ax}) in a first-order theory called Peano Arithmetic axiomatized (PA^{ax}) (whose key part is the analysis of the Gödel's theorems, in a variety of forms) (Chapter 3) and, finally, a modal analysis of the provability predicate Bew(x) of PA^{ax} in terms of the propositional modal system $\mathcal{GL}$ (Chapter 4).

Let us detail.

Propositional Logic (PL). The analysis of this simplest part of mathematical logic is given in both ways, *semantical* and *syntactic*.[2]

Semantical treatment of PL concerns the definitions and the behavior of the semantic notions (e.g., truth function, validity, satisfiability, unsatisfiability, semantical consequence), the proof of some basic theorems of PL (e.g. Substitution Theorem, Replacement Theorem, Duality Theorem, Normality Theorem, Interpolation Theorem) and the discussion of some decision procedures in PL (e.g. truth table method, Quine's method, normal forms), accompanied by relevant examples and proofs.

Syntactical treatment of PL takes as basis of considerations the Lukasiewicz's axiomatic system PL^{ax}, related to which some fundamental results are accurately stated and proved (e.g., Substitution Theorem, Replacement Theorem and Deduction Theorem).

Finally, the analysis of PL ends with stating and proving the theorems connecting both ways of analysis of PL, semantical and syntactic: Soundness Theorem, Completeness Theorem and, as a general result about PL, Decidability Theorem for PL.

---

[1] The most part of these matters represents the content of my university lectures on the specified fields (*Symbolic Logic*, *Modal Logic* and *Recursion Theory*), given at Babeş-Bolyai University Cluj-Napoca.

[2] Also known as "model-theoretic" and "proof-theoretic" treatments of PL (comp. *inter alia* S.C. Kleene [1968]).

First-order Logic (FOL). FOL is an extension of PL and, similarly, it is treated from a double perspective, semantical and syntactic.

Semantical approach refers to the main semantical notions (e.g. a model *M* of a first-order language, satisfiability of a formula α in *M*, truth of α in *M*, validity of α in FOL), methods of testing the validity of a formula α, some basic results in the metatheory of FOL, used later on in stating and proving some theorems and rules of deduction, and the proofs of some fundamental facts of FOL e.g. Duality Theorem, Interpolation Theorem and Beth's Definability Theorem.

Syntactical approach deals with the axiomatic construction FOL^{ax} of FOL, a formal system which extends the formal system PL^{ax} of PL. The main metatheoretical results for FOL^{ax} are stated and integrally proved, e.g. Deduction Theorem, Substitution Theorems, Replacement Theorem, Choice Rule, prenex and Skolem normal forms in FOL^{ax}.

Similar to the analysis of PL, the ways the semantic and syntactic parts are connected, is the subject of two important metatheoretical results: Soundness Theorem and Completeness Theorem for FOL^{ax}. Actually, the completeness and the Löwenheim-Skolem theorems are derived from a basic and more general result concerning the first-order theories, according to which every consistent first-order theory has a model of $\aleph_0$-cardinality. All these results are integrally proved.

Finally, the analysis of FOL ends with the fundamental result known as Church's Undecidability Theorem. But since this section calls for some results of the next chapter (Chapter 3), mainly the notion of essentially recursively undecidable theory, these matters will be perfectly understandable after the reading of Chapter 3.

Formal number theory. This part of the book intends essentially to do an account of a formal system, usually known as Peano Arithmetic axiomatized (PA^{ax}), and then, with reference to this system, to analyse a variety of ways the Gödel's results concerning the incompleteness and undecidability can be formulated and proved. Actually, the first part of this chapter (Sect. 1-3) represents just a preliminary of the second part (Sect. 4), focused on Gödel's theorems.

In short, the first part contains an exposition of the formal system PA^{ax} and a lot of theorems and metatheorems proved within and about PA^{ax} (Sect. 1), some basic notions and number-theoretic functions and relations (Sect. 2), and then some fundamental notions on the recursive functions and relations, including the essential fact of the formal representability (expressibility) of the recursive functions (relations) in PA^{ax} (Sect. 3).

The second part begins with an account of some of the most relevant ways of constructing an undecidable sentence and, correspondingly, in formulating and proving the Gödel's theorem (and its Rosser's form). In all these forms (with or without Diagonal Lemma, or *via* paradoxes) the *diagonalization* plays a key role.

The analysis goes on with an exposal of *Kleene's generalized forms of Gödel's theorems*, on which Smullyan's and Kripke's subsequent investigations are essentially based. Let us review in brief these results.

Kleene's considerations are based on his primitive recursive predicate $T(z, x_1,...,x_n, y)$ which, in turn is based on the Herbrand-Gödel idea of defining the computable functions in terms of the systems of equations. In fact, this predicate is a key notion of a lot of important results, e.g. *inter alia* the Enumeration Theorem. This theorem, *via* diagonalization, leads to a result on which some other significant fact is based, i.e., a form of Church's Theorem, according to which there is no algorithm for either of the predicates $(y)\widetilde{T}(x,x,y)$ or $(Ey)T(x,x,y)$. On the other hand, if $R(x,y)$ is the primitive recursive predicate whose meaning is "*y* is a proof of the formula $\mathrm{A}(\bar{x})$, then $(Ey)R(x,y)$ means "$\vdash \mathrm{A}(\bar{x})$" (i.e., $\mathrm{A}(\bar{x})$ is provable). And therefore if A(x) expresses a predicate $P(x)$, then the expression $\vdash \mathrm{A}(\bar{x}) \leftrightarrow P(x)$ means that the formal system we are referring to is both *correct* and *complete* for $P(x)$. Now, if $P(x)$ is $(y)\widetilde{T}(x,x,y)$, then three generalizations of the first Gödel's incompleteness theorem can be constructed, whose idea is that for this predicate there is no correct and complete formal system.

Using, again, the *T*-predicate, Kleene also gives a generalization of a Rosser form of Gödel's theorem, by the so-called *symmetric form* of Gödel's theorem.

R. Smullyan's and S. Kripke's strategies of deriving undecidability theorems, based on Gödel-Kleene results, apply the idea of recursive enumerability with the ideas of separability and recursive inseparability. Briefly, the way these matters are correlated is the following. Two other remarkable results of mathematical logic are Kleene's Enumeration Theorem for *partial* recursive functions, according to which there is a partial recursive function $\Phi(z, x_1,...,x_n)$ such that for $z = 0,1,2,...$ it gives an enumeration of the *n*-place partial recursive functions, and Kleene's

S^m_n-Theorem. As regards the first result, it is used both in a proof of Diagonal Lemma, using the recursion theory, and in the construction of the so called Kleene's recursively inseparable pair (S_1,S_2) of recursively enumerable sets, a result on which a form of Rosser's theorem is based and proved. The second result, S^m_n-theorem, is relevant both in giving a distinct proof of Gödel's theorem and in the derivation of a fundamental result of the recursion theory: Kleene's Second Recursion Theorem.

Finally, two short sections refer to some other ways of investigation the incompleteness phenomenon. One of these ways is based on the Post's idea of *creative sets*, used in stating and proving a form of incompleteness theorem for any correct and axiomatizable system S. The other way, due to Smullyan, is based on the idea of the complete effective inseparability of a disjoint pair of sets. Using a Kleene function for a recursively enumerable pair of sets, it represents an interesting proof of Kleene's symmetric form of Gödel's theorem.

Modal logic of provability. This logic, as we said, concerns basically the study of the mathematical notion "provability" using the modal apparatus of a modal system. Essentially, it deals with the relationship between two formal systems: PA^{ax} (Peano Arithmetic axiomatized) and $\mathcal{GL}$ (the Gödel-Löb system of modal propositional logic).

The first section (4.1) of this chapter is an exposing of the best-known modal systems, including $\mathcal{GL}$, by considering them both syntactical (i.e., regarding them axiomatically and by proving some theorems) and semantical (using the corresponding notions: frame, model, validity, relativized to each of these systems S, accompanied with the respective theorems, connecting the S-validity with the corresponding properties of S-frames).

Similar to the classical systems PL^{ax} and FOL^{ax}, a relevant question is that of proving the *soundness* and *completeness* of these systems. The soundness easily follows from the theorems mentioned above, theorems connecting S-validity with the respective properties of S-frames. For the proof of completeness two kind of techniques are applied, that of the canonical models and that of the finite models. The use of the later is called for by the fact the system $\mathcal{GL}$, the basis of our considerations in the next section (4.2), is non-canonical, and then the proof of its completeness using canonical models cannot be applied.

The second section (4.2) of this chapter is a concise account of the idea of a modal logic of provability, whose aim is the study of the behavior of "Bew(x)" (the provability predicate for PA^{ax}) using its modal counterpart, expressed by the operator "□" (necessity) (of the modal system $\mathcal{GL}$). Actually, two fundamental items concerning the relationship $\mathcal{GL}$-PA^{ax} make the subject of this account. On the one hand, a result referring to the *arithmetical soundness* and *arithmetical completeness* (due to R. Solovay) of $\mathcal{GL}$, with some of the most relevant consequence for PA^{ax}; and on the other hand, the so-called *Fixed point theorem* (due to Dick de Jongh and Giovanni Sambin).

The elaboration of this book is based on the works of the most notable authors in the above treated fields of mathematical logic. A special indebtedness is acknowledged to the following authors: D. Hilbert and W. Ackermann, H. Scholz and G. Hasenjaeger, K. Gödel, S.C. Kleene, R.M. Smullyan, G. Boolos, G.E. Hughes and M.J. Cresswell, C. Smorynski, whose works were extensively used in the present book. The aim of the book is to give an analysis of the key concepts, characteristic of these fields, all the essential theorems being accompanied by their proofs, given integrally, sometimes by introducing versions of the proofs, illustrating them by examples or by taking them from the original sources of the respective cited authors, in order for this book to be self-contained and maximally explanatory.

This book is for undergraduate, graduate, PhD students and researchers. But having the traits of a textbook, it may be consulted by anyone interested in the mathematical logic.

V.D.

Table of Contents

Chapter 1. PROPOSITIONAL LOGIC

Unlike the traditional syllogistic, in which the structure of a sentence was a part of its analysis, in the propositional logic this aspect is not taken into consideration. Nor the content of a sentence is of some importance here. All that matters in the classical propositional logic is that sentences are either true or false, how are they combined to form more elaborate sentences and how the truth value of a compound sentence depends on the truth values of their components.

Like any other logic the propositional logic has its syntax and its semantics. Let us refer to the classical propositional logic by PL, and to its language by L_{PL}.

1. Syntax of PL

If, for example, p is the sentence *The chalk is white*[1] and q is the sentence *The sky is blue*, then some other sentences can be constructed, like: *The chalk is not white*, *The chalk is white and the sky is blue*, *The chalk is white or the sky is blue*, *If the chalk is white, then the sky is blue*, *The chalk is white if and only if the sky is blue*, etc. In the first case the operation performed was the *negation* ($\neg$) of the sentence p, in the second we connected p and q by *conjunction* ($\wedge$), then by *disjunction* ($\vee$), by *implication* ($\supset$) and by *equivalence* ($\equiv$), respectively. Since the meaning of these sentences doesn't matter, we may represent them by the following symbolic constructions: $\neg p$, $p\wedge q$, $p\vee q$, $p\supset q$, $p\equiv q$. These are formulas, in which p, q are *propositional variables* and $\neg$, $\wedge$, $\vee$, $\supset$ and $\equiv$ *propositional connectives*. But nothing prevents us from constructing more elaborate formulas, using parentheses, like $(p\supset q)\supset r$, $(p\wedge\neg q)\vee s$, $(q\equiv r)\wedge p$, etc. The symbols used in such formulas are the symbols of the language L_{PL}, hence L_{PL} contains the following symbols:

1. The symbols for propositional variables[2]: p_1, p_2, p_3…, written sometimes informally as p, q, r,...

---

[1] We refer to such sentences by writing them in *italics*, without using the quotation marks; the same will hold for other references.

[2] We'll call them simply "variables".

2. The symbols for propositional connectives[3]: $\neg, \wedge, \vee, \supset, \equiv$[4].

3. Auxiliary symbols: (,), written sometimes informally as [,], {,}.

The notion "formula of L_{PL}" will be defined recursively, i.e., using rules allowing to construct more elaborate formulas from those already given. By $\alpha_1, \alpha_2, \alpha_3\ldots$, written sometimes informally as $\alpha, \beta, \gamma,\ldots$ we understand arbitrary formulas of L_{PL}.

Definition 1. a) *Any variable of* L_{PL} *is a formula of* L_{PL} (*it is also called an atomic formula of* L_{PL}).

b) *If* α *is a formula of* L_{PL}*, then* $\neg\alpha$ *is a formula of* L_{PL}.

c) *If* α *and* β *are formulas of* L_{PL}*, then* $\alpha \circ \beta$ *is a formula of* L_{PL}*, where* "$\circ$" *denotes anyone of the connectives* $\wedge, \vee, \supset, \equiv$.

Are formulas of L_{PL} only those syntactic constructions given by a), b) and c).

In what follows by $\Gamma_1, \Gamma_2, \Gamma_3\ldots$, written sometimes informally as Γ, $\Delta, \Lambda,\ldots$, we understand arbitrary sets of formulas of L_{PL}.

Definition 2. *The complexity of a formula* α[5] *of* L_{PL} (compl(α), for short) *is the number of the occurrences of connectives in* α, *i.e.,*

a) *Any variable has the complexity zero.*

b) *If* α *has the complexity* n, *then* $\neg\alpha$ *has the complexity* $n+1$.

c) *If* α *and* β *have the complexities* m *and* n, *respectively, then* $\alpha \circ \beta$ *has the complexity* $m+n+1$.

2. Semantics of PL

2.1. Notions

2.1.1. Truth functions

The propositional logic we study in this section is the classical propositional logic, that is, it is a two valued logic. While the syntax of PL shows the "grammatical" mechanism of constructing its admissible entities, the semantics is interested in the meaning of these entities, and this requires the consideration of the truth values: 1 (true) and 0 (false). To interpret a variable p means to assign it a truth value, 1 or 0. To interpret a connective,

[3] We'll call them simply "connectives" or "operators".

[4] We shall sometimes use + (exclusion), / (incompatibility) and $\downarrow$ (rejection) as abbreviations for anti-equivalence, anti-conjunction and anti-disjunction, respectively; comp. L. Wittgenstein [1], 5.101.

[5] Or the *degree* of α.

$\neg$, $\wedge$, $\vee$, $\supset$, $\equiv$, is to completely specify the truth of the compound sentences containing it as a *function* of the truth values of its arguments (components).[6] This can be done by giving the corresponding *truth tables*, i.e.,

p	$\neg$p
1	0
0	1

;

p	q	p $\wedge$ q
1	1	1
1	0	0
0	1	0
0	0	0

;

p	q	p $\vee$ q
1	1	1
1	0	1
0	1	1
0	0	0

;

p	q	p $\supset$ q
1	1	1
1	0	0
0	1	1
0	0	1

;

p	q	p $\equiv$ q
1	1	1
1	0	0
0	1	0
0	0	1

These truth tables are, therefore, the *semantic definitions* of the respective connectives. A conjunction, p$\wedge$q, is true if and only if both p and q are true, a disjunction, p$\vee$q, is true if and only if at least one argument is true, an implication, p$\supset$q, is false if and only if the antecedent (p) is true and the consequent (q) is false, and an equivalence, p$\equiv$q, is true if and only if its arguments have the same truth value.

Of course, the above truth tables still hold if instead of p, q, we set arbitrary formulas of L_{PL}, α, β, since for any interpretation of their variables these formulas may be either true or false.

Using these connectives, as we saw in 1, we can construct more complex formulas of L_{PL}.

Examples. α=[(p$\supset$q)$\wedge\neg$r]$\equiv$q.[7]

Let us determine the truth value of α with respect to the truth values of its variables, i.e., we have to make the truth table for α. If *n* is the number of variables of a formula α, then the number of interpretations of p, q, r is 2^n, in our case 8.

p	q	r	p $\supset$ q	$\neg$r	(p $\supset$ q) $\wedge\neg$r	[(p $\supset$ q) $\wedge\neg$r)] $\equiv$ q
1	1	1	1	0	0	0
1	1	0	1	1	1	1
1	0	1	0	0	0	1
1	0	0	0	1	0	1
0	1	1	1	0	0	0
0	1	0	1	1	1	1
0	0	1	1	0	0	1
0	0	0	1	1	1	0

[6] I.e. all these operations are truth-functional operations.

[7] In the whole expression "=" is a symbol of metalanguage and means "is". Sometimes we use equivalently ":" instead of "=".

$$\beta = (p \supset q) \equiv (q \vee \neg p)$$

p	q	$p \supset q$	$\neg p$	$q \vee \neg p$	$(p \supset q) \equiv (q \vee \neg p)$
1	1	1	0	1	1
1	0	0	0	0	1
0	1	1	1	1	1
0	0	1	1	1	1

$$\gamma = (p \equiv q) \equiv [(p \wedge \neg q) \vee (q \wedge \neg p)]$$

p	q	$p \equiv q$	$\neg q$	$p \wedge \neg q$	$\neg p$	$q \wedge \neg p$	$(p \wedge \neg q) \vee (q \wedge \neg p)$	γ
1	1	1	0	0	0	0	0	0
1	0	0	1	1	0	0	1	0
0	1	0	0	0	1	1	1	0
0	0	1	1	0	1	0	0	0

As can be observed, for each assignment of truth values to its variables, each formula takes a truth value, 1 or 0. Hence each formula determines a *truth function*, representable by the truth table of the respective formula.

Definition. *Let* $\langle v_1,\ldots,v_n \rangle$, $n \geq 1$, *be an n-tuple of truth values.*[8] *Let* T *be the set of* 2^n *n-tuples. A truth function is a mapping from* T *to the set* $\{1,0\}$, *i.e.*,

$$f^n : T \rightarrow \{1,0\}.$$

Since the cardinal[9] of T is 2^n and the cardinal of the set $\{1,0\}$ is 2, it follows that there are $2^{(2^n)}$ different truth functions.

Let us consider the cases *n* = 1 and *n* = 2.

If *n* = 1 we have the following 4 functions.

p	$f_1^1(p)$	$f_2^1(p)$	$f_3^1(p)$	$f_4^1(p)$
1	1	1	0	0
0	1	0	1	0

If $f^1 = f_1^1$, then f^1 is the function *verum*, true for any assignment of its argument p. f_2^1 is just p, f_3^1 is $\neg p$ and f_4^1 is *falsum* (false for any

[8] Every member of an *n*-tuple being either 1 or 0.

[9] The cardinal number of a finite set is the number of its elements. The cardinal number of a denumerable set is $\aleph_0$.

assignment to p).[10]

If $n = 2$, then f^2 is one of the following 16 binary truth functions from Wittgenstein's Table[11]:

p	q	f_1(p,q)	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}	f_{16}
1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
1	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0
0	1	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0
0	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0
		$\mathcal{T}$	$\vee$	$\subset$	p	$\supset$	q	$\equiv$	$\wedge$	/	+	$\neg$q	$\not\supset$	$\neg$p	$\not\subset$	$\downarrow$	C

These 16 truth functions are respectively: tautology ($\mathcal{T}$), non-exclusive disjunction ($\vee$), converse implication ($\subset$), prependence (p), implication ($\supset$), postpendence (q), equivalence ($\equiv$), conjunction ($\wedge$), incompatibility (/), exclusive disjunction (+), nonpostpendence ($\neg$q), nonimplication ($\not\supset$), nonprependence ($\neg$p), nonconverse implication ($\not\subset$), rejection ($\downarrow$) and contradiction (C).

As can be seen, the functions $f_9 - f_{16}$ can be obtained, accordingly, by negating the truth values of the functions $f_8 - f_1$, a reason for which incompatibility is also named anticonjunction, exclusive disjunction – antiequivalence, rejection – antidisjunction, and so on.

2.1.2. Validity, satisfiability, unsatisfiability

What the truth tables of the above formulas show is the following fact: β is true for every interpretation of its variables, γ is false for every such interpretation, and α is true in some interpretations and false in other. Let us introduce definitionally these differences.

By $[\alpha]^{int}$ in what follows we understand the truth value of the formula α in an interpretation *int*. And if Γ is an *arbitrary* set of formulas of L_{PL}, then $[\Gamma]^{int} = 1$ means Sat(Γ), i.e., there is an interpretation *int* of the variables occurring in the formulas of Γ in which the formulas of Γ are *simultaneously* true.

---

[10] In a notation of the form f^n_m, *n* is the *arity* of f (i.e. the number of its arguments) and *m* is the *index* of f in an enumeration.

[11] Comp. L. Wittgenstein [1], 5.101.

Definition 1. *A formula* α *of* L_{PL} *is valid*[12] (*symbolic*: $\models \alpha$) *if and only if* $[\alpha]^{int} = 1$, *for every interpretation int of its variables*.

This means, equivalently, that its corresponding truth function takes only the value 1.

Definition 2. *A formula* α *of* L_{PL} *is satisfiable if and only if* $[\alpha]^{int} = 1$ *for some*[13] *interpretation int of its variables* (*symbolic*: Sat α).

Equivalently, this means that the corresponding truth function of α takes the value 1 for some *n*-tuple in T.

The following equivalences are immediate consequences of the semantic definitions of the corresponding connectives: $\neg$, $\wedge$, $\vee$, $\supset$, $\equiv$.

$\text{Sat}_{\neg}$ $\quad [\neg\alpha]^{int} = 1$ iff $[\alpha]^{int} = 0$

$\text{Sat}_{\wedge}$ $\quad [\alpha \wedge \beta]^{int} = 1$ iff $[\alpha]^{int} = 1$ and $[\beta]^{int} = 1$

$\text{Sat}_{\vee}$ $\quad [\alpha \vee \beta]^{int} = 1$ iff $[\alpha]^{int} = 1$ or $[\beta]^{int} = 1$

$\text{Sat}_{\supset}$ $\quad [\alpha \supset \beta]^{int} = 1$ iff $[\alpha]^{int} = 0$ or $[\beta]^{int} = 1$

$\text{Sat}_{\equiv}$ $\quad [\alpha \equiv \beta]^{int} = 1$ iff $[\alpha]^{int} = [\beta]^{int}$.

Definition 3. *A formula* α *of* L_{PL} *is unsatisfiable* (*symbolic* notSat α) *if and only if* $[\alpha]^{int} = 0$, *for every interpretation int of its variables*.[14]

Equivalently, the truth function it determines takes only the value 0.

Let us connect the meanings given by Def 1-3 by the theorems stated below. Some of them are immediate and do not require any proof.

Th 1. $\models \alpha$ *iff* notSat$\neg\alpha$.

Th 2. Sat α *iff* not$\models \neg\alpha$.

Th 3. *If* $\models \alpha \equiv \beta$, *then* ($\models \alpha$ iff $\models \beta$).

The converse of the theorem does not hold, since, for example, $\models$ p iff $\models$ q holds but not$\models$ p $\equiv$ q.

Th 4. *If* $\models \alpha \equiv \beta$, *then* (Sat α iff Sat β).

The converse does not hold; as above.

Th 5. $\models \alpha \wedge \beta$ *iff* ($\models \alpha$ *and* $\models \beta$); *by def of* $\wedge$.

Th 6. *If* $\models \alpha$ *or* $\models \beta$, *then* $\models \alpha \vee \beta$; *by def of* $\vee$.

---

[12] Also called *tautology*.

[13] Possible for all interpretations. In a more restricted sense, a formula α is satisfiable iff α is neither valid, nor unsatisfiable.

[14] From now on we shall often time omit the qualifying phrases "of its variables" or "of the variables occurring in a formula α".

The converse does not hold; $\models(p\supset q)\vee(p\vee q)$, but not$\models p\supset q$ and not$\models p\vee q$.

Th 7. *If* Sat $\alpha\wedge\beta$, *then* Sat α *and* Sat β.

Proof. Assume Sat $\alpha\wedge\beta$, i.e., for some interpretation *int* of variables in α and β, $[\alpha\wedge\beta]^{\text{int}}=1$, by Def 2 above. And then $[\alpha]^{\text{int}}=1$ and $[\beta]^{\text{int}}=1$, by $\text{Sat}_{\wedge}$. Whence Sat α and Sat β, by Def 2.

The converse does not hold, since, for example, Sat p and Sat $\neg$p, however notSat $p\wedge\neg p$.

Th 8. Sat $\alpha\vee\beta$ *iff* Sat α *or* Sat β.

Th 9. (*If* Sat α, *then* Sat β), *then* Sat $\alpha\supset\beta$.

Equivalent: If notSat $\alpha\supset\beta$, then Sat α and notSat β.

Proof. Assume notSat $\alpha\supset\beta$. Then $\models\neg(\alpha\supset\beta)$, by Th1, equivalent $\models\alpha\wedge\neg\beta$, (show that!) and then $\models\alpha$ and $\models\neg\beta$, by Th5. Whence, for every *int* $[\alpha]^{\text{int}}=1$ and $[\neg\beta]^{\text{int}}=1$, hence $[\beta]^{\text{int}}=0$. It follows then that Sat α and notSat β.

The converse of this theorem does not hold, since, for example, Sat $p\supset(q\equiv\neg q)$ but the following does not hold: if Sat p, then Sat $q\equiv\neg q$.

Th 10. *If* $\models\alpha\supset\beta$, *then* (*if* Sat α, *then* Sat β).

Equivalent: If Sat α and notSat β, then not$\models\alpha\supset\beta$.

Proof. Assume Sat α, i.e., for some *int* $[\alpha]^{\text{int}}=1$, and nonSat β, i.e., for every *int* $[\beta]^{\text{int}}=0$. It follows that there is an *int* such that $[\alpha\supset\beta]^{\text{int}}=0$, and then not$\models\alpha\supset\beta$.

The converse of this theorem does not hold, since Sat q holds, and then also holds: if Sat p, then Sat q, but not$\models p\supset q$.

Th 11. *If* $\models\beta$, *then* $\models(\alpha\wedge\beta)\equiv\alpha$,

i.e., the truth value of a conjunction does not change if a valid (or true) argument is eliminated (cf. 2.2,9)[15] (below). And then an empty conjunction is valid.

[15] Where the notations of the forms "Sect. 2.2,9" or, simply, "2.2,9", or "Ch. 2, Sect. 3.2.6", etc. indicate the respective section and the corresponding item, including the chapter (if this is the case). Here, 2.2,9 means "the section 2.2, the formula 9".

Th 12. *If* $\models\beta$, *then* $\models(\alpha\vee\beta)\equiv\beta$,

i.e., a disjunction is valid if at least one argument is valid.

Th 13. *If* $\models\beta$, *then* $\models(\beta\supset\alpha)\equiv\alpha$.

Th 14. *If* $\models\beta$, *then* $\models(\beta\equiv\alpha)\equiv\alpha$.

Modus Ponens (MP). If $\models\alpha$ and $\models\alpha\supset\beta$, then $\models\beta$.

Proof (*reductio*). Assume $\models\alpha$, $\models\alpha\supset\beta$ and not$\models\beta$. Then $[\beta]^{int}=0$, for some *int* and then $[\alpha\supset\beta]^{int}=0$ for some *int*, since $[\alpha]^{int}=1$ for every *int*. Whence not$\models\alpha\supset\beta$, contrary to the assumption.

The converse of MP does not hold, since if $\models\beta$, then for *any* formula α of L_{PL} $\models\alpha\supset\beta$ (argue). Hence from the fact that $\models\beta$ does not follow that $\models\alpha\supset\beta$ *and* $\models\alpha$.

MP (variant). If $\models\alpha\supset\beta$, then (if $\models\alpha$, then $\models\beta$).

The two formulations of MP are equivalent; (comp. 2.2,41) (below).

2.2. Remarkable formulas of L_{PL}

Let us consider some remarkable[16] valid formulas of L_{PL}, with respect to the connectives mentioned in the front of every class ($\neg,\wedge,\vee,\supset,\equiv$).

($\neg$) 1. $\neg...\neg^{2k}p\equiv p$; $k=0,1,2,...$; 1 and 2: elimination of multiple negations

2. $\neg...\neg^{2k+1}p\equiv\neg p$; $k=0,1,2,...$

($\wedge$) 3. $(p\wedge p)\equiv p$;

4. $(p\wedge q)\equiv(q\wedge p)$; commutativity of $\wedge$

5. $[p\wedge(q\wedge r)]\equiv[(p\wedge q)\wedge r]$; associativity of $\wedge$

6. $[p\wedge(q\vee r)]\equiv[(p\wedge q)\vee(p\wedge r)]$; distributivity of $\wedge$ with respect to $\vee$

7. $[p\wedge(p\vee q)]\equiv p$;

8. $[p\wedge(p\wedge q)]\equiv(p\wedge q)$

9. $[p\wedge(q\vee\neg q)]\equiv p$

10. $(p\wedge q)\supset p$;

11. $(p\wedge q)\supset q$

[16] "Remarkable" in the sense of being frequently used in our proofs in this book.

12. $\bigwedge_{i=1}^{n} p_i \supset p_j$; $1 \le j \le n$; generalization of 10 and 11
13. $(p \wedge q) \supset (p \vee q)$
14. $(p \wedge q) \supset (p \supset q)$
15. $(p \wedge q) \equiv \neg(\neg p \vee \neg q)$; 15 and 16 De Morgan laws
16. $\neg(p \wedge q) \equiv (\neg p \vee \neg q)$
17. $\neg(p \wedge \neg p)$; *principle of noncontradiction*
18. $(p \wedge \neg p) \supset q$; *ex falso quodlibet*

($\vee$) 19. $(p \vee p) \equiv p$;
20. $(p \vee q) \equiv (q \vee p)$; commutativity of $\vee$
21. $[p \vee (q \vee r)] \equiv [(p \vee q) \vee r]$; associativity of $\vee$
22. $[p \vee (q \wedge r)] \equiv [(p \vee q) \wedge (p \vee r)]$; distributivity of $\vee$ with respect to $\wedge$
23. $[p \vee (q \supset r)] \equiv [(p \vee q) \supset (p \vee r)]$; distributivity of $\vee$ with respect to $\supset$
24. $[p \vee (q \equiv r)] \equiv [(p \vee q) \equiv (p \vee r)]$; distributivity of $\vee$ with respect to $\equiv$
25. $[p \vee (p \vee q)] \equiv (p \vee q)$;
26. $[p \vee (p \wedge q)] \equiv p$
27. $[p \vee (q \wedge \neg q)] \equiv p$
28. $p \supset (p \vee q)$; 28 and 29: introduction of $\vee$
29. $q \supset (p \vee q)$
30. $p_j \supset \bigvee_{i=1}^{n} p_i$; $1 \le j \le n$ (generalization of 28 and 29)
31. $(p \vee q) \equiv \neg(\neg p \wedge \neg q)$; 31 and 32 De Morgan laws
32. $\neg(p \vee q) \equiv (\neg p \wedge \neg q)$
33. $p \vee \neg p$; *tertium non datur*
34. $(p \supset q) \vee (p \supset \neg q)$
35. $(p \supset q) \vee (\neg p \supset q)$
36. $(p \equiv q) \vee (p \equiv \neg q)$
37. $(p \equiv q) \vee (\neg p \equiv q)$
38. $[(p \vee \neg p) \supset q] \equiv q$

$(\supset)$ 39. $p \supset p$; reflexivity of $\supset$

40. $[(p \supset q) \wedge (q \supset r)] \supset (p \supset r)$; transitivity of $\supset$

41. $[p \supset (q \supset r)] \equiv [(p \wedge q) \supset r]$;

42. $[p \supset (q \supset r)] \equiv [q \supset (p \supset r)]$; permutation of premises

43. $[p \supset (p \supset q)] \equiv (p \supset q)$;

44. $[p \supset (q \wedge r)] \equiv [(p \supset q) \wedge (p \supset r)]$; distributivity of $\supset$ with respect to $\wedge$

45. $\bigwedge_{i=1}^{n}(q \supset p_i) \equiv \left(q \supset \bigwedge_{i=1}^{n} p_i\right)$;

46. $\bigwedge_{i=1}^{n}(p_i \supset q) \equiv \left(\bigvee_{i=1}^{n} p_i\right) \supset q$

47. $[p \supset (q \supset r)] \equiv [(p \supset q) \supset (p \supset r)]$; self-distributivity of $\supset$

48. $(p \supset q) \equiv (\neg q \supset \neg p)$; 48 – 50 laws of contraposition

49. $(\neg p \supset q) \equiv (\neg q \supset p)$

50. $(p \supset \neg q) \equiv (q \supset \neg p)$

51. $[(p \wedge q) \supset r] \equiv [(p \wedge \neg r) \supset \neg q] \equiv [(q \wedge \neg r) \supset \neg p]$; partial contrapositions

52. $p \supset (q \supset p)$; 52 and 53 paradoxes of material implication

53. $\neg p \supset (p \supset q)$

54. $p \supset (\neg p \supset q)$; *ex falso quodlibet*

55. $(p \supset q) \supset [(p \wedge r) \supset q]$;

56. $(p_j \supset q) \supset \left(\left(\bigwedge_{i=1}^{n} p_i\right) \supset q\right)$, $1 \le j \le n$

57. $\left(\left(\bigvee_{i=1}^{n} p_i\right) \supset q\right) \supset (p_j \supset q)$, $1 \le j \le n$

58. $(p \supset q) \supset [(r \supset p) \supset (r \supset q)]$

59. $(p \supset q) \supset [(q \supset r) \supset (p \supset r)]$

60. $(p \supset q) \supset [(p \vee r) \supset (q \vee r)]$

61. $[p \wedge (p \supset q)] \supset q$; *modus ponens*

62. $p \supset [(p \supset q) \supset q]$; *modus ponendo ponens*

63. $p \supset [(q \supset \neg p) \supset \neg q]$; 63 şi 64 *modus tollendo tollens*

64. $[(p \supset q) \wedge \neg q] \supset \neg p$ *modus tollens*

65. $[(p \supset q) \wedge (\neg p \supset q)] \supset q$; *constructive dilemma*
66. $(p \supset \neg p) \supset \neg p$; 66 – 67 *reductio ad absurdum*
67. $(\neg p \supset p) \supset p$
68. $[(p \supset q) \wedge (p \supset \neg q)] \supset \neg p$
69. $[p \supset (q \wedge \neg q)] \equiv \neg p$
70. $[(p \supset q) \supset p] \supset p$; Peirce's law

(≡) 71. $p \equiv p$; reflexivity of ≡
72. $(p \equiv q) \supset (q \equiv p)$; symmetry of ≡
73. $[(p \equiv q) \wedge (q \equiv r)] \supset (p \equiv r)$; transitivity of ≡
74. $[p \equiv (q \equiv r)] \equiv [(p \equiv q) \equiv r]$; associativity of ≡
75. $(p \equiv q) \equiv (\neg p \equiv \neg q)$; 75 and 76: contraposition of ≡
76. $(\neg p \equiv q) \equiv (p \equiv \neg q)$
77. $(p \equiv q) \supset (p \supset q)$;
78. $(p \equiv q) \supset (q \supset p)$
79. $\neg(p \equiv q) \equiv (\neg p \equiv q)$; rejection of ≡
80. $(p \equiv q) \equiv [(p \wedge q) \vee (\neg p \wedge \neg q)]$; truth conditions of ≡
81. $(p \equiv q) \supset [(p \circ r) \equiv (q \circ r)]$, where „∘" denotes everyone of the connectives ∧, ∨, ⊃, ≡
82. $[(p \equiv q) \wedge (r \equiv s)] \supset [(p \circ r) \equiv (q \circ s)]$, with „∘" as mentioned
83. $[(p \equiv \neg p) \equiv p] \equiv \neg p$
84. $(p \supset q) \equiv [p \equiv (p \wedge q)]$
85. $(p \supset q) \equiv [q \equiv (p \vee q)]$
86. $[p \supset (q \equiv r)] \equiv [(p \wedge q) \equiv (p \wedge r)]$
87. $[p \supset (q \vee r)] \equiv [(p \supset q) \vee r]$
88. $[p \supset (q \vee r)] \equiv [(p \supset r) \vee q]$
89. $\neg(p \equiv \neg p)$; noncontradiction
90. $[p \equiv (\neg p \wedge q)] \supset \neg q$; noncontradiction
91. $(p \equiv \neg p) \supset q$
92. $[p \equiv (q \wedge \neg q)] \equiv \neg p$
93. $[(p \supset q) \wedge (p \supset \neg q)] \equiv \neg p$
94. $[(p \supset q) \wedge (\neg p \supset q)] \equiv q$
95. $[p \supset (q \equiv \neg q)] \equiv \neg p$.

Remark. If in any valid formula in this list we replace the variables p, q, r with the names of arbitrary formulas of L_{PL}, then we'll obtain the respective *schemas* of valid formulas of L_{PL}. From the formula $(p\supset q)\supset(\neg q\supset\neg p)$, for example, we get the valid schema $(\alpha\supset\beta)\supset(\neg\beta\supset\neg\alpha)$. As can be seen, such a schema represents an infinity of valid formulas, obtained by replacing the names α and β with *formulas* of L_{PL}.

2.3. Reducibility in PL

Let us take some examples illustrating the idea of reducibility.

If $\alpha = p \equiv q$, then α can be equivalently expressed by the formula $\alpha^* = \neg[\neg(\neg p\vee q)\vee\neg(\neg q\vee p)]$, in which only $\neg$ and $\vee$ occur (show that!). If $\beta = p + q$, then β can be written, equivalently, as $\beta^* = \neg(\neg p\vee q)\vee\neg(\neg q\vee p)$, i.e., it can be expressed by a formula containing only the connectives $\neg$ and $\vee$. If $\gamma = (p \supset q)$, then γ^* will be $\neg p\vee q$, which also contains only $\neg$ and $\vee$. All these cases are examples of *reducibility*. Let us introduce this notion definitionally.

Definition 1. *Let* $M = \{con_1,...,con_n\}$, $n \geq 1$, *be a set of connectives. A formula* α *of* L_{PL} *is called* M-*reducible if and only if* α *can be equivalently expressed by a formula* α^* *containing only the connectives of* M.

Definition 2. *Let* M *be a set of connectives*. M *is called complete if and only if any formula of* L_{PL} *is* M-*reducible*.

Theorem. *Each of the following sets is complete*: $M_1 = \{\neg,\wedge\}$, $M_2 = \{\neg,\vee\}$ *and* $M_3 = \{\neg,\supset\}$.

Proof. Using the truth tables *it can be shown* that the equivalences below hold (exercise).

$M_1 = \{\neg,\wedge\}$; with respect to the connectives $\vee$, $\supset$, $\equiv$, $+$, $/$, $\downarrow$

$(p\vee q) \equiv \neg(\neg p\wedge\neg q)$

$(p\supset q) \equiv \neg(p\wedge\neg q)$

$(p \equiv q) \equiv [\neg(p\wedge\neg q)\wedge\neg(q\wedge\neg p)]$

$(p + q) \equiv \neg[\neg(p\wedge\neg q)\wedge\neg(q\wedge\neg p)]$

$(p/q) \equiv \neg(p\wedge q)$

$(p\downarrow q) \equiv (\neg p\wedge\neg q)$

$M_2 = \{\neg,\vee\}$; with respect to the connectives $\wedge$, $\supset$, $\equiv$, $+$, $/$, $\downarrow$

$(p\wedge q) \equiv \neg(\neg p\vee\neg q)$

$(p\supset q) \equiv (\neg p\vee q)$

$(p \equiv q) \equiv \neg[\neg(\neg p\vee q)\vee\neg(\neg q\vee p)]$

$(p + q) \equiv [\neg(\neg p\vee q)\vee\neg(\neg q\vee p)]$

$(p/q) \equiv (\neg p\vee\neg q)$

$(p\downarrow q) \equiv \neg(p\vee q)$

$M_3 = \{\neg,\supset\}$; with respect to the connectives $\wedge$, $\vee$, $\equiv$, $+$, $/$, $\downarrow$

$(p\wedge q) \equiv \neg(p\supset\neg q)$

$(p\vee q) \equiv (\neg p\supset q)$

$(p \equiv q) \equiv \neg[(p\supset q)\supset\neg(q\supset p)]$

$(p + q) \equiv [(p\supset q)\supset\neg(q\supset p)]$

$(p/q) \equiv (p\supset\neg q)$

$(p\downarrow q) \equiv \neg(\neg p\supset q)$

Exercise. Show that $M_4 = \{/\}$ and $M = \{\downarrow\}$ are the only complete sets containing just one connective.

2.4. Substitution and replacement in PL

2.4.1. Substitution in PL

Let us illustrate the idea of substitution in PL. Let $\alpha = (p\supset q)\supset(\neg q\supset\neg p)$. If in this formula instead of p, in all of its occurrences in α, we put, for example, $p\wedge\neg q$, and instead of q, again in all of its occurrences, we put, for example, $\neg p$, then the formula obtained α^* is $((p\wedge\neg q)\supset\neg p)\supset(\neg\neg p\supset\neg(p\wedge\neg q))$. But α is valid (show that!), and we can easy show that α^* is also valid (show that!). This is the content of an important theorem in PL, i.e.,

Substitution Theorem. (Subst_{PL}) *Let* $\alpha(p_1,\ldots,p_n)$ *be a formula of* L_{PL} *containing the variables* $p_1,\ldots,p_n$. *Let* $\alpha^*(\beta_1/p_1,\ldots,\beta_n/p_n)$ *be the formula obtained from* α *by substituting arbitrary formulas of* L_{PL}, $\beta_1,\ldots,\beta_n$, *for* $p_1,\ldots,p_n$, *respectively. Then the following holds*:

If $\models\alpha$, *then* $\models\alpha^*$.

Proof. Assume $\models\alpha(p_1,\ldots,p_n)$. Let *int* be an arbitrary interpretation of variables of α^*, and let $v_1,\ldots,v_n$ be the truth values taken by $\beta_1,\ldots,\beta_n$ in *int*. Now, if we assign these values $v_1,\ldots,v_n$ to the variables $p_1,\ldots,p_n$, respectively, then the truth value of α will be exactly the truth value of α^* in *int*. Since, by hypothesis, α is valid, it follows that $[\alpha]^{int} = 1$, and then $[\alpha^*]^{int} = 1$. But *int* was arbitrary, hence α^* is also valid.

Remarks. 1. The converse of Substitution Theorem does not hold generally. Since a *valid* formula can also be obtained by substitution from a non-valid formula; e.g., $\alpha^* = (p \equiv q) \supset (p \supset q)$ is valid and can result by substitution from the non-valid formula $p \supset q$!

2. Though the idea of substitution contains the idea of a *simultaneous* substitution, since $\beta_1, \ldots, \beta_n$ are *arbitrary* formulas of L_{PL}, nothing prevent us from choosing for substitution only some of the variables $p_1, \ldots, p_n$ (even one), the other remaining unchanged.

3. By Substitution Theorem from a valid formula α of L_{PL} we obtain an infinite number of valid formulas of L_{PL}, whose validity is guaranteed just by this theorem.

2.4.2. Replacement in PL

Let us firstly illustrate the content of this theorem. Let $\alpha_\beta = (p \supset q) \supset (q \equiv r)$, containing the subformula $\beta = p \supset q$ (a fact hinted by the notation α_β). Let $\gamma = \neg p \vee q$ and $\alpha_\gamma = (\neg p \vee q) \supset (q \equiv r)$. As can be observed, $\models (p \supset q) \equiv (\neg p \vee q)$, viz. $\models \beta \equiv \gamma$. Then we have the following result: $\models \alpha_\beta = \alpha_\gamma$ (show that!)

Replacement theorem ($Repl_{PL}$). *Let* α_β *be a formula of* L_{PL} *containing the subformula* (*proper or not*) β, *and let* α_γ *be the formula resulting from* α_β *by replacing one or more occurrences of* β *with* γ. *Then the following holds:*

If $\models \beta \equiv \gamma$, *then* $\models \alpha_\beta \equiv \alpha_\gamma$.

Proof (induction on *n* = the complexity of α).

Basis. $n = 0$. In this case $\alpha_\beta = \beta = p$, and evidently the theorem holds, since if $\models \beta \equiv \gamma$, then $\models \beta \equiv \gamma$.

Induction. Suppose that compl(α) = $n \neq 0$ and that the theorem holds for any $k < n$. And show that it holds for n.

Since $M = \{\neg, \supset\}$ is complete (by 2.3) is enough to consider only the following forms of α_β:

(a) $\alpha_\beta = \neg\delta_\beta$

(b) $\alpha_\beta = (\delta \supset \varepsilon)_\beta$ (i.e., $\delta_\beta \supset \varepsilon_\beta$).

(a) $\alpha_\beta = \neg\delta_\beta$, where compl($\beta_\delta$) < n, and then the theorem holds for δ_β. We have the following derivations:

(1) $\models \beta \equiv \gamma$; hyp.

(2) $\models \delta_\beta \equiv \delta_\gamma$; (1) by ind. hyp.

(3) $\models (\delta_\beta \equiv \delta_\gamma) \supset (\neg\delta_\beta \equiv \neg\delta_\gamma)$; by Subst$_{PL}$.[17]

(4) $\models \neg\delta_\beta \equiv \neg\delta_\gamma$; (2), (3), MP; i.e., $\models \alpha_\beta \equiv \alpha_\gamma$.

(b) $\alpha_\beta = \delta_\beta \supset \varepsilon_\beta$, where compl(δ_β) < n and compl(ε_β) < n, and then the theorem holds for δ_β and δ_γ. So,

(1) $\models \beta \equiv \gamma$; hyp.

(2) $\models \delta_\beta \equiv \delta_\gamma$; by ind. hyp.

(3) $\models \varepsilon_\beta \equiv \varepsilon_\gamma$; by ind. hyp.

(4) $\models (\delta_\beta \equiv \delta_\gamma) \supset [(\varepsilon_\beta \equiv \varepsilon_\gamma) \supset ((\delta_\beta \supset \varepsilon_\beta) \supset (\delta_\gamma \supset \varepsilon_\gamma))]$; PL

(5) $\models (\delta_\beta \supset \varepsilon_\beta) \supset (\delta_\gamma \supset \varepsilon_\gamma)$; (2), (3), (4), MP (twice)

i.e., $\models \alpha_\beta \equiv \alpha_\gamma$.

Comments. 1. In the proof of (b) we have considered the case when the subformula β occurs in both formulas, δ and ε. Actually, we also have the following two cases:

(a) β occurs only in antecendent of α_β (i.e., in δ). In this case α_β is the formula $\delta_\beta \supset \varepsilon$ (as in our example above).

(b) β occurs only in the consequent of α_β (i.e., in ε). In this case α_β is the formula $\delta \supset \varepsilon_\beta$.

In both cases the theorem holds (argue!).

2. In our example above we considered the case when only one occurrence of β in α_β was replaced by γ. But α_β may contain more occurrences of β, and then we repeat the procedure. In fact, the number of occurrences of β in α_β in which β is replaced with γ is arbitrary.

[17] Often the simple cases of some theorems will be justified using the label "by PL", or simply "PL".

Corollary (Replacemant Rule). *If* $\models \beta \equiv \gamma$ *and* $\models \alpha_\beta$, *then* $\models \alpha_\gamma$ (argue!).

Remark. Sometimes Replacement Theorem is given in the following form: Let α_β and α_γ be as above, where this time γ is an *arbitrary* formula of L_{PL} (i.e., where the condition $\models \beta \equiv \gamma$ may or may not hold!). Then $\models (\beta \equiv \gamma) \supset (\alpha_\beta \equiv \alpha_\gamma)$ (construct an argument!). Evidently, from this form, the preceding form of the Replacement Theorem can be derived.

The two theorems, $Subst_{PL}$ and $Repl_{PL}$, combined, may be used in proving the validity of formulas of L_{PL}. Let us take an example.

Let $\alpha = (p \equiv q) \supset (p / \neg q)$. We must show that α is valid using both theorems. Now, a simple application of the truth table method shows us that $\models (p/\neg q) \equiv (p \supset q)$ and $\models (p \equiv q) \supset ((p \supset q) \wedge (q \supset p))$. Hence using $Repl_{PL}$ from α we derive, *equivalently*, $\alpha^* = ((p \supset q) \wedge (q \supset p)) \supset (p \supset q)$. But $\models \alpha^*$, since it can be obtained, using $Subst_{PL}$, from the evidently valid formula of L_{PL}: $(p \wedge q) \supset p$. And then α is also valid.

Exercises. Using $Subst_{PL}$ and $Repl_{PL}$, argue the validity of the following formulas:

α_1: $(p \equiv q) \supset (\neg p \vee q)$

α_2: $(p/q) \supset (r \vee \neg(p \wedge q))$

α_3: $(p/q) \supset [q \supset ((p/q) \wedge \neg\neg\neg q)]$

α_4: $(p \equiv q) \supset (r \supset \neg(p + q))$.

2.5. Duality in PL

As we saw (comp. 2.3), the sets $M_1 = \{\neg, \wedge\}$ and $M_2 = \{\neg, \vee\}$ are *complete* sets of connectives. So, *a fortiori*, the set $M = \{\neg, \wedge, \vee\}$ [18] is also complete. This means that any formula α of L_{PL} can be converted *equivalently*[19] into a formula α^* which contains only connectives of M.

Notational convention. In what follows "eq" is a metalinguistic operator whose meaning is: **α eq β iff $\models \alpha \equiv \beta$** (i.e., "eq" says that α and β are *logically* equivalent). So we'll use these notations interchangeably.

---

[18] These connectives, $\neg$, $\wedge$, $\vee$, are called *Boolean connectives* (*operators*). As we'll see below (2.7.4), they play a key role in the construction of the so called *Boolean normal forms*.

[19] "Equivalently", since all these changes are in fact applications of $Repl_{PL}$.

By using $M = \{\neg, \wedge, \vee\}$, for example, the formula $(p \equiv q) \wedge (p / r)$ can be transformed in $(p \supset q) \wedge (q \supset p) \wedge \neg(p \wedge r)$, and finally in α: $(\neg p \vee q) \wedge (\neg q \vee p) \wedge (\neg p \vee \neg r)$ (containing only the Boolean connectives).

Definition. *Let* α *be a formula of* L_{PL} *which contain only Boolean connectives. Then the dual* α^{δ} *of* α *is the formula obtained from* α *by interchanging the binary*[20] *connectives* $\wedge$ *and* $\vee$ *in all of their occurrences in* α.

Example. The dual of the formula above mentioned α is

$$\alpha^{\delta} : (\neg p \wedge q) \vee (\neg q \wedge p) \vee (\neg p \wedge \neg r).$$

Remarks. By the above definition of the duality the following evidently hold:

(1) $\models \alpha^{\delta\delta} \equiv \alpha$

(2) $\models (\alpha \wedge \beta)^{\delta} \equiv (\alpha^{\delta} \vee \beta^{\delta})$

(3) $\models (\alpha \vee \beta)^{\delta} \equiv (\alpha^{\delta} \wedge \beta^{\delta})$

(4) $\models \neg(\alpha^{\delta}) \equiv (\neg\alpha)^{\delta}$.

(Argue that!)

From this definition it also follows that the truth table (matrix) of α^{δ} can be obtained from the truth table of α by interchanging the truth values 1 and 0 (throughout in α), and conversely (show that by an example!).

Theorem. *Let* $\alpha(p_1,...,p_n)$ *be a formula of* L_{PL} *containing only the Boolean connectives and the variables* $p_1,...,p_n$. *Let* α^{*} *be the formula of* L_{PL} *arising from* α *by interchanging* $\wedge$ *and* $\vee$ *and by substituting each variable with its negation. Then* $\models \neg\alpha \equiv \alpha^{*}$; *equivalently*, $\models \alpha \equiv \neg\alpha^{*}$.

What this theorem says is the following thing: the negation $\neg\alpha$ of a formula α (of the type specified) can be obtained by, firstly, constructing its dual, α^{δ}, and then substitute every variable with its negation (and finally eliminate the multiple negation symbols). Let $\alpha^{\delta} + \mathrm{Subst}(\neg p_i / p_i)$

[20] Evidently, "binary" can be omitted from the definition, since by (4) below the negation is *self*-dual! Moreover, this definition is not restricted to the pair $(\wedge, \vee)$ of dual connectives. It also holds if instead we consider the following pairs of binary connectives: $(\equiv, +)$, $(/, \downarrow)$, $(\supset, \not\subset)$, $(\subset, \not\supset)$; for details, comp. A. Church [1956], §§16, 37.

symbolize these two operations: "construct α^δ *and* then substitute $\neg p_i$ for p_i". Then α^* of the theorem is the formula $\alpha^\delta + \text{Subst}(\neg p_i / p_i)$.

Example. Let α: $(\neg p \vee q) \wedge (\neg q \vee p) \wedge (\neg p \vee \neg r)$. Then
$\neg\alpha$: $\neg[(\neg p \vee q) \wedge (\neg q \vee p) \wedge (\neg p \vee \neg r)]$ eq
$\neg(\neg p \vee q) \vee \neg(\neg q \vee p) \vee \neg(\neg p \vee \neg r)$ eq $(p \wedge \neg q) \vee (q \wedge \neg p) \vee (p \wedge r)$.
Now, α^*: $(\neg\neg p \wedge \neg q) \vee (\neg\neg q \wedge \neg p) \vee (\neg\neg p \wedge \neg\neg r)$ eq
$(p \wedge \neg q) \vee (q \wedge \neg p) \vee (p \wedge r)$ eq $\neg\alpha$.

Proof of the theorem (induction on n = complexity of α).

Basis. $n = 0$. Then α is a variable p, $\alpha^\delta = p$ and $\alpha^* = \neg p$, and then the theorem holds.

Induction. Compl(α) = n, and suppose that the theorem holds for any α with compl(α) < n. We have, accordingly, the following three cases:

(a) $\alpha = \neg\beta$, and the theorem holds for β. Then $\models \neg\beta \equiv \beta^*$, and therefore $\models \neg\neg\beta \equiv \neg\beta^*$ (by PL). But $\neg\beta^*$ is just α^*.

(b) $\alpha = \beta \wedge \gamma$. Then $\neg\alpha$ is $\neg(\beta \wedge \gamma)$, equivalently $\neg\beta \vee \neg\gamma$. Since the theorem holds for β and γ (by ind. hyp.), it follows that we have: $\models \neg\beta \equiv \beta^*$ and $\models \neg\gamma \equiv \gamma^*$. This does imply that $\models (\neg\beta \vee \neg\gamma) \equiv (\beta^* \vee \gamma^*)$ (by Repl$_{PL}$), equivalently $\models \neg\alpha \equiv (\beta^* \vee \gamma^*) \equiv (\beta \wedge \gamma)^* \equiv \alpha^*$.

(c) $\alpha = \beta \vee \gamma$. Then $\neg\alpha$ is the formula $\neg(\beta \vee \gamma)$, equivalently $\neg\beta \wedge \neg\gamma$. As above, we have the following derivations: $\models \neg\beta \equiv \beta^*$ and $\models \neg\gamma \equiv \gamma^*$, and then $\models \neg\alpha \equiv (\beta^* \wedge \gamma^*) \equiv (\beta \vee \gamma)^* \equiv \alpha^*$.

Remark. Intuitively, the proof of this theorem proceeds as follows:

(1) From $\neg\alpha$ and using the De Morgan formulas (cf. 2.2, 16, 32) we gradually remove the negation from the front of a formula by negating its arguments and by interchanging $\wedge$ and $\vee$.

(2) Using 2.2, 1, 2) we eliminate the multiple negations, such that any variable occurs negated at most once. The formula so obtained is just α^*.

Example. $\alpha = [(\neg p \wedge q) \vee \neg\neg r] \wedge \neg q$
Then $\alpha^* = [(p \vee \neg q) \wedge \neg r] \vee q$, since
(1) $\neg\alpha = \neg\{[(\neg p \wedge q) \vee \neg\neg r] \wedge \neg q\} \equiv \{\neg[(\neg p \wedge q) \vee \neg\neg r] \vee \neg\neg q\}$
$\equiv \{[\neg(\neg p \wedge q) \wedge \neg\neg\neg r] \vee \neg\neg q\} \equiv \{[(\neg\neg p \vee \neg q) \wedge \neg\neg\neg r] \vee \neg\neg q\}$

(2) $[(p \vee \neg q) \wedge \neg r] \vee q = \alpha^*$

By the preceding theorem $\models \neg\alpha \equiv \alpha^*$ (show that using truth table).

Duality Theorem. *Let* α *be a formula of* L_{PL} *containing only Boolean connectives, let* α^δ *be its dual. Then the following holds*:

1. $\models \alpha$ *iff* $\models \neg\alpha^\delta$.
2. $\models \alpha \supset \beta$ *iff* $\models \beta^\delta \supset \alpha^\delta$.
3. $\models \alpha \equiv \beta$ *iff* $\models \alpha^\delta \equiv \beta^\delta$.

Proof. 1a) If $\models \alpha$, then $\models \neg\alpha^\delta$.

(1) Assume $\models \alpha$. Then $\models \neg\alpha^*$; by Theorem and $Repl_{PL}$ (Corollary).

(2) If $\models \neg\alpha^*$, then $\models \neg\alpha^{**}$, where α^{**} is obtained from α^* by substituting $\neg p_i$ for each p_i of α (by $Subst_{PL}$).

(3) If $\models \neg\alpha^{**}$, then $\models \neg\alpha^\delta$, where $\neg\alpha^\delta$ is obtained from $\neg\alpha^{**}$ by eliminating multiple negations.

(4) If $\models \alpha$, then $\models \neg\alpha^\delta$; (1)-(3); PL.

1b) If $\models \neg\alpha^\delta$, then $\models \alpha$ (by 1a) and Remarks (1) and (4)).

2. $\models \alpha \supset \beta$, iff $\models \beta^\delta \supset \alpha^\delta$.

We have the following derivation:

$\models \alpha \supset \beta$ iff $\models \neg\alpha \vee \beta$ iff $\models \neg(\neg\alpha \vee \beta)^\delta$ (by 1) iff

iff $\models \neg(\neg\alpha^\delta \wedge \beta^\delta)$ (by Remarks (3) above) iff $\models \neg\neg\alpha^\delta \vee \neg\beta^\delta$

iff $\models \alpha^\delta \vee \neg\beta^\delta$ iff $\beta^\delta \supset \alpha^\delta$ (by PL).

3. $\models \alpha \equiv \beta$ iff $\models \alpha^\delta \equiv \beta^\delta$

We have: $\models \alpha \equiv \beta$ iff $\models (\alpha \supset \beta) \wedge (\beta \supset \alpha)$ (by PL) iff

$\models \alpha \supset \beta$ and $\models \beta \supset \alpha$ (by 2.1.2, Th.5) iff $\models \beta^\delta \supset \alpha^\delta$ and

$\models \alpha^\delta \supset \beta^\delta$ (by 2) iff $\models (\beta^\delta \supset \alpha^\delta) \wedge (\alpha^\delta \supset \beta^\delta)$ (by 2.1.2, Th.5)

iff $\models \alpha^\delta \equiv \beta^\delta$ (by PL).

2.6. Semantic consequence in PL

Let us consider the following formulas of L_{PL}: $\alpha = p \wedge q$ and $\beta = p \supset q$. By comparing their truth tables we observe that whenever α is true β is also true (the first line of the truth table). In this case we say that β is a semantical consequence of α; symbolically: $\alpha \models \beta$. The same is true if we take $\alpha_1 = p \wedge q$, $\alpha_2 = p \equiv q$ and $\beta = p \supset q$: whenever α_1 and α_2 are *simultaneously* true, the formula β is also true. Similarly, we say that β is a semantical consequence of α_1 and α_2; symbolically: $\alpha_1, \alpha_2 \models \beta$.

Definition. *Let* $\alpha_1,...,\alpha_n,\beta$ *be formulas of* L_{PL}. β *is a semantical consequence of the formulas* $\alpha_1,...,\alpha_n$ (*symbolically*: $\alpha_1,...,\alpha_n \models \beta$) *if for any interpretation of variables occurring in* $\alpha_1,...,\alpha_n,\beta$, *the formula* β *is true whenever* $\alpha_1,...,\alpha_n$ *are simultaneously true* ($\alpha_1,...,\alpha_n$ *are the premises and* β *is the conclusion*).

If $n = 0$, we have the special case of the derivation of β from zero premisses, i.e., $\varnothing \models \beta$ (where $\varnothing$ is the empty set), equivalently $\models \beta$ (and then the valid formulas of L_{PL} are formulas derived from an empty set of formulas).

Note that when we make derivations of this kind we use arbitrary sets of formulas. So, in what follows, let Γ, Δ be arbitrary set of such formulas (they may be finite, denumerable or even empty sets).[21]

Let us write down some properties of the relation "semantic consequence", where Γ, Δ, α, β are arbitrary.

Prop. 1. $\alpha \models \alpha$.

Prop. 2. If $\Gamma \models \alpha$, then $\Gamma \cup \Delta \models \alpha$ (written sometimes as $\Gamma, \Delta \models \alpha$).

Prop. 3. If $\Gamma \models \alpha$ and $\alpha \models \beta$, then $\Gamma \models \beta$.

Prop. 4. If $\Gamma \models \alpha$ and $\Gamma \models \alpha \supset \beta$, then $\Gamma \models \beta$.

Prop. 5. If $\models \alpha$, then $\Gamma \models \alpha$.

These facts about semantic consequence can be argued by using the above definition and some semantic notions given in Sect. 2.1 and 2.2. For example, Prop. 3 can be argued as follows. Suppose, by *reductio*, that $\Gamma \models \alpha$, $\alpha \models \beta$ and $\Gamma \not\models \beta$. It follows that there is an interpretation *int* of the propositional variables from Γ and β such that $[\Gamma]^{int} = 1$ and $[\beta]^{int} = 0$. From the first supposition, $\Gamma \models \alpha$, it follows that if $[\Gamma]^{int} = 1$, then $[\alpha]^{int} = 1$.

[21] If Γ is denumerable, then given the *finitist* character of "semantic consequence in PL" (comp. 3.3.3), if $\Gamma \models \alpha$, then there is a *finite* set $\Gamma_0 \subseteq \Gamma$ such that $\Gamma_0 \models \alpha$.

And then since $[\beta]^{int}=0$, it follows that $\alpha \not\models \beta$, contradicting the second assumption.

Similarly, Prop. 5 can be argued using Prop. 2 etc.

Normality Theorem (Norm). $\alpha_1,...,\alpha_n \models \beta$ *iff* $\alpha_1,...,\alpha_{n-1} \models \alpha_n \supset \beta$.

Proof. a) If $\alpha_1,...,\alpha_n \models \beta$, then $\alpha_1,...,\alpha_{n-1} \models \alpha_n \supset \beta$.

Assume $\alpha_1,...,\alpha_n \models \beta$. Let *int* be an interpretation of variables in $\alpha_1,...,\alpha_n,\beta$ such that $[\alpha_i]^{int}=1$ for $1 \le i \le n$. Then, by definition, $[\beta]^{int}=1$. It follows that $[\alpha_n]^{int}=1$ and $[\beta]^{int}=1$ and hence $[\alpha_n \supset \beta]^{int}=1$. Therefore, if $\alpha_1,...,\alpha_n \models \beta$, then $\alpha_1,...,\alpha_{n-1} \models \alpha_n \supset \beta$.

b) If $\alpha_1,...,\alpha_{n-1} \models \alpha_n \supset \beta$, then $\alpha_1,...,\alpha_n \models \beta$.

Assume $\alpha_1,...,\alpha_{n-1} \models \alpha_n \supset \beta$. Let *int* be an interpretation such that $[\alpha_i]^{int}=1$, $1 \le i \le n-1$. Then, by definition $[\alpha_n \supset \beta]^{int}=1$, and hence if $[\alpha_n]^{int}=1$, then $[\beta]^{int}=1$. Therefore, If $[\alpha_i]^{int}=1$, $1 \le i \le n$, then $[\beta]^{int}=1$, i.e., $\alpha_1,...,\alpha_n \models \beta$.

Corollary. $\alpha_1,...,\alpha_n \models \beta$ *iff* $\models \alpha_1 \supset (... \supset (\alpha_n \supset \beta)...)$.

Proof (repeated applications of Norm).

Theorem. $\alpha_1,...,\alpha_n \models \beta$ *iff* $\alpha_1 \wedge ... \wedge \alpha_n \models \beta$.

Proof. (1) $\alpha_1,...,\alpha_n \models \beta$ iff $\models \alpha_1 \supset (... \supset (\alpha_{n-1} \supset (\alpha_n \supset \beta))...)$; by Corollary.

(2) iff $\alpha_1 \supset (... \supset ((\alpha_{n-1} \wedge \alpha_n) \supset \beta)$; (1) PL[22], Subst$_{PL}$, Repl$_{PL}$.

(n) iff $\models (\alpha_1 \wedge ... \wedge \alpha_n) \supset \beta$; repeated applications of PL, Subst$_{PL}$ and Repl$_{PL}$.

(n+1) iff $\alpha_1 \wedge ... \wedge \alpha_n \models \beta$; Norm.

Exercises

1. Prove by contraposition the following equivalence:

$$\alpha \models \beta \text{ iff } \models \alpha \supset \beta.$$

2. Prove by *reductio ad absurdum*[23] the following equivalence:

$$\alpha_1, \alpha_2 \models \beta \text{ iff } \alpha_1 \wedge \alpha_2 \models \beta.$$

3. Prove the following equivalence

$$(\alpha \models \beta \text{ and } \beta \models \alpha) \text{ iff } \models \alpha \equiv \beta.$$

---

[22] By PL here we understand the application of the following formula of L_{PL}: $[p_{n-1} \supset (p_n \supset q)] \equiv [(p_{n-1} \wedge p_n) \supset q]$; comp. 2.2, 41.

[23] In what follows, *reductio*.

4. Prove that the following holds:

a) $\alpha, \neg\alpha \models \beta$, where β is arbitrary

b) ($\alpha \models \beta$ and $\alpha \models \neg\beta$) iff $\models \neg\alpha$

c) ($\alpha \models \beta$ and $\neg\alpha \models \beta$) iff $\models \beta$.

2.7. Decision procedures in PL

To decide in PL means to have an effective procedure (method) to determine whether a given formula of L_{PL} is valid, unsatisfiable, or neither.

2.7.1. Truth table method [24]

As we saw above (cf. 2.1) the truth table method enable us to answer the questions concerning the decidability. All we have to do is to make the truth table for the given formula α and to decide according to the Def 1-Def 3 of 2.1.2.

Exercises. Using truth table method test the validity of the following formulas of L_{PL}:

$$\alpha = [(p+q) \equiv \neg r] \supset (p \vee r)$$
$$\beta = [(p/q) \wedge r] \vee \neg p$$
$$\gamma = (\neg p \vee q) \supset [\neg(q \wedge \neg r) \supset (p/\neg r)].$$

2.7.2. Quine's method

This method is based on the reduction of a formula α to a truth value by application of some rules characterizing its connectives. More exactly, every binary[25] connective of L_{PL} is characterized by two rules, one for the case the formula containing it has a *true* argument, and the other for the case the same formula has a *false* argument. In both cases this formula can be reduced either to a truth value or to an argument negated or not. Let us take some examples.

Conjunction. $p \wedge q$

R_1. $(1 \wedge q) \equiv q$. Hence by this rule, if in a conjuction one argument is true (either p or q, conjunction being commutative) then this formula

[24] Also called *matrix method.*

[25] For negation the result is immediate.

containing $\wedge$ is reducible to the other argument.

R_2. $(0\wedge q)\equiv 0$, i.e., if in a conjunction one argument is false, then the conjunction is false.

Notice that these rules are not arbitrary conventions, they are based on the semantic definitions (truth tables) of the respective connectives, and can be stated by a simple inspection of these truth tables.

Disjunction. $p\vee q$

R_1. $(1\vee q)\equiv 1$. If in a disjunction one argument is true (any of them), then the respective disjunction is reducible to the truth value 1.

R_2. $(0\vee q)\equiv q$. If in a disjunction one argument is false (either p or q), then the disjunction is reducible to the other argument.

Implication. $p\supset q$

R_1. If in an implication one argument is true (consider the first three interpretations of p and q in the truth table of $\supset$), then $p\supset q$ is reducible to q.

R_2. If in an implication one argument is false (consider the last three interpretations of p and q in the truth table of $\supset$), then $p\supset q$ is reducible to $\neg p$.

Equivalence. $p\equiv q$

R_1. $(1\equiv q)\equiv q$

R_2. $(0\equiv q)\equiv\neg q$

Exercise. State the rules for +, / and $\downarrow$.

Example 1. $\alpha=[(p\equiv q)\wedge\neg r]\supset(\neg p\vee r)$

First of all, we choose the variable occurring most frequently,[26] and give it alternatively the truth values 1 and 0. If $p=1$, for example, then we replace p in α with its value 1, i.e.,

$p=1$

$[(1\equiv q)\wedge\neg r]\supset(0\vee r)$.

The antecedent of this implication contains an equivalence whose left argument is true. Then by R_1 of equivalence $1\equiv q$ is equivalent to q, and we can therefore replace, by $Repl_{PL}$, $1\equiv q$ with q. And by R_2 of disjunction, the consequent of the formula, $0\vee r$, is reducible to r, and then, by $Repl_{PL}$, from the preceding formula, we obtain

$(q\wedge\neg r)\supset r$.

Now, we choose again one variable, this time r, and give it alternatively the truth values 1 and 0, and repeat the process, i.e.,

[26] For *practical* reasons, for a faster reduction of α.

$r=1$

$(q\wedge 0)\supset 1$

By R_2 of conjunction, $(q\wedge 0)\equiv 0$, and hence $(0\supset 1)\equiv 1$. Whence if $p=1$ and $r=1$, regardless of the value of q, the formula α is 1.

Now we return to the preceding formula, $(q\wedge\neg r)\supset r$, and to r we assign 0, and obtain

$r=0$

$(q\wedge 1)\supset 0$

$q\supset 0$,

and by R_2 of $\supset$, $q\supset 0$ is equivalent to $\neg q$

$\neg q$.

Now if $q=0$, then $\neg q=1$, and if $q=1$, then $\neg q=0$. Hence, if $p=1$ and $r=0$, the truth value of α is reducible to the negated value of q.

Finally, we return to our formula α and for p we take the value 0, and repeat the preceding process, i.e.,

$[(0\equiv q)\wedge\neg r]\supset(1\vee r)$,

a formula reducible, by R_2 for equivalence and R_1 for disjunction, to

$(\neg q\wedge\neg r)\supset 1$,

a formula reducible, by R_1 for implication, to 1. Hence, if $p=0$, then $\alpha=1$ irrespective of the values of q and r. So, finally, α is satisfiable.

Let us take another example.

Example 2. $\beta=[(p\supset q)\wedge\neg r]\equiv(p/\neg q)$

$p=1$			$p=0$	
$[(1\supset q)\wedge\neg r]\equiv(1/\neg q)$			$[(0\supset q)\wedge\neg r]\equiv(0/\neg q)$	
$(q\wedge\neg r)\equiv\neg\neg q$			$(1\wedge\neg r)\equiv 1$	
$q=1$		$q=0$	$\neg r\equiv 1$	
$(1\wedge\neg r)\equiv 1$		$(0\wedge\neg r)\equiv 0$	$\neg r$	
$\neg r\equiv 1$		$0\equiv 0$	$r=1$	$r=0$
$\neg r$		1	0	1
$r=1$	$r=0$			
0	1			

Let us make the truth table of β: $[(p\supset q)\wedge\neg r]\equiv(p/\neg q)$

p	q	r	$p \supset q$	$\neg r$	$(p \supset q) \wedge \neg r$	$\neg q$	$p / \neg q$	β
1	1	1	1	0	0	0	1	0
1	1	0	1	1	1	0	1	1
1	0	1	0	0	0	1	0	1
1	0	0	0	1	0	1	0	1
0	1	1	1	0	0	0	1	0
0	1	0	1	1	1	0	1	1
0	0	1	1	0	0	1	1	0
0	0	0	1	1	1	1	1	1

If we compare both results (obtained by truth table and Quine's method) we observe that they coincide, i.e., if p=1 and q=1, the value of β is the negation of the value of r; if p=1 and q=0, then β=1, irrespective of the value of r. And if p=0, then the value of β is equivalent to the negation of r, irrespective of the value of q.

Exercises. Given the following formulas of L_{PL}:

$$\alpha = (p \wedge q) \supset [(p \supset q) \equiv r]$$
$$\beta = (p + q) \supset [(p \downarrow r) \vee q]$$
$$\gamma = (p \supset q) \supset [(\neg q \vee r) \supset (\neg r \supset \neg p)]$$

a) Using Quine's method, decide on the above formulas.

b) Verify the result using truth table method.

2.7.3. Reductio Test

Reductio Test is a variant of *reductio ad absurdum*. If we want to test the validity of a formula α, we begin by supposing its invalidity, and if this supposition leads to a contradiction, then our supposition is inconsistent and then α is valid. Let us take an example.

$$\alpha = [(p \supset q) \wedge (q \supset r)] \supset (p \supset r)$$

Suppose that α is not valid. It follows, by 2.1.2, Def 1, that there is an interpretation of the variables p, q and r such that (1) $[\alpha]^{int} = 0$. This means (2) $[(p \supset q) \wedge (q \supset r)]^{int} = 1$ and (3) $[p \supset r]^{int} = 0$. From (3) follows (4) $[p]^{int} = 1$ and (5) $[r]^{int} = 0$. From (2) it follows (6) $[p \supset q]^{int} = 1$ and (7) $[q \supset r]^{int} = 1$. But by (4) p is true in *int*, hence by (6) q must be true in *int*, i.e., (8) $[q]^{int} = 1$, and by (7) r must be true in *int*, i.e., (9) $[r]^{int} = 1$, a result contradicting (5). Therefore, α cannot be falsified.

Exercises. Using *Reductio Test*, decide on the validity of the following formulas of L_{PL}:

$\alpha = (p \supset q) \supset [(q \supset r) \supset (p \supset r)]$
$\beta = (p \supset q) \supset (\neg q \supset \neg p)$
$\gamma = [(p \equiv q) \wedge \neg p] \supset (p \supset q)$.

2.7.4. Normal forms in PL

Some other procedure to decide in PL is that of the normal forms. In contrast to the preceding semantical procedures, this is a syntactical one, i.e., it is based on the combination of symbols of L_{PL}, without appealing to the idea of truth value.

The normal forms we expose in this section are *Boolean normal forms*, a naming due to the fact that the only connectives contained in them are the Boolean ones: $\neg$, $\wedge$ and $\vee$. Two such normal forms will be analysed here: *conjunctive* and *disjunctive* normal forms.

2.7.4.1. Conjunctive normal forms

Definition 1. *A formula* α *of* L_{PL} *is in conjunctive normal form if it has the form of a conjunction* $D_1 \wedge ... \wedge D_n$, $n \geq 1$, *where each* D_i, $1 \leq i \leq n$ *has the form of a disjunction* $C_1 \vee ... \vee C_m$, $m \geq 1$, *where each* C_i, $1 \leq i \leq m$, *is a variable unnegated or negated at most once*.

Theorem 1. *Every formula* α *of* L_{PL} *can be equivalently converted in a conjunctive normal form* α_c.

Proof. A simple proof of this theorem can be given by indicating the steps of constructing α_c.

1. Replace all subformulas of α containing not-Boolean connectives with equivalent formulas containing only $\neg$, $\wedge$ and $\vee$; by applying $Repl_{PL}$.

The equivalent formulas we refer to can be obtained using these equivalences in PL[27]:

$(\alpha_1 \supset \alpha_2) \equiv (\neg \alpha_1 \vee \alpha_2)$
$(\alpha_1 \equiv \alpha_2) \equiv [(\neg \alpha_1 \vee \alpha_2) \wedge (\neg \alpha_2 \vee \alpha_1)]$
$(\alpha_1 + \alpha_2) \equiv \neg [(\neg \alpha_1 \vee \alpha_2) \wedge (\neg \alpha_2 \vee \alpha_1)]$
$(\alpha_1 / \alpha_2) \equiv \neg (\alpha_1 \wedge \alpha_2)$
$(\alpha_1 \downarrow \alpha_2) \equiv \neg (\alpha_1 \vee \alpha_2)$

[27] Based on the fact that the set $M = \{\neg, \wedge, \vee\}$ is complete; comp. 2.3.

The first step being performed, a formula α^1 is obtained such that it contains only $\neg$, $\wedge$, $\vee$ and the following holds: $\models\alpha\equiv\alpha^1$.

2. Replace all subformulas of α^1 of the form $\neg(\alpha_1\wedge\alpha_2)$ or $\neg(\alpha_1\vee\alpha_2)$; by $\mathrm{Repl}_{\mathrm{PL}}$, using the following equivalences:

$$\neg(\alpha_1\wedge\alpha_2)\equiv(\neg\alpha_1\vee\neg\alpha_2) \text{ (cf. 2.2,16)}$$

$$\neg(\alpha_1\vee\alpha_2)\equiv(\neg\alpha_1\wedge\neg\alpha_2) \text{ (cf.2.2,32)}$$

In the formula so obtained, α^2, only the variables occur negated, and the following holds: $\models\alpha^1\equiv\alpha^2$.

3. Replace all multiple negated variables of α^2, i.e., the variables with more than one negation, with their equivalents; by $\mathrm{Repl}_{\mathrm{PL}}$, using equivalences 1 and 2 of 2.2, e.g.

$$\neg\neg p\equiv p$$

$$\neg\neg\neg p\equiv\neg p$$

In the formula so obtained, α^3, any variable has at most one negation, and $\models\alpha^2\equiv\alpha^3$.

4. Replace all formulas of the form $\alpha_1\vee(\alpha_2\wedge\alpha_3)$ or $(\alpha_2\wedge\alpha_3)\vee\alpha_1$ with their equivalents; by $\mathrm{Repl}_{\mathrm{PL}}$, using 22 of 2.2, i.e.,

$$[\alpha_1\vee(\alpha_2\wedge\alpha_3)]\equiv[(\alpha_1\vee\alpha_2)\wedge(\alpha_1\vee\alpha_3)]$$

$$[(\alpha_2\wedge\alpha_3)\vee\alpha_1]\equiv[(\alpha_1\vee\alpha_2)\wedge(\alpha_1\vee\alpha_3)]$$

But if α^3 has the form $(\alpha_1\wedge\alpha_2)\vee(\alpha_3\wedge\alpha_4)$, then we use the following equivalence:

$$[(\alpha_1\wedge\alpha_2)\vee(\alpha_3\wedge\alpha_4)]\equiv[(\alpha_1\vee\alpha_3)\wedge(\alpha_1\vee\alpha_4)\wedge(\alpha_2\vee\alpha_3)\wedge(\alpha_2\vee\alpha_4)].$$

The formula so obtained, α_c, will have the form required by Def 1, i.e., it is the sought formula such that $\models\alpha^3\equiv\alpha_c$. Hence, by transitivity of equivalence, (cf. 2.2, 73), we finally have:

$$\models\alpha\equiv\alpha_c.$$

Example. $\alpha = (\neg p\equiv q)\vee(q/\neg r)$

1. $[(\neg\neg p\vee q)\wedge(\neg q\vee\neg p)]\vee\neg(q\wedge\neg r) = \alpha^1$
2. $[(\neg\neg p\vee q)\wedge(\neg q\vee\neg p)]\vee(\neg q\vee\neg\neg r) = \alpha^2$
3. $[(p\vee q)\wedge(\neg q\vee\neg p)]\vee(\neg q\vee r) = \alpha^3$
4. $(p\vee q\vee\neg q\vee r)\wedge(\neg q\vee\neg p\vee\neg q\vee r) = \alpha_c$

The way we decide whether or not a given formula is valid is indicated by the following theorem.

Theorem 2. *A formula* α *of* L_{PL} *in conjunctive normal form* $\alpha_c = \mathrm{D}_1\wedge...\wedge\mathrm{D}_n$, $n\geq 1$, *is valid if and only if every conjunct* D_i, $1\leq i\leq n$, *contains some variable and its negation*.

Proof. If every conjunct D_i has at least one variable p_j together with its negation, $\neg p_j$, then every conjunct contains a valid disjunction of the form $p_j \vee \neg p_j$, and hence it cannot be falsified. Therefore, the formula α_c is valid, and then α is also valid (by Theorem 1).

Example. $\alpha_c = (p \vee \neg q \vee q) \wedge (\neg p \vee \neg q \vee p) \wedge (r \vee p \vee \neg r)$

The formula α_c satisfies the preceding theorem, and then $\models \alpha_c$. Indeed, this is the case, since α_c can be re-written as

$$\alpha_c = [p \vee (q \vee \neg q)] \wedge [(p \vee \neg p) \vee \neg q] \wedge [(r \vee \neg r) \vee p].$$

Each conjunct contains a valid subformula, a disjunction of a variable with its negation. Hence in each conjunct D_i an argument is always true, and then, by the semantic definition of "$\vee$", it cannot be falsified. Finally, by Theorem 1, $\models \alpha \equiv \alpha_c$, and then $\models \alpha$.

On the contrary, if α_c would have a conjunct C_j not containing some variable negated and unnegated, then α would not be valid (argue).

Definition 2. *A formula α of L_{PL} is in **perfect** conjunctive normal form if and only if α is in conjunctive normal form and each conjunct contains all the variables of α.*

Theorem 3. *Every formula α of L_{PL} can be equivalently converted in a perfect conjunctive normal form α_c^*.*

Proof. As in Theorem 1 we only indicate the steps of constructing α_c^*.

1. Construct the conjunctive normal form α_c of α.
2. If a conjunct D_i does not contain a variable p, then this variable will be introduced using the following equivalences:

$$D_i \equiv [D_i \vee (p \wedge \neg p)] \equiv [(D_i \vee p) \wedge (D_i \vee \neg p)].$$

The formula so obtained is just α_c^*.

Example. $\alpha_c = (\neg p \vee q \vee \neg r) \wedge (p \vee r)$

As can be seen the second conjunct of α_c, $p \vee r$, does not contain the variable q. Hence we introduce it in the way indicated above by the step 2,

$$(p \vee r) \equiv [(p \vee r) \vee (q \wedge \neg q)] \equiv [(p \vee r \vee q) \wedge (p \vee r \vee \neg q)],$$

and then $\alpha_c^* = (\neg p \vee q \vee \neg r) \wedge (p \vee r \vee q) \wedge (p \vee r \vee \neg q)$.

Of course, also in this case we have $\models \alpha_c \equiv \alpha_c^*$, and then $\models \alpha \equiv \alpha_c \equiv \alpha_c^*$. And, finally, Theorem 2 will also hold for perfect conjunctive normal form α_c^*.

2.7.4.2. Disjunctive normal forms [28]

Definition 1. *A formula* α *of* L_{PL} *is in disjunctive normal form if it has the form of a disjunction* $C_1 \vee ... \vee C_n$, $n \geq 1$, *where each* C_i, $1 \leq i \leq n$ *has the form of a conjunction* $D_1 \wedge ... \wedge D_m$, $m \geq 1$, *where each* D_i, $1 \leq i \leq m$, *is a variable unnegated or negated at most once.*

Theorem 1. *Every formula* α *of* L_{PL} *can be equivalently converted in a disjunctive normal form* α_d.

Proof. As in the proof of Theorem 1 of the preceding section, it is enough to indicate the steps of constructing α_d.

The steps (1)-(3), as in Theorem 1 of 2.7.4.1, and so the corresponding formula α^3 is obtained.

4. Replace all formulas of the form $\alpha_1 \wedge (\alpha_2 \vee \alpha_3)$ or $(\alpha_2 \vee \alpha_3) \wedge \alpha_1$ with their equivalents; by Repl_{PL}, using 6 of 2.2, i.e.,

$$[\alpha_1 \wedge (\alpha_2 \vee \alpha_3)] \equiv [(\alpha_1 \wedge \alpha_2) \vee (\alpha_1 \wedge \alpha_3)]$$
$$[(\alpha_2 \vee \alpha_3) \wedge \alpha_1] \equiv [(\alpha_1 \wedge \alpha_2) \vee (\alpha_1 \wedge \alpha_3)]$$

But if α^3 has the form $(\alpha_1 \vee \alpha_2) \wedge (\alpha_3 \vee \alpha_4)$, then we use the following equivalence:

$$[(\alpha_1 \vee \alpha_2) \wedge (\alpha_3 \vee \alpha_4)] \equiv [(\alpha_1 \wedge \alpha_3) \vee (\alpha_1 \wedge \alpha_4) \vee (\alpha_2 \wedge \alpha_3) \vee (\alpha_2 \wedge \alpha_4)].$$

The formula so obtained, α_d, will have the form required by Definition 1, i.e., it is the disjunctive normal form of α and the following holds $\models \alpha \equiv \alpha_d$.

Example. If α is the formula given in the preceding section, $(\neg p \equiv q) \vee (q / \neg r)$, then in the step 3 we obtained $\alpha^3 = [(p \vee q) \wedge (\neg q \vee \neg p)] \vee (\neg q \vee r)$. To obtain α_d, this time we apply the step 4 mentioned above, with respect to the formula in brackets, and obtain
$\{[(p \vee q) \wedge (\neg q \vee \neg p)] \vee (\neg q \vee r)\} \equiv \{[(p \wedge \neg q) \vee (p \wedge \neg p) \vee (q \wedge \neg q) \vee (q \wedge \neg p)] \vee \neg q \vee r\}$
$\equiv [(p \wedge \neg q) \vee (q \wedge \neg p) \vee \neg q \vee r] = \alpha_d$ (from the preceding formula by eliminating, equivalently, the unsatisfiable conjunctions $p \wedge \neg p$ and $q \wedge \neg q$.

Using the disjunctive normal forms we decide whether or not a given formula is unsatisfiable.

Theorem 2. *A formula* α *of* L_{PL} *in disjunctive normal form* $\alpha_d = C_1 \vee ... \vee C_n$, $n \geq 1$, *is unsatisfiable if and only if every disjunct* C_i, $1 \leq i \leq n$, *contains some variable and its negation.*

[28] As can be expected, since $\vee$ is the dual of $\wedge$, this section will be a parallel to the preceding section, in a way easy recognizable.

Proof. If every disjunct C_i has at least one variable p_j occuring both negated and unnegated, then every disjunct contains an unsatisfiable conjunction of the form $p_j \wedge \neg p_j$, and hence the whole disjunct is false. Therefore, the formula α_d is always false, and then, by Theorem 1, α will be unsatisfiable.

Example. Let us suppose that α_d is $(p \wedge \neg p \wedge q) \vee (q \wedge r \wedge \neg q) \vee (r \wedge \neg r)$. It is equivalent to $[(p \wedge \neg p) \wedge q] \vee [(q \wedge \neg q) \wedge r] \vee (r \wedge \neg r)$, each disjunct containing one unsatisfiable subformula of the form $p_j \wedge \neg p_j$, and then being always false (by the semantic definition of $\wedge$). Whence α_d is always false (by the semantic definition of $\vee$). And since $\models \alpha \equiv \alpha_d$, it follows that α is also unsatisfiable.

On the contrary, if α_d does not conform to Theorem 2, then it is not unsatisfiable (argue!).

Definition 2. *A formula α of L_{PL} is in **perfect** disjunctive normal form if and only if α is in disjunctive normal form and each disjunct contains all the variables of α.*

Theorem 3. *Every formula α of L_{PL} can be equivalently converted in a perfect disjunctive normal form α_d^*.*

Proof. The steps of constructing α_d^* are the following:

1. Construct the disjunctive normal form α_d of α.
2. If a disjunct C_i does not contain a variable p, then this variable will be introduced using the following equivalences:

$$C_i \equiv [C_i \wedge (p \vee \neg p)] \equiv [(C_i \wedge p) \vee (C_i \wedge \neg p)],$$

and the formula so obtained is just α_d^*.

Example. $\alpha_d = (p \wedge \neg q) \vee (q \wedge \neg p \wedge r)$.

α_d has the form $C_1 \vee C_2$, but C_1 does not contain the variable r. We introduce this variable, using the preceding equivalences, i.e.,

$$(p \wedge \neg q) \equiv [(p \wedge \neg q) \wedge (r \vee \neg r)] \equiv [(p \wedge \neg q \wedge r) \vee (p \wedge \neg q \wedge \neg r)]$$

and then $\alpha_d^* = (p \wedge \neg q \wedge r) \vee (p \wedge \neg q \wedge \neg r) \vee (q \wedge \neg p \wedge r)$.

Of course, we also have $\models \alpha_d \equiv \alpha_d^*$, and then $\models \alpha \equiv \alpha_d \equiv \alpha_d^*$, and then Theorem 2 will also hold for perfect disjunctive normal form α_d^* of α.

Exercise. Are the following formulas of L_{PL} valid, unsatisfiable or neither?

$$\alpha_1 = (p \wedge q) \supset (p \equiv q)$$
$$\alpha_2 = [p \vee (q/r)] \vee [(p \downarrow r) \vee \neg q]$$
$$\alpha_3 = (p \downarrow \neg q) \supset [(q/r) \vee \neg r]$$
$$\alpha_4 = \{[(p/q) \wedge \neg r] \supset (r \downarrow r)\}$$
$$\alpha_5 = (p \equiv q) \wedge (p \wedge \neg q).$$

2.8. Representability of truth functions in PL

2.8.1. Representation Theorem

As we saw in 2.1.1, Definition, an *n*-ary truth function is a mapping from the set T of 2^n *n*-tuples of truth values[29] to the set {1,0}. If, for example, T={(1,1),(1,0),(0,1),(0,0)} and if $f^2(1,1)=1$, $f^2(1,0)=0$, $f^2(0,1)=0$ and $f^2(0,0)=1$, then f^2 is the equivalence.

Let $\alpha(p_1\dots,p_n)$ be a formula of L_{PL}. This formula defines exactly an *n*-ary function f^n satisfying the following condition:

*Cond***.** For every interpretation i: $[\alpha]^i = f^n([p_1]^i,...,[p_n]^i)$.[30]

Definition. *A formula* α *of* L_{PL} *is called a representant of a truth function* f^n *if the only variables in* α *are* $p_1\dots,p_n$ *and Cond is satisfied.*

By the following theorem one can answer the question whether any *n*-ary function f^n is representable by at least one[31] formula of L_{PL}.

Representation Theorem. *Any n*-ary *function* f^n *can be represented by a formula* $\alpha(p_1,...,p_n)$ *of* L_{PL} *and this formula can be effectively constructed.*

Proof. (1) Let $T=\{T_1,...,T_{2^n}\}$ be the set of all *n*-tuples built up using 1 and 0. For every $1\le i\le 2^n$ we set $f^n(i)=f^n(T_i)$. If, for example, $n=3$ and $T_i=(1,0,1)$, then $f^3(i)=f^3(1,0,1)$. By $T_{i[k]}$ we understand the *k*-th member of the *n*-tuple T_i.

(2) Now we form the 2^n conjunctions α_i, $1\le i\le 2^n$, of variables in the following way: if in the interpretation i the variable p_k, $1\le k\le n$, has the value 1, then it appears in the conjunction α_i as p_k, and if its value is 0, then we set $\neg p_k$.

Example. If $\alpha = p_1\equiv p_2$, then $i_1,...,i_4$ are the 2^2 possible interpretations of the variables p_1, p_2. For i_2 the conjunction α_2 will be $p_1\wedge\neg p_2$, and for i_4 $\alpha_4=\neg p_1\wedge\neg p_2$ etc.

More generally,

$$\alpha_i = \alpha_{i1}\wedge\alpha_{i2}\wedge\dots\wedge\alpha_{in}, \text{ where } \alpha_{ik} = \begin{cases} p_k & \text{if} \quad T_{i[k]}=1 \\ \neg p_k & \text{if} \quad T_{i[k]}=0 \end{cases}$$

[29] This means that every number of an *n*-tuple is either 1 or 0.

[30] Here "i" is another symbolization (more convenient sometimes) for the *interpretation* of the propositional variables of a formula α, given above by "int".

[31] Evidently, if f^n is represented by a formula $\alpha(p_1,...,p_n)$ ($n\ge 1$), then f^n is also represented by an *equivalent* formula of L_{PL}.

Let now α_i^+ be the conjunction α_i if $f^n(i)=1$, and α_i^- the conjunction α_i if $f^n(i)=0$. For every interpretation i we'll set $[p_k]^i = T_{i[k]}$. Hence,

a) $[\alpha_i]^i = 1$.

b) For any $j \neq i$, $1 \leq j \leq 2^n$, $[\alpha_j]^i = 0$, since, by construction, every conjunction α_j different from α_i, contains at least one member α_{jk} such that $\alpha_{jk} = \neg\alpha_{ik}$ or $\alpha_{ik} = \neg\alpha_{jk}$. Hence if $[\alpha_{ik}]^i = 1$, then $[\alpha_{jk}]^i = 0$ or vice versa. And then, since $[\alpha_i]^i = 1$, it follows that $[\alpha_j]^i = 0$.

Example. For the formula $\alpha = p_1 \equiv p_2$, $\alpha_1 = p_1 \wedge p_2$. In the interpretation i_1 α_1 is, evidently, true. But any other α_j, $j \neq 1$, will be false in i_1; $\alpha_2 = p_1 \wedge \neg p_2$, for example, contains $\alpha_{22} = \neg p_2$ and then $\alpha_{22} = \neg\alpha_{12}$. And since α_{12}, i.e., p_2, is true in the interpretation i_1, $[\alpha_{22}]^{i_1} = 0$. Similar, α_3 and α_4 can be shown to be false in the interpretation i_1.

(3) Let us consider all interpretations i such that $f^n(i)=1$. Let $m \neq 0$ be their number and let $\alpha_1^+, \ldots, \alpha_m^+$ be the corresponding conjunctions. Let *Disj* $\alpha^+ = \alpha_1^+ \vee \ldots \vee \alpha_m^+$. All we have to do is to show that the formula *Disj* α^+ represents the function f^n, i.e., for every interpretation i *Cond* holds, i.e.,

$$[\mathit{Disj}\ \alpha^+]^i = f^n([p_1]^i, \ldots, [p_n]^i) = f^n(i).$$

By (2) a) and b) above, for each i, $1 \leq i \leq m$, exactly one formula α_i, is true in i, i.e., $[\alpha_i]^i = 1$ and then $[\mathit{Disj}\ \alpha^+]^i = 1$. But for such an i $f^n(i) = 1$. On the other hand, for any interpretation $j \neq i_1, \ldots, i_m$, $[\mathit{Disj}\ \alpha^+]^j = 0$, since in j only α_j is true, i.e., $[\alpha_j^-]^i = 1$, but α_j^- is not a member of Disj α^+. And since $f^n(j) = 0$, it follows that *Cond* also holds, and then *Disj* α^+ represents f^n.

(4) Now let us consider all interpretations i such that $f^n(i) = 0$. Let $r \neq 0$ be their number and let $\alpha_1^-, \ldots, \alpha_r^-$ be the corresponding conjunctions. Let *Conj* $\neg\alpha^- = \neg\alpha_1^- \wedge \ldots \wedge \neg\alpha_r^-$. We have to show that the formula *Conj* $\neg\alpha^-$ also represents the function f^n, i.e., *Cond* holds.

Again, by (2) a) and b) above, for each i, $1 \leq i \leq r$, exactly one formula, α_i, is true in i, and then $\neg\alpha_i$ is the only conjunction false in the interpretation i. It follows that the whole conjunction $\neg\alpha_1^- \wedge \ldots \wedge \neg\alpha_r^-$ will be false in i. Hence $[\text{Conj}\neg\alpha^-]^i = 0$. But for each i, $1 \le i \le r$, $f^n(i) = 0$.

On the other hand, for any interpretation $j \ne i_1,...,i_r$, $[\text{Conj}\neg\alpha^-]^j = [\neg\alpha_1^- \wedge ... \wedge \neg\alpha_m^-]^j = 1$, since in j only α_j is true, i.e., $[\alpha_j^-]^j = 1$ and then all the formulas $\alpha_1^-,...,\alpha_r^-$ will be false, and hence their negations will be true in j. But for j $f^n(j) = 1$, and then *Cond* holds. Therefore, *Conj* $\neg\alpha^-$ also represents the function f^n.

(5) If $r = 0$, then $f^n(i) = 1$ for every $1 \le i \le 2^n$. In this case the formula $\alpha = \alpha_1 \vee \alpha_2 \vee ... \vee \alpha_{2^n}$ represents the function f^n. And if $m = 0$, then $\alpha = \neg\alpha_1 \wedge \neg\alpha_2 \wedge ... \wedge \neg\alpha_{2^n}$ represents the function f^n.

Example. $\alpha = p_1 \equiv p_2$

The formulas α_i corresponding to the 4 interpretations, $i_1,...,i_4$, are: $\alpha_1^+ = p_1 \wedge p_2$, $\alpha_2^- = p_1 \wedge \neg p_2$, $\alpha_3^- = \neg p_1 \wedge p_2$ and $\alpha_4^+ = \neg p_1 \wedge \neg p_2$. It is easy to see that for each i $[\alpha_i]^i = 1$ and for any $j \ne i$ $[\alpha_j]^i = 0$. In this example, $m = 2$ and $r = 2$, corresponding to the cases in which $f^2(i) = 1$ and $f^2(i) = 0$, respectively. Since for $i = 1$, $[\alpha_1]^{i_1} = 1$, i.e., $[p_1 \wedge p_2]^{i_1} = 1$, it follows that $[(p_1 \wedge p_2) \vee (\neg p_1 \wedge \neg p_2)]^{i_1} = 1$, and since for $i = 4$, $[\alpha_4]^{i_4} = 1$, i.e., $[\neg p_1 \wedge \neg p_2]^{i_4} = 1$, it follows that $[(p_1 \wedge p_2) \vee (\neg p_1 \wedge \neg p_2)]^{i_4} = 1$. And i_1 and i_4 are the only interpretations for which *Disj* $\alpha^+ = (p_1 \wedge p_2) \vee (\neg p_1 \wedge \neg p_2)$ is true. Then *Cond* holds and thus *Disj* α^+ is a representant of f^2.

Now, we take the cases for which $f^2(i) = 0$, the interpretations i_2 and i_3, respectively. Since $[\alpha_2^-]^{i_2} = 1$, it follows that $[\neg\alpha_2^-]^{i_2} = 0$, and hence $[\neg\alpha_2^- \wedge \neg\alpha_3^-]^{i_2} = 0$. The same holds of α_3^-, and then $[\neg\alpha_2^- \wedge \neg\alpha_3^-]^{i_3} = 0$. And since i_2 and i_3 are the only interpretations for which *Conj* $\neg\alpha^- = \neg\alpha_2^- \wedge \neg\alpha_3^- = \neg(p_1 \wedge \neg p_2) \wedge \neg(\neg p_1 \wedge p_2)$ is false, it follows that *Conj* $\neg\alpha^-$ represents the function f^2.

Of course, f^2 will be also represented by any formula equivalent to *Disj* α^+ or *Conj* $\neg\alpha^-$, e.g. $(p \supset q) \wedge (q \supset p)$ or $(\neg p \vee q) \wedge (\neg q \vee p)$, or $\neg(p \wedge \neg q) \wedge \neg(q \wedge \neg p)$ etc.

2.8.2. Representation Theorem and Boolean normal forms

Since the formula representing a function f^n is either of the form *Disj* α^+, i.e., $\alpha_1^+ \vee \ldots \vee \alpha_m^+$, where each α_i^+, $1 \le i \le m$, is a conjunction of variables (negated or not), or of the form *Conj* $\neg\alpha^-$, i.e., $\neg\alpha_1^- \wedge \ldots \wedge \neg\alpha_r^-$, where each α_i^-, $1 \le i \le r$, is a conjunction of the variable (negated or not)[32], is to be expected that the idea of representation be directly correlated to Boolean normal forms. Let us take an example.

Example 1. Let $\alpha = [(p_1 \equiv p_2) \vee p_3] \equiv (p_1 \supset \neg p_2)$

p_1	p_2	p_3	$p_1 \equiv p_2$	$(p_1 \equiv p_2) \vee p_3$	$\neg p_2$	$p_1 \supset \neg p_2$	α
1	1	1	1	1	0	0	0
1	1	0	1	1	0	0	0
1	0	1	0	1	1	1	1
1	0	0	0	0	1	1	0
0	1	1	0	1	0	1	1
0	1	0	0	0	0	1	0
0	0	1	1	1	1	1	1
0	0	0	1	1	1	1	1

As we saw above, by Representation Theorem, the representant of a function f^n can be effectively constructed. By the above truth table we have the truth function corresponding to the given formula α. Now, if we proceed as in 2.8.1,(3), we choose the interpretations i for which f(i) = 1, that is 3, 5, 7 and 8, and construct the formula *Disj* α^+. The conjunctions α_i^+ will be

$\alpha_3 = p_1 \wedge \neg p_2 \wedge p_3$; $\qquad \alpha_5 = \neg p_1 \wedge p_2 \wedge p_3$

$\alpha_7 = \neg p_1 \wedge \neg p_2 \wedge p_3$; $\qquad \alpha_8 = \neg p_1 \wedge \neg p_2 \wedge \neg p_3$,

and then $m = 4$. These 4 interpretations represent all possible interpretations for which *Disj* α_i^+ is true. And then a formula representing f^3 is:

Disj $\alpha^+ = (p_1 \wedge \neg p_2 \wedge p_3) \vee (\neg p_1 \wedge p_2 \wedge p_3) \vee (\neg p_1 \wedge \neg p_2 \wedge p_3) \vee (\neg p_1 \wedge \neg p_2 \wedge \neg p_3)$.

As can be seen, *Disj* α^+ is just the *perfect disjunctive normal form* of α.

For the construction of the other representant of the function f^3 we choose the interpretations i such that f(i) = 0, that is the interpretations 1, 2, 4, and 6 (i.e., $r = 4$), and construct the formula *Conj* $\neg\alpha^-$. This conjunction, according to 2.8.1,(4) will be:

[32] By distribution of $\neg$ each conjunction α_i^- becomes a disjunction.

$Conj\,\neg\alpha^{-} = \neg(p_1\wedge p_2\wedge p_3)\wedge\neg(p_1\wedge p_2\wedge\neg p_3)\wedge\neg(p_1\wedge\neg p_2\wedge\neg p_3)\wedge\neg(\neg p_1\wedge p_2\wedge\neg p_3)$, equivalently, $(\neg p_1\vee\neg p_2\vee\neg p_3)\wedge(\neg p_1\vee\neg p_2\vee p_3)\wedge(\neg p_1\vee p_2\vee p_3)\wedge(p_1\vee\neg p_2\vee p_3)$, and this last formula is just the *perfect conjunctive normal form* of the given formula α.

Evidently, if formula α were valid, then $Disj\,\alpha^{+}$ would have 8 members, corresponding to the 8 possible interpretations of p_1, p_2 and p_3. And if α were unsatisfiable, then $Conj\ \neg\alpha^{-}$ would contain 8 formulas α_i^{-}.

These considerations on the idea of representation can be further on developed. Let us take an example.

Example 2. $\alpha = (p_1\equiv p_2)\supset(\neg p_1\vee p_2)$

The corresponding function will be given by the following truth table.

p_1	p_2	$p_1 \equiv p_2$	$\neg p_1$	$\neg p_1 \vee p_2$	α
1	1	1	0	1	1
1	0	0	0	0	1
0	1	0	1	1	1
0	0	1	1	1	1

The formula α is therefore valid, and then $Disj\ \alpha^{+}$ will be

$$Disj\ \alpha^{+} = (p_1\wedge p_2)\vee(p_1\wedge\neg p_2)\vee(\neg p_1\wedge p_2)\vee(\neg p_1\wedge\neg p_2).$$

This formula can be equivalently re-written as $\alpha^{*} = (p_1\vee\neg p_1)\wedge(p_2\vee\neg p_2)$ (comp. 2.7.4.2, the step 4 in the proof of Theorem 1), or in the abbreviated form $\alpha^{*} = \bigwedge_{i=1}^{2}(p_i \vee \neg p_i)$. Consequently, $\models Disj\ \alpha^{+} = \bigwedge_{i=1}^{2}(p_i \vee \neg p_i)$.

By a partial disjunction of $Disj\ \alpha^{+}$ we understand any disjunct of $Disj\ \alpha^{+}$ or any disjunction of disjuncts of $Disj\ \alpha^{+}$ including $Disj\ \alpha^{+}$.

Let now D_1 and D_2 be any partial disjunctions of $Disj\ \alpha^{+}$ such that $Disj\ \alpha^{+} = D_1\vee D_2$. Then the following holds:

a) $\models D_1\vee D_2$, evident.

b) notSat $D_1\wedge D_2$.

Argument. If, for example, $D_1 = (p_1\wedge p_2)\vee(p_1\wedge\neg p_2)\vee(\neg p_1\wedge p_2)$, and $D_2 = \neg p_1\wedge\neg p_2$, then $D_1\wedge D_2$ will be the formula

$$[(p_1\wedge p_2)\vee(p_1\wedge\neg p_2)\vee(\neg p_1\wedge p_2)]\wedge(\neg p_1\wedge\neg p_2),$$

equivalent, by distribution, to the formula

$$(p_1\wedge p_2\wedge\neg p_1\wedge\neg p_2)\vee(p_1\wedge\neg p_2\wedge\neg p_1\wedge\neg p_2)\vee(\neg p_1\wedge p_2\wedge\neg p_1\wedge\neg p_2).$$

As can be observed, each disjunct contains a variable negated and unnegated, and then this formula is unsatisfiable (by 2.7.4.2, Theorem 2), i.e., notSat $D_1\wedge D_2$.

But from a) and b) it follows

c) $\models D_1 \equiv \neg D_2$, or $\models \neg D_1 \equiv D_2$, respectively.

Proof. (1) $\models (D_1 \vee D_2) \equiv (\neg D_2 \supset D_1)$; PL.

(2) $\models (D_1 \vee D_2)$ iff $\models \neg D_2 \supset D_1$; (1), 2.1.2, Th 3.

(3) $\models \neg(D_1 \wedge D_2) \equiv (D_1 \supset \neg D_2)$; PL.

(4) $\models \neg(D_1 \wedge D_2)$ iff $\models D_1 \supset \neg D_2$, 2.1.2, Th 3.

(5) notSat $D_1 \wedge D_2$ iff $\models D_1 \supset \neg D_2$; by 2.1.2, Th 1.

(6) If $\models D_1 \vee D_2$ and notSat $D_1 \wedge D_2$, then $\models \neg D_2 \supset D_1$ and $\models D_1 \supset \neg D_2$; (2), (5) PL.[33]

(7) If $\models D_1 \vee D_2$ and notSat $D_1 \wedge D_2$, then $\models D_1 \equiv \neg D_2$; (6) PL.

The formulas D_1 and D_2 are called *complementary*.
This result can be used in the simplification of the formulas, in the following sense: if $\models D_1 \equiv \neg D_2$, then in our Example 2 D_1 can be equivalently reduced to $\neg(\neg p_1 \wedge \neg p_2)$, or to $p_1 \vee p_2$, respectively.

These considerations enable us to reformulate the representation theorem in the following way. If $\alpha(p_1,...,p_n)$ is a formula of L_{PL} containing the variables $p_1,...,p_n$, by *Disj*$(p_1,...,p_n)$ we denote the formula obtained by disjunction of *all* 2^n formulas α_i constructed as in 2.8.1,(2).

Let us consider *Disj*$(p_1,...,p_n)$ and let α be a representant of an *n*-ary truth function f^n. The following holds:

1. $\models \alpha \equiv Disj\ \alpha^+$ $(m \neq 0)$
2. $\models \alpha \equiv Conj\ \neg\alpha^-$ $(r \neq 0)$
3. $\models Conj\ \neg\alpha^- \equiv \neg Disj\ \alpha^-$; cf. 2.2, 32.
4. $Disj\ \alpha^+ \vee Disj\ \alpha^- = Disj\ (p_1,...,p_n)$ $(m+r = 2^n)$.

Now, if the empty partial disjunction of *Disj* $(p_1,...,p_n)$ is defined as the *negation* of the formula *Disj* $(p_1,...,p_n)$, the representation theorem can be stated in the following terms.

Representation Theorem (variant). *Any n-ary function* f^n *can be represented by a partial disjunction of the formula Disj* $(p_1,...,p_n)$ *or, equivalently, by the negation of its complement.*

Corollary. *Any formula* $\alpha(p_1,...,p_n)$ *of* L_{PL} *can be expressed by a partial disjunction of Disj* $(p_1,...,p_n)$ *or by the negation of its complement. If*

[33] I.e. by a metalingvistic application of a valid formula of L_{PL}; cf. 2.2, 82.

$\models \alpha(p_1,...,p_n)$, *then* $\alpha = Disj\,(p_1,...,p_n)$ *and if* nonSat $\alpha(p_1,...,p_n)$, *then* $\alpha = \neg Disj\,(p_1,...,p_n)$.

Proof. Let us firstly remark that the number of nonequivalent formulas containing the variables $p_1,...,p_n$ is $2^{(2^n)}$, since the number of distinct interpretations of the *n* variables is 2^n and every formula $\alpha(p_1,...,p_n)$ defines, for these 2^n interpretations, a certain *n*-ary truth function f^n. And by *Cond* of 2.8.1, we deduce that for any interpretation i:

$$[\alpha_j]^i = [\alpha_k]^i \text{ iff } f_j^n([p_1]^i,...[p_n]^i) = f_k^n([p_1]^i,...[p_n]^i),$$

and then $\models \alpha_j \equiv \alpha_k$ iff $f_j^n = f_k^n$. And since the number of *n*-ary functions is $2^{(2^n)}$ it follows that the number of nonequivalent formulas in $p_1,...,p_n$ is $2^{(2^n)}$.

Then, since the empty partial disjunction of $Disj\,(p_1,...,p_n)$ is $\neg Disj\,(p_1,...,p_n)$, according to the mention made above, the number of all partial disjunctions of $Disj\,(p_1,...,p_n)$ is $2^{(2^n)}$, coinciding with the number of functions f^n. Now, by the preceding variant of representation theorem, we can conclude that each of the $2^{(2^n)}$ distinct formulas $\alpha(p_1,...,p_n)$ can be equivalently expressed by a partial disjunction of $Disj\,(p_1,...,p_n)$ or by the negation of its complement, since each formula α defines a function f^n, representable by this partial disjunction or by the negation of its complement.

Example 3. For $n = 2$, by the preceding considerations, the number of formulas $\alpha(p,q)$ constructible with p and q is $2^{(2^n)} = 16$, and this number is just the number of the 16 binary truth functions of Wittgenstein's Table. We have to show that any such formula, representing a truth function, can be equivalently expressed by a partial disjunction of *Disj*(p,q), or by the negation of its complement.

$$Disj(p,q) = (p\wedge q)\vee(p\wedge\neg q)\vee(\neg p\wedge q)\vee(\neg p\wedge\neg q).$$

For pTq (tautology), we set, simply, $\alpha_1 = Disj(p,q)$ or, equivalently, $(p\vee\neg p)\wedge(q\vee\neg q)$, equivalent $p\vee\neg p$. For pCq (contradiction), we set $\alpha_{16} = \neg Disj(p,q)$ or, equivalently $(p\wedge\neg p)\vee(q\wedge\neg q)$, equivalent $p\wedge\neg p$.

$\alpha_2 = p\vee q$, equivalent $(p\wedge q)\vee(p\wedge\neg q)\vee(\neg p\wedge q)$; cf. 2.8.2, a)-c), since $(p\vee q)\equiv\neg(\neg p\wedge\neg q)=\neg D_2$ and since $D_1(=(p\wedge q)\vee(p\wedge\neg q)\vee(\neg p\wedge q))\equiv\neg D_2$ $(=\neg(\neg p\wedge\neg q))$, by c).

$\alpha_3 = p\subset q$, equivalent $(p\wedge q)\vee(p\wedge\neg q)\vee(\neg p\wedge\neg q)$; cf. 2.8.2, a)-c).

α_4 = p, equivalent $(p\wedge q)\vee(p\wedge\neg q)$; PL.
α_5 = $p\supset q$, equivalent $(p\wedge q)\vee(\neg p\wedge q)\vee(\neg p\wedge\neg q)$; cf. 2.8.2, a)-c).
α_6 = q, equivalent $(p\wedge q)\vee(\neg p\wedge q)$; PL.
α_7 = $p\equiv q$, equivalent $(p\wedge q)\vee(\neg p\wedge\neg q)$; PL.
α_8 = $p\wedge q$.
α_9 = p/q, equivalent $(p\wedge\neg q)\vee(\neg p\wedge q)\vee(\neg p\wedge\neg q)$; cf. 2.8.2, a)-c).
α_{10} = p+q, equivalent $(\neg p\wedge q)\vee(p\wedge\neg q)$; PL.
α_{11} = $\neg q$, equivalent $(p\wedge\neg q)\vee(\neg p\wedge\neg q)$; cf. 2.8.2, a)-c).
α_{12} = $p\not\supset q$, equivalent $(p\wedge\neg q)$, from α_5, 2.8.2, a)-c).
α_{13} = $\neg p$, equivalent $(\neg p\wedge q)\vee(\neg p\wedge\neg q)$; PL.
α_{14} = $p\not\subset q$, equivalent $\neg p\wedge q$; from α_3, 2.8.2, a)-c).
α_{15} = $p\downarrow q$, equivalent $\neg p\wedge\neg q$.

Remark. By *Corollary*, given a formula $\alpha(p_1,...,p_n)$ of L_{PL}, the partial disjunction of *Disj* $(p_1,...,p_n)$ by which it can be equivalently expressed may be constructible, by constructing the *perfect* disjunctive normal form of α (comp. 2.7.4.2).

2.9. Interpolation in PL

Interpolation Theorem for PL. *If* $\models\alpha\supset\beta$ *and* α *and* β *have at least one variable in common, then there is a formula* γ *of* L_{PL} *all of whose variables occur in both* α *and* β *such that* $\models\alpha\supset\gamma$ *and* $\models\gamma\supset\beta$. γ *is called the interpolant for the implication* $\alpha\supset\beta$.

Proof. (outline). 1. Assume that all variables of α occur in β. Then we let γ be α. Then evidently holds: if $\models\alpha\supset\beta$, then $\models\alpha\supset\gamma$ and $\models\gamma\supset\beta$.

2. Assume that α and β differ by just a variable p occurring in α but not in β. By hypothesis $\models\alpha\supset\beta$, hence the truth value of $\alpha\supset\beta$ does not depend on the interpretation of p. Now, if q is the variable common to α and β, let α_1 be the formula obtained from α by substituting $q\vee\neg q$ for p, and let α_2 be the formula resulting from α by substituting $q\wedge\neg q$ for p. As can be seen the following holds: $\models\alpha_1\supset\beta$ and $\models\alpha_2\supset\beta$. But $\models\alpha\supset(\alpha_1\vee\alpha_2)$, since α_1 and α_2 result from α by substituting for p alternatively 1 and 0. And from $\models\alpha_1\supset\beta$ and $\models\alpha_2\supset\beta$ follows that $\models(\alpha_1\vee\alpha_2)\supset\beta$ (*via* $\models(p\supset r)\supset[(q\supset r)\supset[(p\vee q)\supset r)]$, $Subst_{PL}$ and MP twice). Hence the formula γ of the theorem is just $\alpha_1\vee\alpha_2$.

3. If α has more than one variable not occurring in β, then we construct α_1 and α_2 as before but by substituting $q\vee\neg q$ and $q\wedge\neg q$, respectively, for each such an occurrence of the respective variable. And reason as above.

Remark 1. This proof can be easily turned into a proof by induction on *n*, the number of propositional variables occurring in α but not in β, in the following way.

Basis. $n = 0$. This is just the first step in the proof above: all the variables of α are in β. Then γ is α itself and the theorem holds.

Induction. $n > 0$. Suppose that the theorem holds for any $k < n$ and show that it holds for $k = n$. Suppose that α contains just *n* propositional variables that do not occur in β. Let p be such a variable, and q be an arbitrary variable occurring in both α and β. As in the preceding proof, construct α_1 and α_2, two formulas obtained from α by substituting $q\vee\neg q$ for p and $q\wedge\neg q$ for p, respectively, and then we have the following derivations:

(1) $\alpha_1 = \alpha(q \vee \neg q / p)$

(2) $\alpha_2 = \alpha(q \wedge \neg q / p)$

(3) $\models \alpha \supset (\alpha_1 \vee \alpha_2)$; (1), (2)

(4) $\models \alpha_1 \supset \beta$; since $\models \alpha \supset \beta$ by hyp.

(5) $\models \alpha_2 \supset \beta$; similar to (4)

(6) $\models (\alpha_1 \vee \alpha_2) \supset \beta$; (4), (5), PL

Since $(\alpha_1 \vee \alpha_2) \supset \beta$ contains $n-1$ variables, then by ind. hyp. the theorem holds for this formula, and then there is a formula γ such that

(7) $\models (\alpha_1 \vee \alpha_2) \supset \gamma$ and

(8) $\models \gamma \supset \beta$

(9) $\models \alpha \supset \gamma$; (3), (7).

(8) and (9) show us that the theorem holds for $k = n$.

Remark 2. Still another proof is that based on the idea of *Craig-consistency* of a finite set S, and on the Theorem 2 of 3.3.3 (below). Let us sketch it.

Definition. *A finite set* S *is Craig-consistent if there is a partition of* S *into* S_1 *and* S_2 (*where* S_1 *and* S_2 *are subsets of* S, $S = S_1 \cup S_2$ *and* $S_1 \cap S_2 = \varnothing$), *such that the implication* $\mathrm{Conj}(S_1)\supset\neg\mathrm{Conj}(S_2)$ *has no interpolant*.

What must be proved is: *if* $\models \alpha\supset\beta$, *then* $\alpha\supset\beta$ *has an interpolant*.

Proof (*reductio*). Suppose that $\models \alpha\supset\beta$ and $\alpha\supset\beta$ has no interpolant. Let S be the set $\{\alpha, \neg\beta\}$ and the partition $S_1 = \{\alpha\}$ and $S_2 = \{\neg\beta\}$. Let us observe that the implication Conj(α)$\supset$ $\neg$Conj($\neg\beta$) has no interpolant (otherwise it would be an interpolant for the formula $\alpha\supset\neg\neg\beta$, i.e., $\alpha\supset\beta$, contrary to our supposition). It follows that S is Craig-consistent and therefore it is satisfiable (by Th. 2 of 3.3.3). I.e., there is an interpretation *int* such that $[\alpha]^{int}=1$ and $[\neg\beta]^{int}=1$, that is $[\beta]^{int}=0$, and then $[\alpha\supset\beta]^{int}=0$. And this does imply that $\not\models\alpha\supset\beta$, contrary to our supposition.

3. PL axiomatized (PL^{ax})

3.1. An axiomatic system

An axiomatic system (or an axiomatic calculus) is a pure syntactical procedure to obtaining some special formulas called theorems. Such systems are of a great variety depending on the chosen axioms, logical connectives or deduction rules. Another difference between them lies in the *expressions* taken as axioms, they can be formulas of L_{PL} or *axiom schemata*, each schema standing for an infinite number of axioms. Let us illustrate this last point.

Let PL^{ax} be the following axiomatic system:[34]

Axioms. Ax1. $\alpha\supset(\beta\supset\alpha)$

Ax2. $(\alpha\supset(\beta\supset\gamma))\supset((\alpha\supset\beta)\supset(\alpha\supset\gamma))$

Ax3. $(\neg\alpha\supset\neg\beta)\supset(\beta\supset\alpha)$.

Rule of inference: *Modus ponens* (MP) $\dfrac{\alpha,\quad \alpha\supset\beta}{\beta}$

Let $PL^{ax(*)}$ be the following axiomatic system:

Axioms. Ax1*. $p\supset(q\supset p)$

Ax2*. $(p\supset(q\supset r))\supset((p\supset q)\supset(p\supset r))$

Ax3*. $(\neg p\supset\neg q)\supset(q\supset p)$

Rules of inference: MP (as above)

Subst. $\dfrac{\vdash\alpha(p_1,\ldots,p_n)}{\vdash\alpha^*(\beta_1/p_1,\ldots,\beta_n/p_n)}$

Definitions

Def$_\vee$. $\alpha\vee\beta =_{df} \neg\alpha\supset\beta$

Def$_\wedge$. $\alpha\wedge\beta =_{df} \neg(\neg\alpha\vee\neg\beta)$

---

[34] This axiomatic system is the Lukasiewicz's simpler version (cf. Lukasiewicz and Tarski [1930]) of Frege's system (cf. Frege [1879]).

Def$_{\equiv}$. $\alpha\equiv\beta =_{df} (\alpha\supset\beta)\wedge(\beta\supset\alpha)$

A simple comparison shows us the following: the axioms of $PL^{ax(*)}$ are formulas of L_{PL}, containing the specified variables. The axioms of PL^{ax},[35] instead, are not formulas of L_{PL}, but axiom *schemata*, that is, indications on how a formula must be in order to be an axiom of PL^{ax}, i.e., such a formula must have the form of Ax1-Ax3, where α, β and γ are *arbitrary* formulas of L_{PL}.[36] As evident, for PL^{ax} the substitution rule is useless.

Examples of proof in PL^{ax}

Th1. $\alpha\supset\alpha$

(1) $[\alpha\supset((\alpha\supset\alpha)\supset\alpha)]\supset[(\alpha\supset(\alpha\supset\alpha))\supset(\alpha\supset\alpha)]$; Ax2

(2) $\alpha\supset((\alpha\supset\alpha)\supset\alpha)$; Ax1

(3) $(\alpha\supset(\alpha\supset\alpha))\supset(\alpha\supset\alpha)$; (1) (2) MP

(4) $\alpha\supset(\alpha\supset\alpha)$; Ax1

(5) $\alpha\supset\alpha$; (3) (4) MP

Example of proof in $PL^{ax(*)}$

Th1*. $p\supset p$

(1) $p\supset(q\supset p)$; Ax1*

(2) $p\supset((q\supset p)\supset p)$; (1) $q\supset p/q$

(3) $(p\supset(q\supset r))\supset((p\supset q)\supset(p\supset r))$; Ax2*

(4) $(p\supset((q\supset p)\supset r))\supset[(p\supset(q\supset p))\supset(p\supset r)]$; (3) $q\supset p/q$

(5) $(p\supset((q\supset p)\supset p))\supset[(p\supset(q\supset p))\supset(p\supset p)]$; (4), p/r

(6) $(p\supset(q\supset p))\supset(p\supset p)$; (2) (5) MP

(7) $p\supset p$; (1) (6) MP.

The symbols occurring after semicolon, to the right of every formula of this proof, indicate the way the respective formula was obtained (or its justification), e.g. by Ax1* (in (1)), by substitution of $q\supset p$ for q in the preceding formula (in (2) and (4)), by MP applied to some preceding formulas in this row (as in (6) or (7)) etc.

The formula in (7) $p\supset p$ is the last formula of this sequence of seven formulas, it is the proved formula. In other words, $p\supset p$ is a *theorem* of

[35] Though we'll also call them axioms, the difference is evident.

[36] Moreover, will be also an axiom schema any schema of the *form* of Ax1-Ax3, e.g. since $\alpha\supset(\beta\supset\alpha)$ is an axiom schema (Ax1), the following expressions will be also axiom schemata: $\beta\supset(\alpha\supset\beta)$, $(\beta\supset\gamma)\supset((\alpha\supset(\beta\supset\gamma))$, $\alpha\supset((\beta\supset\alpha)\supset\alpha)$ etc., since all these expressions have the *form* of Ax1.

$PL^{ax(*)}$. Similarly, the formula in (5), $\alpha \supset \alpha$, in the first proofs, is a theorem[37] of PL^{ax}. Evidently, the proofs in PL^{ax} are simpler than the proofs in $PL^{ax(*)}$ since PL^{ax} has not the substitution rule. And this is not a weakness of PL^{ax} since, as can be argued, the following holds: a formula α of L_{PL} is a theorem of PL^{ax} if and only if α is a theorem of $PL^{ax(*)}$ (argue!). However, this does not hold for any *deduction*, in the sense of Definition 3 and Definition 4 below.

In what follows we take the system PL^{ax} to be the axiomatization of PL. Since the notions "proof" and "deduction" are central to axiomatic considerations, let us introduce them definitionally.

Definition 1. *A proof in* PL^{ax} *is a finite sequence of formulas of* L_{PL} *each one of which is either an axiom of* PL^{ax} *or follows from two preceding formulas of the sequence by one application of* MP.

Definition 2. *A formula* α *of* L_{PL} *is provable in* PL^{ax} *or is a theorem of* PL^{ax} (*symbolic*: $PL^{ax} \vdash \alpha$) *if and only if there is a proof in* PL^{ax} *whose last formula is* α.

Definition 3. *A deduction in* PL^{ax} *of a formula* α *from a set* Γ *of formulas of* L_{PL} *is a* ***finite*** *sequence of formulas of* L_{PL} *each one of which is either an axiom of* PL^{ax} *or a member of* Γ *or follows from two preceding formulas of the sequence by one application of* MP.

Definition 4. *A formula* α *of* L_{PL} *is deductible in* PL^{ax} *from a set* Γ *of formulas* (*symbolic* $\Gamma \vdash \alpha$) *if and only if there is a deduction in* PL^{ax} *whose last formula is* α.

If in $\Gamma \vdash \alpha$ the set $\Gamma = \varnothing$, then we have the case $\varnothing \vdash \alpha$, equivalently $\vdash \alpha$. And then we may correlate these two notions, "proof" and "deduction", in the following way: a *proof* of a formula α in S is a deduction of it from zero premises, and a *deduction* of α in S is a proof of it *from* n ($n \neq 0$) premises.

Remark. The definitions *Def* 1 and *Def* 2 do not provide us with a decision procedure for "theoremhood". For "valid formula of L_{PL}" such a procedure exists (cf. 2.7), but this procedure can be taken as a decision procedure for "theoremhood" only *via* completeness theorem of PL^{ax}, by which "theorem"

[37] Properly speaking, it is a theorem *schema*, since α can be *any* formula of L_{PL}, though we'll call it also theorem.

and "valid formula" are proved to be co-extensive.[38] But for *S* to have a proof procedure means something more (comp. Sect. 3.3.5.2).

The deduction relation[39] has the properties of its relative (semantic consequence) (comp. 2.6). Let us write them down (where Γ, Δ, α, β are arbitrary).

Prop. 1*. $\alpha \vdash \alpha$.

Prop. 2*. If $\Gamma \vdash \alpha$, then $\Gamma \cup \Delta \vdash \alpha$ (written sometimes as $\Gamma, \Delta \vdash \alpha$).

Prop. 3*. If $\Gamma \vdash \alpha$ and $\alpha \vdash \beta$, then $\Gamma \vdash \beta$.

Prop. 4*. If $\Gamma \vdash \alpha$ and $\Gamma \vdash \alpha \supset \beta$, then $\Gamma \vdash \beta$.

Prop. 5*. If $\vdash \alpha$, then $\Gamma \vdash \alpha$.

Similar to the corresponding properties pertaining to semantic consequence, these properties of deduction relation can be proved. Let us take as example Prop 3*.

Suppose, for *reductio*, that (1) $\Gamma \vdash \alpha$, (2) $\alpha \vdash \beta$ and (3) $\Gamma \nvdash \beta$. Then there is a finite set Γ_0 of Γ such that (4) $\Gamma_0 \vdash \alpha$ (from (1) by Def. 3). If Conj(Γ_0) is the conjunction of all formulas of Γ_0, then (5) Conj(Γ_0)$\vdash \alpha$ (by PL[40]) and then (6) $\vdash$Conj(Γ_0)$\supset \alpha$ (from (5) by Deduction Theorem[41]). From (2) it follows (7) $\vdash \alpha \supset \beta$ (by Ded. Th.) and from (6) and (7) we have (8) $\vdash$Conj(Γ_0)$\supset \beta$ (by PL), and therefore (9) Conj(Γ_0)$\vdash \beta$. Whence (10) $\Gamma_0 \vdash \beta$ (from (9), comp. footnote 40, and finally, (11) $\Gamma \vdash \beta$, since $\Gamma_0 \subseteq \Gamma$ and Prop. 2*. But (11) and (3) are contradictory. All these properties will be used in our arguments.

3.2. Basic results on PL^{ax}

3.2.1. Substitution and replacement in PL^{ax}

The following theorems are syntactical counterparts of the corresponding semantic theorems (cf. 2.4).

---

[38] Comp. 3.3 (below).

[39] A.k.a. "deducibility relation" or "syntactic relation".

[40] By *syntactic* counterpart of Theorem (Sect. 2.6). (As we'll see, by soundness and completeness of PL^{ax} (comp. Sect. 3.3) the metalogical symbols "$\models$" and "$\vdash$" are interchangeable).

[41] Comp. Sect. 3.2.2 (below).

3.2.1.1. Substitution Theorem

Substitution Theorem. *Let* $\alpha(p_1,...,p_n)$ *be a formula of* L_{PL} *containing the variables* $p_1,...,p_n$. *Let* $\alpha^*(\beta_1/p_1,...,\beta_n/p_n)$ *be the formula obtained from* α *by substituting arbitrary formulas of* L_{PL}, $\beta_1,...,\beta_n$, *for* $p_1,...,p_n$, *respectively. Then the following holds*:

If $\vdash\alpha$, *then* $\vdash\alpha^*$.

Proof. Assume $\vdash\alpha$. Hence, by 3.1, Def 1 and Def 2, there is a sequence of formulas, Seq, called the proof of α in PL^{ax}. Now in this proof we replace everywhere the variables $p_1,...,p_n$ with corresponding formulas $\beta_1,...,\beta_n$, and what is obtained is a new sequence Seq^* such that all the axioms in Seq are converted into the axioms in Seq^* and all the applications of MP are converted into the new applications of the same rule. Hence Seq^* will be a proof of the formula α^* in PL^{ax}.

3.2.1.2. Replacement Theorem

The idea of replacement in PL is that given in 2.4.2, but this time we shall consider its syntactical counterpart.

Replacement Theorem. *Let* α_β *be a formula of* L_{PL} *containing the subformula* β (*proper or not*) *and let* α_γ *be the formula resulting from* α_β *by replacing one or more occurrences of* β *with* γ. *Then the following holds*:

If $\vdash\beta\equiv\gamma$, then $\vdash\alpha_\beta\equiv\alpha_\gamma$.

Proof. (by induction on n = the complexity of α_β minus the complexity of β). As in the proof of its semantic form (cf. 2.4.2) we confine the proof to the case when only one occurrence of β is replaced by γ, since this process can be repeated.

Basis. $n = 0$. In this case $\alpha_\beta = \beta$, and then if $\vdash\beta\equiv\gamma$, then $\vdash\beta\equiv\gamma$.

Induction. Assume the theorem holds for $k < n$ and show that it holds for n. Then the formula α_β can be expressed in the form α_{β^*}, where $\alpha_{\beta^*} \neq \beta^*$ and $\beta^* = \beta^*_\beta$ or $\beta^* = \beta$. Then $\alpha_\beta = \alpha_{\beta^*_\beta}$, and by induction hypothesis if $\vdash\beta\equiv\gamma$, then $\vdash\beta^*_\beta\equiv\beta^*_\gamma$, since $\mathrm{compl}(\beta^*) - \mathrm{compl}(\beta) < n$. But α_β may have one of the following forms: $\neg\beta^*_\beta$, $\beta^*_\beta\supset\delta$, $\delta\supset\beta^*_\beta$. Let us consider these cases.

1. $\alpha_\beta = \neg\beta^*_\beta$.

(1) $\vdash\beta\equiv\gamma$; hyp.

(2) $\vdash \beta^*_\beta \equiv \beta^*_\gamma$; (1), ind. hyp.

(3) $\vdash (\beta^*_\beta \equiv \beta^*_\gamma) \supset (\neg\beta^*_\beta \equiv \neg\beta^*_\gamma)$; by 3.2.3, Th. 33 (below)

(4) $\vdash \neg\beta^*_\beta \equiv \neg\beta^*_\gamma$; (2), (3), MP

And then, if $\vdash \beta \equiv \gamma$, then $\vdash \alpha_\beta \equiv \alpha_\gamma$.

2a. $\alpha_\beta = \beta^*_\beta \supset \delta$

(1) $\vdash \beta \equiv \gamma$; hyp.

(2) $\vdash \beta^*_\beta \equiv \beta^*_\gamma$; (1), ind. hyp.

(3) $\vdash (\beta^*_\beta \equiv \beta^*_\gamma) \supset [(\beta^*_\beta \supset \delta) \equiv (\beta^*_\gamma \supset \delta)]$; cf. 3.2.3 Th. 40 (below).

(4) $\vdash (\beta^*_\beta \supset \delta) \equiv (\beta^*_\gamma \supset \delta)$; (2), (3), MP.

And then, if $\vdash \beta \equiv \gamma$, then $\vdash \alpha_\beta \equiv \alpha_\gamma$.

2b. $\alpha_{\beta^*} = \delta \supset \beta^*$

(as in the proof of 2a but using in (3) the following theorem
$\vdash (\beta^*_\beta \equiv \beta^*_\gamma) \supset [(\delta \supset \beta^*_\beta) \equiv (\delta \supset \beta^*_\gamma)]$); cf. 3.2.3, Th. 41 (below).

3.2.2. Deduction Theorem in PL^{ax}

The Deduction Theorem[42] is an important tool for proving theorems in mathematical logic. It is based on the following idea: if a formula β is provable from the assumption α, then the implication α⊃β can be asserted. Let us detail it in a more general fashion.

Deduction Theorem. *If* $\alpha_1,...,\alpha_m \vdash \beta$, *then* $\alpha_1,...,\alpha_{m-1} \vdash \alpha_m \supset \beta$.

I.e., if the formula β is provable from the assumptions (hypotheses) $\alpha_1,...,\alpha_m$, then the formula $\alpha_m \supset \beta$ is provable from the remaining assumptions.

Proof. Assume that $\alpha_1,...,\alpha_m \vdash \beta$, i.e., there is a deduction $Ded = \beta_1,...,\beta_n$ (where $\beta_n = \beta$) of the formula β from $\alpha_1,...,\alpha_m$. We transform this deduction in a deduction Ded^* of the implication $\alpha_m \supset \beta$ from $\alpha_1,...,\alpha_{m-1}$ in the following way. Construct a sequence of formulas, *Seq*, by prefixing each formula β_i ($1 \le i \le n$) from *Ded* with "$\alpha_m \supset$", i.e.,

[42] Also called Herbrand Theorem; cf. J. Herbrand [1930].

$Seq = (\alpha_m \supset \beta_1, \alpha_m \supset \beta_2, ..., \alpha_m \supset \beta_n)$.

Seq is not the deduction Ded^*, we are seeking, but it can be transformed in Ded^* if we *justify* all the formulas in *Seq*. And this fact depends on what the formulas β_i of *Ded* may be: (1) the assumption α_m, (2) an assumption $\alpha_k \neq \alpha_m$, (3) an axiom or (4) it results from two preceding formulas in *Ded* by one application of MP.

1. $\beta_i = \alpha_m$. Then the corresponding formula in *Seq* is $\alpha_m \supset \beta_i$, i.e., $\alpha_m \supset \alpha_m$. We replace this formula in *Seq* with the whole proof of it in PA^{ax} (according to 3.1, Th. 1).
2. $\beta_i = \alpha_k$ $(k \neq m)$. Then $\alpha_m \supset \alpha_k \in Seq$, and therefore replace $\alpha_m \supset \alpha_k$ in *Seq* with its justification in PA^{ax}, i.e.,

   (1) α_k; hyp. (note that α_k is also an assumption in Ded^*)

   (2) $\alpha_k \supset (\alpha_m \supset \alpha_k)$; Ax1

   (3) $\alpha_m \supset \alpha_k$; (1), (2), MP
3. β_i is an axiom. Then $\alpha_m \supset \beta_i \in Seq$, and therefore we replace it with its justification:

   (1) β_i; axiom

   (2) $\beta_i \supset (\alpha_m \supset \beta_i)$; Ax1

   (3) $\alpha_m \supset \beta_i$; (1), (2), MP
4. β_i results in *Ded* by one application of MP from formulas β_k and $\beta_k \supset \beta_i$. Then the corresponding formulas in *Seq* are $\alpha_m \supset \beta_k$, $\alpha_m \supset (\beta_k \supset \beta_i)$ and $\alpha_m \supset \beta_i$. Then replace $\alpha_m \supset \beta_i \in Seq$ with its justification, i.e.,

   (1) $\alpha_m \supset \beta_k \in Seq$

   (2) $\alpha_m \supset (\beta_k \supset \beta_i) \in Seq$

   (3) $(\alpha_m \supset (\beta_k \supset \beta_i)) \supset ((\alpha_m \supset \beta_k) \supset (\alpha_m \supset \beta_i))$; Ax2

   (4) $(\alpha_m \supset \beta_k) \supset (\alpha_m \supset \beta_i)$; (2), (3), MP

   (5) $\alpha_m \supset \beta_i$; (1), (4), MP.

By proceeding in this way we obtained a deduction of the formula $\alpha_m \supset \beta$ from the assumptions $\alpha_1, ..., \alpha_{m-1}$.

As can be observed, the theoretical resources used in the proof of Deduction Theorem are Ax1, Ax2, Th 1 and MP.

Remark. For the system $PL^{ax(*)}$ Deduction Theorem does not hold generally, as the following simple example shows:

(1) p; hyp

(2) q; (1) Subst q/p.

Hence (3) $p \vdash q$.

But from this deduction in $PL^{ax(*)}$ we cannot deduce, by one application of Deduction Theorem that $\vdash p \supset q$. More generally, from the fact that in $PL^{ax(*)}$ we have a derivation $\Gamma \vdash \beta$, we do not conclude that the same derivation holds in PL^{ax}. Otherwise, from $p \vdash q$ in $PL^{ax(*)}$ we conclude that $p \vdash q$ in PL^{ax}, and then by Deduction Theorem $\vdash p \supset q$ in PL^{ax}, whence, by Substitution Theorem (cf. 3.2.1.1) $\vdash (\alpha \supset \alpha) \supset \beta$. And from this, by 3.1 Th1 and MP, we obtain $\vdash \beta$, where β is arbitrary. This means that PL^{ax} is inconsistent, but as we shall see[43] this is not the case.

Corollary. *If* $\alpha_1,...,\alpha_n \vdash \beta$, *then* $\vdash \alpha_1 \supset (\alpha_2 \supset ... \supset (\alpha_n \supset \beta)...)$.

Substitution Theorem, Replacement Theorem and Deduction Theorem are usually called *metatheorems*, since their proofs use theoretical resources transcending the formal resources of PL^{ax} itself. In general, the importance of such metatheorems consists of the fact that they do guarantee *the existence of some derivations*, without displaying them explicitly. But any use of such a metatheorem in proving a theorem of PL^{ax} does *not* make the proof of it a *meta*-theoretical one. Since if a proof of such a formula of L_{PL} is sought, then it can be constructed from the proof of the corresponding metatheorem.

By Deduction Theorem, for example, if β is deducible from the assumptions $\alpha_1,...,\alpha_m$, then $\alpha_m \supset \beta$ is deducible from the assumptions $\alpha_1,...,\alpha_{m-1}$. And this can be proved by converting the given deduction *Ded* of β from $\alpha_1,...,\alpha_m$ into a deduction Ded^* of $\alpha_m \supset \beta$ from the $\alpha_1,...,\alpha_{m-1}$, in the way explained above. Now let us take an example of a formula of L_{PL}, proved using Deduction Theorem, and then show how from the proof of Deduction Theorem the proof of it in PL^{ax} can be effectively displayed.

Example. Th2. $\neg\alpha \supset (\alpha \supset \beta)$

Using Deduction Theorem the proof of Th2 is:

(1) $\neg\alpha$; hyp

[43] Cf. 3.3.1, Consistency Theorem.

(2) $\neg\alpha\supset(\neg\beta\supset\neg\alpha)$; Ax1

(3) $\neg\beta\supset\neg\alpha$; (1) (2) MP

(4) $(\neg\beta\supset\neg\alpha)\supset(\alpha\supset\beta)$; Ax3

(5) $\alpha\supset\beta$; (3) (4) MP

Hence $\neg\alpha\vdash\alpha\supset\beta$; (1)-(5), whence $\vdash\neg\alpha\supset(\alpha\supset\beta)$, by Deduction Theorem.

Let us now construct a proof in $\mathrm{PL}^{\mathrm{ax}}$ of Th2, from the proof of *Deduction Theorem* but *without* any use of Deduction Theorem.

The preceding sequence of formulas (1)-(5) contains the formulas of the deduction *Ded* of $\alpha\supset\beta$ from $\neg\alpha$. i.e., in our symbolism it is the formulas $\beta_1,\ldots,\beta_5$. Replace these formulas with the formulas $\neg\alpha\supset\beta_i$ $(1\leq i\leq 5)$ and obtain the following sequence *Seq*:

1'. $\neg\alpha\supset\neg\alpha$

2'. $\neg\alpha\supset[\neg\alpha\supset(\neg\beta\supset\neg\alpha)]$

3'. $\neg\alpha\supset(\neg\beta\supset\neg\alpha)$

4'. $\neg\alpha\supset[(\neg\beta\supset\neg\alpha)\supset(\alpha\supset\beta)]$

5'. $\neg\alpha\supset(\alpha\supset\beta)$

This is a sequence of formulas but not a deduction. In order to obtain a deduction *Ded* * we have to *insert* in *Seq* [44]all the formulas justifying the members of *Seq*, depending on what the formula β_i, $1\leq i\leq 5$, in *Ded* is.

β_1 *is the formula* $\neg\alpha$. Hence, we have the case 1 in the proof of Deduction Theorem, and then we insert before the formula 1' all the formulas from which it follows in $\mathrm{PL}^{\mathrm{ax}}$, i.e., just Th1 (of 3.1).

β_2 *is* Ax1. Here we have the case 3 in the proof of Deduction Theorem. Then we insert before 2' in *Seq* all the formulas from which it follows, that is $\neg\alpha\supset(\neg\beta\supset\neg\alpha)$ (Ax1) and $[\neg\alpha\supset(\neg\beta\supset\neg\alpha)]\supset$ $\supset[\neg\alpha\supset(\neg\alpha\supset(\neg\beta\supset\neg\alpha))]$ (Ax1), whence, by MP, 2' follows.

β_3 *follows from the formulas* (1) *and* (2) in *Ded* by MP. This is the case 4 in Deduction Theorem, and then before the formula 3' in *Seq* we set the following formulas: $\neg\alpha\supset\neg\alpha$, $\neg\alpha\supset(\neg\alpha\supset(\neg\beta\supset\neg\alpha))$ and Ax2: $\{\neg\alpha\supset[\neg\alpha\supset(\neg\beta\supset\neg\alpha)]\}\supset[(\neg\alpha\supset\neg\alpha)\supset(\neg\alpha\supset\neg\beta\supset\neg\alpha))]$.

By two applications of MP we get 3'.

β_4 *is* Ax3. Again, we have the case 3 in the proof of Deduction Theorem, and proceed similarly, i.e., we insert before the formula 4' in *Seq* all the formulas from which it follows: Ax3 and $\beta_4\supset(\neg\alpha\supset\beta_4)$: $[(\neg\beta\supset\neg\alpha)\supset\alpha\supset\beta)]\supset[\neg\alpha\supset((\neg\beta\supset\neg\alpha)\supset(\alpha\supset\beta))]$, (Ax1) and by one

[44] Or, equivalently, as we proceed above, replace each formula in *Seq* with its *justification*.

application of MP we obtain 4'.

β_5 follows in *Ded* from (3) and (4) by MP. And again we have the case 4. Hence before the formula 5' in *Seq* we insert the following formulas: $\neg\alpha\supset(\neg\beta\supset\neg\alpha)$, $\neg\alpha\supset[(\neg\beta\supset\neg\alpha)\supset(\alpha\supset\beta)]$ and Ax2: $\{\neg\alpha\supset[(\neg\beta\supset\neg\alpha)\supset(\alpha\supset\beta)]\}\supset\{[\neg\alpha\supset(\neg\beta\supset\neg\alpha)]\supset[\neg\alpha\supset(\alpha\supset\beta)]\}$, from which by two applications of MP the formula 5' in *Seq* is obtained.

Finally, if we string all the formulas obtained by inserting in *Seq* the formulas mentioned above, then we obtain exactly the deduction Ded^* of the formula $\neg\alpha\supset(\alpha\supset\beta)$, that is, a proof of Th2 in PL^{ax}, constructed according to the proof of Deduction Theorem, but not using it.

3.2.3. Proofs in PL^{ax}

Th1. $\alpha\supset\alpha$; cf. 3.1, Th1.

Th2. $\neg\alpha\supset(\alpha\supset\beta)$; cf. 3.2.2, Example.

Th3. $[\alpha\supset(\alpha\supset\beta)]\supset(\alpha\supset\beta)$.

(1) $\alpha\supset(\alpha\supset\beta)$; hyp
(2) α; hyp
(3) $\alpha\supset\beta$; (1) (2) MP
(4) β; (2) (3) MP.

Hence $\alpha\supset(\alpha\supset\beta), \alpha\vdash\beta$, and then $\vdash[\alpha\supset(\alpha\supset\beta)]\supset(\alpha\supset\beta)$; by 3.2.2 Coroll.

Th4. $\alpha\supset(\neg\alpha\supset\beta)$

(1) α; hyp
(2) $\neg\alpha$; hyp
(3) $\alpha\supset\beta$; (2) Th2
(4) β; (1) (3)

Hence $\alpha,\neg\alpha\vdash\beta$, whence $\vdash\alpha\supset(\neg\alpha\supset\beta)$; 3.2.2 Coroll.

Th5. $(\alpha\supset\beta)\supset[(\beta\supset\gamma)\supset(\alpha\supset\gamma)]$

(1) $\alpha\supset\beta$; hyp
(2) $\beta\supset\gamma$; hyp
(3) α; hyp
(4) β; (1) (3) MP
(5) γ; (2) (4) MP

Hence $\alpha\supset\beta, \beta\supset\gamma, \alpha\vdash\gamma$, whence $\vdash(\alpha\supset\beta)\supset[(\beta\supset\gamma)\supset(\alpha\supset\gamma)]$, by 3.2.2 Coroll.

Th6. $\neg\neg\alpha\supset\alpha$

(1) $\neg\neg\alpha$; hyp

(2) $\neg\neg\alpha\supset(\neg\alpha\supset\neg\neg\neg\alpha)$; Th2

(3) $\neg\alpha\supset\neg\neg\neg\alpha$; (1) (2) MP

(4) $(\neg\alpha\supset\neg\neg\neg\alpha)\supset(\neg\neg\alpha\supset\alpha)$; Ax3

(5) $\neg\neg\alpha\supset\alpha$; (3) (4) MP

Th7. $\alpha\supset\neg\neg\alpha$ (exercise)

Th8. $(\alpha\supset\beta)\supset(\neg\beta\supset\neg\alpha)$

(1) $\alpha\supset\beta$; hyp

(2) $\neg\neg\alpha\supset\alpha$; Th6

(3) $\neg\neg\alpha\supset\beta$; (1) (2) Th5 MP

(4) $\beta\supset\neg\neg\beta$; Th7

(5) $\neg\neg\alpha\supset\neg\neg\beta$; (3) (4) Th5, MP

(6) $(\neg\neg\alpha\supset\neg\neg\beta)\supset(\neg\beta\supset\neg\alpha)$; Ax3

(7) $\neg\beta\supset\neg\alpha$; (5) (6) MP

Hence $\alpha\supset\beta\vdash\neg\beta\supset\neg\alpha$; whence $\vdash(\alpha\supset\beta)\supset(\neg\beta\supset\neg\alpha)$; Ded Th.

Th9. $(\alpha\supset\neg\beta)\supset(\beta\supset\neg\alpha)$ (exercise)

Th10. $(\neg\alpha\supset\beta)\supset(\neg\beta\supset\alpha)$ (exercise)

Th11. $\alpha\supset(\neg\beta\supset\neg(\alpha\supset\beta))$

(1) $\alpha\supset[(\alpha\supset\beta)\supset\beta]$; by MP and 3.2.2 Coroll.

(2) $[(\alpha\supset\beta)\supset\beta]\supset((\neg\beta\supset\neg(\alpha\supset\beta))$; Th8

(3) $\alpha\supset(\neg\beta\supset\neg(\alpha\supset\beta))$; (1) (2) Th5, MP

Th12. $(\alpha\supset\beta)\supset((\neg\alpha\supset\beta)\supset\beta)$ (exercise)

Th13. $(\alpha\supset(\beta\supset\gamma))\supset(\beta\supset(\alpha\supset\gamma))$

(1) $\alpha\supset(\beta\supset\gamma)$; hyp

(2) β; hyp

(3) $(\alpha\supset(\beta\supset\gamma))\supset((\alpha\supset\beta)\supset(\alpha\supset\gamma))$; Ax2

(4) $(\alpha\supset\beta)\supset(\alpha\supset\gamma)$; (1) (3) MP

(5) $\beta\supset(\alpha\supset\beta)$; Ax1

(6) $\alpha\supset\beta$; (2) (5) MP

(7) $\alpha\supset\gamma$; (4) (6) MP

Hence $\alpha\supset(\beta\supset\gamma), \beta\vdash\alpha\supset\gamma$, whence $\vdash(\alpha\supset(\beta\supset\gamma))\supset(\beta\supset(\alpha\supset\gamma))$; 3.2.2 Coroll.

Th14. $(\alpha\supset\beta)\supset((\gamma\vee\alpha)\supset(\gamma\vee\beta))$

(1) $\alpha\supset\beta$; hyp

(2) $(\alpha\supset\beta)\supset(\neg\gamma\supset(\alpha\supset\beta))$; Ax1

(3) $\neg\gamma\supset(\alpha\supset\beta)$; (1), (2) MP

(4) $(\neg\gamma\supset(\alpha\supset\beta))\supset((\neg\gamma\supset\alpha)\supset(\neg\gamma\supset\beta))$; Ax2

(5) $(\neg\gamma\supset\alpha)\supset(\neg\gamma\supset\beta)$; (3), (4) MP

(6) $(\gamma\vee\alpha)\supset(\gamma\vee\beta)$; (5) 3.1 Def$_\vee$

Hence $\alpha\supset\beta\vdash(\gamma\vee\alpha)\supset(\gamma\vee\beta)$; whence $\vdash(\alpha\supset\beta)\supset((\gamma\vee\alpha)\supset(\gamma\vee\beta))$; Ded. Th.

Th15. $(\alpha\supset\neg\alpha)\supset\neg\alpha$

(By Th4: $\alpha\supset(\neg\alpha\supset\neg(\beta\supset\beta))$, Ax2, Th9,Th5,Th13,Th1,MP) (exercise)

Th16. $(\neg\alpha\supset\alpha)\supset\alpha$ (exercise)

Th17. $(\alpha\vee\alpha)\supset\alpha$; Th16, Def$_\vee$

Th18. $\alpha\supset(\alpha\vee\beta)$ (exercise)

Th19. $(\alpha\vee\beta)\supset(\beta\vee\alpha)$

(1) $\alpha\vee\beta$; hyp

(2) $\neg\alpha\supset\beta$; (1), 3.1 Def$_\vee$

(3) $\neg\beta\supset\alpha$; (2), Th8 MP

(4) $\beta\vee\alpha$; (3), 3.1 Def$_\vee$

Th20. $(\alpha\wedge\beta)\supset\alpha$ (exercise)

Th21. $(\alpha\wedge\beta)\supset\beta$ (exercise)

Th22. $\alpha\vee\neg\alpha$ (exercise)

Th23. $(\alpha\vee(\beta\vee\gamma))\supset(\beta\vee(\alpha\vee\gamma))$

(1) $\alpha\vee(\beta\vee\gamma)$; hyp

(2) $\neg\alpha\supset(\neg\beta\supset\gamma)$; (1) 3.1 Def$_\vee$

(3) $\neg\beta\supset(\neg\alpha\supset\gamma)$; (2) Th13, MP

(4) $\beta\vee(\alpha\vee\gamma)$; (3) 3.1 Def$_\vee$

Th24. $(\alpha\vee(\beta\vee\gamma))\supset((\alpha\vee\beta)\vee\gamma)$

(1) $\alpha\vee(\beta\vee\gamma)$; hyp

(2) $\neg\alpha\supset(\beta\vee\gamma)$; 3.1 Def$_\vee$

(3) $(\beta\vee\gamma)\supset(\gamma\vee\beta)$; Th19

(4) $\neg\alpha\supset(\gamma\vee\beta)$; (2) (3), Th5 MP

(5) $\alpha\vee(\gamma\vee\beta)$; (4), 3.1 Def$_\vee$

(6) $\gamma\vee(\alpha\vee\beta)$; (5), Th23, MP

(7) $(\alpha\vee\beta)\vee\gamma$; Th19, (6), MP

Th25. $((\alpha\vee\beta)\vee\gamma)\supset(\alpha\vee(\beta\vee\gamma))$ (exercise)

Th26. $\alpha,\beta\vdash\alpha\wedge\beta$

(1) $\neg(\neg\alpha\vee\neg\beta)\supset\neg(\neg\alpha\vee\neg\beta)$; Th1

(2) $(\neg\alpha\vee\neg\beta)\vee\neg(\neg\alpha\vee\neg\beta)$; (1) 3.1 Def$_\vee$

(3) $\neg\alpha\vee(\neg\beta\vee\neg(\neg\alpha\vee\neg\beta))$; (2) Th25

(4) $\neg\alpha\vee(\neg\beta\vee(\alpha\wedge\beta))$; (3), 3.1 Def$_\wedge$

(5) $\alpha\supset(\beta\supset(\alpha\wedge\beta))$; (4) 3.1 Def$_\vee$

(6) α; hyp

(7) $\beta\supset(\alpha\wedge\beta)$; (5) (6) MP

(8) β; hyp

(9) $\alpha\wedge\beta$; (7) (8) MP

Hence $\alpha,\beta\vdash\alpha\wedge\beta$.

Th27. $(\alpha\supset(\beta\supset\gamma))\supset((\alpha\wedge\beta)\supset\gamma)$

(1) $\alpha\supset(\beta\supset\gamma)$; hyp

(2) $\alpha\wedge\beta$; hyp

(3) α; (2) Th20, MP

(4) β; (2) Th21, MP

(5) $\beta\supset\gamma$; (1) (3) MP

(6) γ; (5) (4) MP

Hence $\alpha\supset(\beta\supset\gamma)$, $\alpha\wedge\beta\vdash\gamma$, and then $\vdash(\alpha\supset(\beta\supset\gamma))\supset((\alpha\wedge\beta)\supset\gamma)$, by 3.2.2 Coroll.

Th28. $((\alpha\wedge\beta)\supset\gamma)\supset(\alpha\supset(\beta\supset\gamma))$ (exercise)

Th29. $(\alpha\supset(\beta\supset\gamma))\equiv((\alpha\wedge\beta)\supset\gamma)$

(by Th27, Th28, Th26, 3.1 Def$_\equiv$)

Th30. $((\alpha\equiv\beta)\wedge(\beta\equiv\gamma))\supset(\alpha\equiv\gamma)$ (exercise)

Th31. $\alpha\equiv\beta$, $\beta\equiv\gamma\vdash\alpha\equiv\gamma$ (exercise)

Th32. $\alpha\equiv\neg\neg\alpha$

(by Th6, Th7, Th26, 3.1 Def$_\equiv$)

Th33. $(\alpha\equiv\beta)\supset(\neg\alpha\equiv\neg\beta)$

(1) $\alpha\equiv\beta$; hyp

(2) $(\alpha\supset\beta)\wedge(\beta\supset\alpha)$; (1) 3.1 Def$_\equiv$

(3) $\alpha\supset\beta$; (2), Th20, MP

(4) $\beta\supset\alpha$; (2) Th21, MP

(5) $\neg\beta\supset\neg\alpha$; (3) Th8, MP

(6) $\neg\alpha\supset\neg\beta$; (4) Th8, MP

(7) $(\neg\alpha\supset\neg\beta)\wedge(\neg\beta\supset\neg\alpha)$; (5) (6) Th26

(8) $\neg\alpha\equiv\neg\beta$; (7) 3.1 Def$_\equiv$

(9) $\alpha\equiv\beta\vdash\neg\alpha\equiv\neg\beta$

(10) $(\alpha\equiv\beta)\supset(\neg\alpha\equiv\neg\beta)$; 9 Ded. Th.

Th34. $(\neg\alpha\equiv\neg\beta)\supset(\alpha\equiv\beta)$ (exercise)

Th35. $(\alpha\equiv\beta)\equiv(\neg\alpha\equiv\neg\beta)$

(by Th33, Th34, 3.1 Def$_\equiv$)

Th36. $\neg(\alpha\wedge\neg\alpha)$

(1) $\neg\alpha\vee\neg\neg\alpha$; Th22

(2) $(\neg\alpha\vee\neg\neg\alpha)\supset\neg\neg(\neg\alpha\vee\neg\neg\alpha)$; Th7

(3) $\neg\neg(\neg\alpha\vee\neg\neg\alpha)$; (1) (2) MP

(4) $\neg(\alpha\wedge\neg\alpha)$; (3), 3.1 Def$_\wedge$, Repl. Th.

Th37. $((\alpha\supset\beta)\wedge\neg\beta)\supset\neg\alpha$

(1) $(\alpha\supset\beta)\wedge\neg\beta$; hyp

(2) $\alpha\supset\beta$; (1) Th20, MP

(3) $\neg\beta$; (1) Th21, MP

(4) $\neg\beta\supset\neg\alpha$; (2) Th8, MP

(5) $\neg\alpha$; (3) (4) MP

Hence $(\alpha\supset\beta)\wedge\neg\beta\vdash\neg\alpha$, and then $\vdash((\alpha\supset\beta)\wedge\neg\beta)\supset\neg\alpha$ by Ded. Th.

Th38. $\neg(\alpha\vee\beta)\equiv(\neg\alpha\wedge\neg\beta)$

(1) $\neg(\alpha\vee\beta)\supset\neg(\alpha\vee\beta)$; Th1

(2) $\neg(\alpha\vee\beta)\supset\neg(\neg\neg\alpha\vee\neg\neg\beta)$; (1) Th32, Repl. Th.

(3) $\neg(\alpha\vee\beta)\supset(\neg\alpha\wedge\neg\beta)$; (2) Def$_\wedge$

(4) $(\neg\alpha\wedge\neg\beta)\supset(\neg\alpha\wedge\neg\beta)$; Th1

(5) $(\neg\alpha\wedge\neg\beta)\supset\neg(\neg\neg\alpha\vee\neg\neg\beta)$; (4), 3.1 Def$_\wedge$

(6) $(\neg\alpha\wedge\neg\beta)\supset\neg(\alpha\vee\beta)$; (5), Th32, Repl.Th; whence Th38, by (3) and(6), Def$_\equiv$.

Th39. $\neg(\alpha\supset\beta)\equiv(\alpha\wedge\neg\beta)$

(1) $\neg(\alpha\supset\beta)\equiv\neg(\neg\alpha\vee\beta)$; 3.1 Def$_\vee$

(2) $\neg(\neg\alpha\vee\beta)\equiv(\neg\neg\alpha\wedge\neg\beta)$; Th38

(3) $(\neg\neg\alpha\wedge\neg\beta)\equiv(\alpha\wedge\neg\beta)$; Th32, Repl. Th.

(4) $\neg(\alpha\supset\beta)\equiv(\alpha\wedge\neg\beta)$; (1)-(3), Th. 31, Ded. Th.

Th40. $(\alpha\equiv\beta)\supset((\alpha\supset\gamma)\equiv(\beta\supset\gamma))$

(1) $\alpha\equiv\beta$; hyp

(2) $\alpha\supset\beta$; (1) 3.1 Def$_\equiv$, Th20, MP

(3) $\beta\supset\alpha$; (1) 3.1 Def$_\equiv$, Th21, MP

(4) $(\alpha\supset\beta)\supset((\beta\supset\gamma)\supset(\alpha\supset\gamma))$; Th5

(5) $(\beta\supset\gamma)\supset(\alpha\supset\gamma)$; (2) (4) MP

(6) $(\beta\supset\alpha)\supset((\alpha\supset\gamma)\supset(\beta\supset\gamma))$; Th5

(7) $(\alpha\supset\gamma)\supset(\beta\supset\gamma)$; (3) (6) MP

(8) $((\alpha\supset\gamma)\supset(\beta\supset\gamma))\wedge((\beta\supset\gamma)\supset(\alpha\supset\gamma)))$; (5) (7), Th26

(9) $(\alpha\supset\gamma)\equiv(\beta\supset\gamma)$; 8, 3.1 Def$_\equiv$

Hence $(\alpha\equiv\beta)\vdash(\alpha\supset\gamma)\equiv(\beta\supset\gamma)$, whence $\vdash(\alpha\equiv\beta)\supset((\alpha\supset\gamma)\equiv(\beta\supset\gamma))$; Ded Th.

Th41. $(\alpha\equiv\beta)\supset((\gamma\supset\alpha)\equiv(\gamma\supset\beta))$

(as Th.40 but in the steps (4) and (6) use Th14 and 3.1 Def$_\vee$).

Th42. $(\alpha\supset\beta)\supset[(\alpha\supset\gamma)\supset(\alpha\supset(\beta\wedge\gamma))]$

(1) $\alpha\supset\beta$; hyp

(2) $\alpha\supset\gamma$; hyp

(3) α; hyp

(4) β; (1) (3) MP

(5) γ; (2) (3) MP

(6) $\beta\wedge\gamma$; (4) (5) Th26

Hence $\alpha\supset\beta, \alpha\supset\gamma, \alpha\vdash\beta\wedge\gamma$ and then $\vdash(\alpha\supset\beta)\supset[(\alpha\supset\gamma)\supset(\alpha\supset(\beta\wedge\gamma))]$; by 3.2.2 Corollary.

Th43. $(\alpha\supset\beta)\supset[(\alpha\supset\neg\beta)\supset\neg\alpha]$

(1) $\alpha\supset\beta$; hyp

(2) $\alpha\supset\neg\beta$; hyp

(3) $(\alpha\supset\beta)\supset[(\alpha\supset\neg\beta)\supset(\alpha\supset(\beta\wedge\neg\beta))]$; Th42

(4) $\alpha\supset(\beta\wedge\neg\beta)$; (1) (2) (3) MP twice

(5) $\neg(\beta\wedge\neg\beta)\supset\neg\alpha$; (4) Th8, MP

(6) $\neg(\beta\wedge\neg\beta)$; Th36

(7) $\neg\alpha$; (5) (6) MP

Hence $\alpha\supset\beta, \alpha\supset\neg\beta\vdash\neg\alpha$, and then $\vdash(\alpha\supset\beta)\supset[(\alpha\supset\neg\beta)\supset\neg\alpha]$; by 3.2.2 Corollary.

3.3. Soundness and completeness of PL^{ax}

3.3.1. Soundness of PL^{ax}

Definition 1. *An axiomatic system S is called sound if every formula provable in S is a valid formula of* L_{PL}.

Soundness Theorem for PL^{ax} (weak form). *If* $\vdash\alpha$, *then* $\models\alpha$.

Proof. All we have to prove is that the axioms of PL^{ax} are valid formulas of L_{PL} and that MP preserves in conclusion the validity of premises. Using any decision procedure in PL (cf. 2.7) we can show that Ax1-Ax3 are valid formulas of L_{PL}. And by 2.1.2 Modus Ponens, the rule MP preserves the validity. Hence every theorem of PL^{ax} is valid.

Definition 2. *An axiomatic system S is called sound in a stronger sense if the following holds*: *If* $\alpha_1,...,\alpha_n\vdash\beta$, *then* $\alpha_1,...,\alpha_n\models\beta$.

Remark 1. As we mentioned in 3.1, for the systems PL^{ax} and $PL^{ax(*)}$ the following holds $PL^{ax}\vdash\alpha$ iff $PL^{ax(*)}\vdash\alpha$. It follows that the soundness theorem also holds for $PL^{ax(*)}$.[45] But the stronger form of soundness does not hold for $PL^{ax(*)}$, for the simple reason that $Subst_{PL}$ (comp. 2.4.1) preserves in its applications the *validity* of a given formula α, but does *not* preserve the simple[46] truth of α. This means the following thing. By its syntactic counterpart (cf. 3.2.1.1), if p is given, then by substitution q/p, the formula q is obtained, i.e., $p \vdash q$ is a deduction in $PL^{ax(*)}$. But from this deduction we cannot conclude that $p \models q$, since we have no guarantee that whenever p is true in some interpretation, the sentence q will also be true.[47]

As we saw above (Sect. 2.6 and 3.1), often we are interested not only in which formulas are provable (valid) but also in which formulas *follow* sintactic and semantic from which formulas. And this fact does imply to consider also a *stronger* form of soundness theorem,[48] given by the following result.

Soundness Theorem for PL^{ax} (stronger form). *If* $\Gamma \vdash \alpha$, *then* $\Gamma \models \alpha$.

Proof. Assume that (1) $\Gamma \vdash \alpha$. Then there is a finite set Γ_0 of Γ such that (2) $\Gamma_0 \vdash \alpha$ (by def.). And then (3) $Conj(\Gamma_0) \vdash \alpha$ (by PL). Whence (4) $\vdash Conj(\Gamma_0) \supset \alpha$ (by Ded. Th.). From (4), by Soundness Theorem (weak form), it follows (5) $\models Conj(\Gamma_0) \supset \alpha$, and therefore (6) $Conj(\Gamma_0) \models \alpha$ (by Normality Theorem, comp. Sect. 2.6), and then (7) $\Gamma_0 \models \alpha$ (by PL). Hence (8) $\Gamma \models \alpha$ (by Prop. 2*, Sect. 3.1).

Definition. *A system S is consistent if there is no formula* α *such that* $\vdash \alpha$ *and* $\vdash \neg\alpha$.

Consistency Theorem. PL^{ax} *is consistent.*

Proof. The consistency of PL^{ax} follows from its soundness. Since by soundness if $\vdash \alpha$, then $\models \alpha$, but $\neg\alpha$ cannot be valid, hence $\neg\alpha$ is not provable.

Remark 2. Other definitions of consistency can also be given: e.g. a) A system *S* is consistent if not every formula is provable in *S*, and b) A system *S* is consistent if there is a formula p (or $p \wedge \neg p$) not provable in *S*. For PL^{ax} all the three definitions given are equivalent, since if in *S* the formula p were provable (against b), then by Substitution Theorem (comp. 3.2.1.1) any

[45] It follows that $PL^{ax(*)}$ is also consistent (in the sense of definition given below).

[46] I.e. the truth of α in some interpretation.

[47] Comp. also 3.2.2 Remark.

[48] This is also the case with completeness theorem (comp. 3.3.2 below).

formula would be provable (against a) and then $\vdash \alpha$ and $\vdash \neg\alpha$ will also be provable (contrary to Definition). On the other hand, if in PL^{ax} we had $\vdash \alpha$ and $\vdash \neg\alpha$ (contrary to Definition), then by 3.2.3 Th2: $\neg\alpha\supset(\alpha\supset\beta)$ and MP any formula β would be provable (against a) and then p will also be provable (against b).

3.3.2. Completeness of PL^{ax} (Henkin-style[49])

As we saw above, PL^{ax} is a sound system of PL, i.e., every theorem of PL^{ax} is a valid formula of L_{PL}. Now we are interested in the converse of this assertion and as we shall see it also holds for PL^{ax}.

Definition 1. *An axiomatic system S is called complete if every valid formula of* L_{PL} *is a theorem of S. S is called complete in a stronger sense if the following holds*: *if* $\Gamma\models\alpha$, *then* $\Gamma\vdash\alpha$.

Completeness Theorem (stronger form). *If in* PL $\Gamma\models\alpha$, *then in* PL^{ax} $\Gamma\vdash\alpha$.

The proof of this theorem requires some notions and corresponding lemmas.

Definition 2. *A set* Γ *of formulas of* L_{PL} *is consistent if there is no formula* β *of* L_{PL} *such that* $\Gamma\vdash\beta$ *and* $\Gamma\vdash\neg\beta$, *otherwise it is inconsistent*.

Since the relation "$\vdash$" is finitist[50], $\Gamma\vdash\beta$ means that there are formulas $\alpha_1,\dots,\alpha_n\in\Gamma$ such that $\alpha_1,\dots,\alpha_n\vdash\beta$.

Lemma 1. $\Gamma\vdash\alpha$ *if and only if* $\Gamma\cup\{\neg\alpha\}$ *is inconsistent*.

To prove this lemma means to prove the following conditionals:

(a) If $\Gamma\vdash\alpha$, then $\Gamma\cup\{\neg\alpha\}$ is inconsistent.

(b) If $\Gamma\cup\{\neg\alpha\}$ is inconsistent, then $\Gamma\vdash\alpha$.

Proof. (a) Suppose that (1) $\Gamma\vdash\alpha$. But (2) $\vdash\alpha\supset(\neg\alpha\supset\beta)$ (by Th. 4, Sect. 3.2.3), and then (3) $\Gamma\vdash\neg\alpha\supset\beta$ (from (1) and (2) *via* Prop. 4* and 5* of Sect. 3.1), where β is an arbitrary formula of L_{PL}. Whence (4) $\Gamma,\neg\alpha\vdash\beta$ (by the converse of Ded. Th.). Since β is arbitrary (i.e., take $\beta=\gamma\wedge\neg\gamma$), it follows

[49] Cf. L. Henkin [1949]. In fact, there is a variety of such proofs of completeness, e.g. E. Post [1921], L. Kalmár [1935] (see 3.3.4 below), W.V.O. Quine [1938], K. Schütte [1956], A.R. Anderson and N. Belnap [1959], S.C. Kleene [1967], Ch. VI, R.R. Stoll [1963], Ch. 6.
[50] Comp. 3.1, Def. 3.

(7) $\Gamma\cup\{\neg\alpha\}$ is inconsistent.

(b) Suppose that (1) $\Gamma\cup\{\neg\alpha\}$ is inconsistent. Then there is a formula β such that (2) $\Gamma, \neg\alpha\vdash\beta$ and (3) $\Gamma, \neg\alpha\vdash\neg\beta$. By Deduction Theorem from (2) it follows (4) $\Gamma\vdash\neg\alpha\supset\beta$, and from (3) it follows (5) $\Gamma\vdash\neg\alpha\supset\neg\beta$. But (6) $\vdash(\neg\alpha\supset\beta)\supset[(\neg\alpha\supset\neg\beta)\supset\alpha]$ (by Th. 43, Sect. 3.2.3), and then (7) $\Gamma\vdash\alpha$ (from (4), (5), (6), by Prop. 4* and 5* of Sect. 3.1).

Definition 3. *A set* Γ *of formulas of* L_{PL} *is called maximal consistent if* Γ *is consistent and for any formula* α *of* L_{PL} *the following holds*: *if* $\Gamma\cup\{\alpha\}$ *is consistent, then* $\alpha\in\Gamma$.

Lemma 2 (Lindenbaum's Lemma). *Let* Γ *be a consistent set of formulas of* L_{PL}. *Then there is a maximal consistent set* Γ_∞ *such that* $\Gamma\subseteq\Gamma_\infty$.

Proof. With the language L_{PL} of PL an infinite number of formulas may be constructed. Let us consider such an enumeration: $\alpha_1,\alpha_2,\alpha_3,\ldots$[51] Now we define an infinite sequence of sets in the following way:

$$\Gamma_0=\Gamma$$

$$\Gamma_{n+1}=\begin{cases}\Gamma_n\cup\{\alpha_{n+1}\} \text{ if the so obtained set is consistent}\\ \Gamma_n \text{ otherwise.}\end{cases}$$

We set $\Gamma_\infty=\cup\Gamma_n$ ($n=0,1,2,\ldots$).

a) Γ_∞ *is consistent*.

Proof (*reductio*). Assume Γ_∞ is inconsistent, hence there are the formulas $\alpha_{i_1},\ldots,\alpha_{i_m}$ of L_{PL} such that $\alpha_{i_1},\ldots,\alpha_{i_m}\vdash\beta$ and $\alpha_{i_1},\ldots,\alpha_{i_m}\vdash\neg\beta$. Let k be the greatest number of $i_1,\ldots,i_m$, such that $\alpha_{i_1},\ldots,\alpha_{i_m}\in\Gamma_k$. In this case we have $\Gamma_k\vdash\beta$ and $\Gamma_k\vdash\neg\beta$, and then Γ_k would be inconsistent. But this is not possible, since by construction of this sequence of sets $\Gamma_0=\Gamma$ is consistent (*ex hypothesi*) and if Γ_n is consistent so is Γ_{n+1} (by construction of Γ_{n+1}).

b) Γ_∞ *is maximal*.

Suppose that a formula $\alpha\notin\Gamma_\infty$. Let n be its number in the above enumeration. Then by construction of the sets Γ_i we have $\Gamma_{n-1}\cup\{\alpha_n\}$ is *inconsistent* (otherwise $\Gamma_n=\Gamma_{n-1}\cup\{\alpha_n\}$). Hence $\Gamma_{n-1},\alpha_n\vdash\beta$ and $\Gamma_{n-1},\alpha_n\vdash\neg\beta$, and therefore, by Ded. Th., $\Gamma_{n-1}\vdash\alpha_n\supset\beta$ and $\Gamma_{n-1}\vdash\alpha_n\supset\neg\beta$. But since $PL^{ax}\vdash(\alpha_n\supset\beta)\supset((\alpha_n\supset\neg\beta)\supset\neg\alpha_n)$ (comp.

[51] The fact that the set of formulas of L_{PL} is effectively enumerable can be proved by a method similar to that of Ch. 2, Sect. 3.5.3, Lemma 1.

3.2.3, Th. 43) it follows that $\Gamma_{n-1} \vdash \neg\alpha$ (by Prop. 4* and 5*, 3.1) and then $\Gamma_\infty \vdash \neg\alpha$ (by Prop. 2*, 3.1). Whence $\neg\alpha \in \Gamma_\infty$ (by Lemma 3(*b*) below).

Remark. A shorter proof can also be given, using the following result: *If* Γ *is consistent, then for any formula* $\alpha \in L_{PL}$ *the following holds:* $\Gamma \cup \{\alpha\}$ *is consistent or* $\Gamma \cup \{\neg\alpha\}$ *is consistent*. Prove this result and construct a proof for b).

Lemma 3. *If* Γ *is any maximal consistent set, then the following holds of* Γ:

(*a*) *For any formula* α *of* L_{PL} exactly one of α and $\neg\alpha$ is in Γ.
And this means that $\alpha \in \Gamma$ iff $\neg\alpha \notin \Gamma$.

(*b*) *For any formula* α *of* L_{PL}: $\Gamma \vdash \alpha$ *iff* $\alpha \in \Gamma$.

(*c*) $\alpha \supset \beta \in \Gamma$ iff $\alpha \notin \Gamma$ or $\beta \in \gamma$.

Proof. (a) Γ being consistent, it is not the case that $\alpha \in \Gamma$ and $\neg\alpha \in \Gamma$. Let us suppose that $\alpha \notin \Gamma$ and $\neg\alpha \notin \Gamma$. Hence $\Gamma \cup \{\alpha\}$ and $\Gamma \cup \{\neg\alpha\}$ are both inconsistent sets, by Def. 3. Then, by Lemma 1 $\Gamma \vdash \neg\alpha$ and $\Gamma \vdash \alpha$, and then Γ is inconsistent, contrary to the assumed consistency of Γ.

Proof. (b_1) If $\Gamma \vdash \alpha$, then $\alpha \in \Gamma$.

(*Reductio*). Suppose that $\Gamma \vdash \alpha$ and $\alpha \notin \Gamma$. It follows that $\neg\alpha \in \Gamma$ (by max of Γ). Then $\Gamma \vdash \neg\alpha$ (since, by Prop. 1* (Sect. 3.1), $\neg\alpha \vdash \neg\alpha$, and then, by Prop. 2* $\Gamma, \neg\alpha \vdash \neg\alpha$, i.e., since $\neg\alpha \in \Gamma$: $\Gamma \vdash \neg\alpha$. Hence Γ is inconsistent since it is also the case $\Gamma \vdash \alpha$ (by hyp.).

(b_2) If $\alpha \in \Gamma$, then $\Gamma \vdash \alpha$.

(*Reductio*). Suppose that $\alpha \in \Gamma$ and $\Gamma \nvdash \alpha$. Then $\Gamma \cup \{\neg\alpha\}$ is consistent (by Lemma 1) and therefore $\neg\alpha \in \Gamma$ (by Def. 3). But by hypothesis $\alpha \in \Gamma$, and then Γ is inconsistent.

Proof. (c_1) If $\alpha \supset \beta \in \Gamma$, then $\alpha \notin \Gamma$ or $\beta \in \Gamma$.

(*Reductio*). Suppose that $\alpha \supset \beta \in \Gamma$ and Non($\alpha \notin \Gamma$ or $\beta \in \Gamma$), equivalent $\alpha \in \Gamma$ and $\beta \notin \Gamma$. Then $\neg\beta \in \Gamma$. But if $\alpha \in \Gamma$ and $\neg\beta \in \Gamma$ it follows (by (b)) that $\Gamma \vdash \alpha$ and $\Gamma \vdash \neg\beta$. Then, since $PL^{ax} \vdash \alpha \supset (\neg\beta \supset \neg(\alpha \supset \beta))$, it follows (by 3.1, Prop. 4* and 5*) that $\Gamma \vdash \neg(\alpha \supset \beta)$, contradicting the consistency of Γ since $\alpha \supset \beta \in \Gamma$ (by hyp.) and then $\Gamma \vdash \alpha \supset \beta$ (by (b)).

(c_2) If $\alpha \notin \Gamma$ or $\beta \in \Gamma$, then $\alpha \supset \beta \in \Gamma$. Suppose that $\alpha \notin \Gamma$. Then $\neg\alpha \in \Gamma$ (by max), and therefore $\Gamma \vdash \neg\alpha$ (by (b)). But $PL^{ax} \vdash \neg\alpha \supset (\alpha \supset \beta)$ and then $\Gamma \vdash \alpha \supset \beta$ (by 3.1, Prop. 4* and 5*). Now suppose that $\beta \in \Gamma$. Then

$\Gamma \vdash \beta$ (by (b)). But $PL^{ax} \vdash \beta \supset (\alpha \supset \beta)$, it follows (by Prop. 4* and 5*) that $\Gamma \vdash \alpha \supset \beta$. In both cases it follows that $\alpha \supset \beta \in \Gamma$.

Definition 4. *A set* Γ *of formulas of* L_{PL} *is called satisfiable if there exists an interpretation* int *of variables occurring in the formulas of* Γ *such that for any* $\alpha \in \Gamma$: $[\alpha]^{int} = 1$ (*symbolic*: Sat Γ, *or* $[\Gamma]^{int} = 1$).

Lemma 4. *If* Γ *is a consistent set of formulas of* L_{PL}, *then* Γ *is satisfiable*.

Proof. Since Γ is consistent, by Lemma 2 there is a maximal consistent set Γ_∞ such that $\Gamma \subseteq \Gamma_\infty$. If we show that Γ_∞ is satisfiable, then the satisfiability of Γ follows. Hence what we shall prove is that Γ_∞ is satisfiable. Actually, we prove the stronger fact that there is an interpretation *int* such that for any formula α of Γ_∞ the following holds:

Eq. $[\alpha]^{int} = 1$ iff $\alpha \in \Gamma_\infty$.

For this task we define an interpretation *int* for propositional variables of L_{PL} in the following way:

$[p]^{int} = 1$ iff $p \in \Gamma$,

and then show that *Eq* holds for any formula $\alpha \in L_{PL}$.

We prove this equivalence by induction on the complexity of α.

Basis. Compl(α)=0. This means that α is a variable and then Lemma holds by definition.

Induction. Assume that *Eq* holds for every formula whose complexity <*n*, and show that *Eq* holds for every formula with complexity *n*. The formula α has one of the following forms: $\neg\beta$ or $\beta \supset \gamma$.

1. $\alpha = \neg\beta$.

$[\alpha]^{int} = 1$ iff $[\neg\beta]^{int} = 1$ iff $[\beta]^{int} = 0$ iff $\beta \notin \Gamma_\infty$ (by ind. hyp.) iff $\neg\beta \in \Gamma_\infty$ (by Lemma 3 (a)) iff $\alpha \in \Gamma_\infty$.

2. $\alpha = \beta \supset \gamma$.

$[\alpha]^{int} = 1$ iff $[\beta \supset \gamma]^{int} = 1$ iff $[\beta]^{int} = 0$ or $[\gamma]^{int} = 1$ iff $\beta \notin \Gamma_\infty$ (by ind. hyp.) or $\gamma \in \Gamma_\infty$ (by ind. hyp.) iff $\beta \supset \gamma \in \Gamma_\infty$ (by Lemma 3 (*c*)).

Finally, since $\Gamma \subseteq \Gamma_\infty$, it follows that if Γ is consistent then Γ is satisfiable, i.e., Lemma 4.

Lemma 5. $\Gamma \models \alpha$ *iff* $\Gamma \cup \{\neg\alpha\}$ *is not satisfiable*.

Proof. $\Gamma\models\alpha$ iff there is Γ_0 *finite* ($\Gamma_0\subseteq\Gamma$) such that $\Gamma_0\models\alpha$[52] iff $\mathrm{Conj}(\Gamma_0)\models\alpha$ (by PL) iff $\models\mathrm{Conj}(\Gamma_0)\supset\alpha$ (by Normality) iff NonSat$\neg(\mathrm{Conj}(\Gamma_0)\supset\alpha)$ (by PL) iff NonSat$(\mathrm{Conj}(\Gamma_0)\wedge\neg\alpha)$ (by PL) iff NonSat$(\Gamma_0\cup\{\neg\alpha\})$ iff NonSat $\Gamma\cup\{\neg\alpha\}$ (since $\Gamma_0\cup\{\neg\alpha\}$ is a subset of $\Gamma\cup\{\neg\alpha\}$ and if a set is unsatisfiable, then so is any extension of it).

Now we turn to the proof of the stronger completeness theorem for PL^{ax}: if $\Gamma\models\alpha$, then $\Gamma\vdash\alpha$. Suppose, by contraposition, that $\Gamma\nvdash\alpha$, then $\Gamma\cup\{\neg\alpha\}$ is consistent (by Lemma 1) and hence $\Gamma\cup\{\neg\alpha\}$ is satisfiable (by Lemma 4), i.e., there is an interpretation *int* satisfying all the formulas in Γ and the formula $\neg\alpha$. Whence $[\alpha]^{int}=0$, and therefore $\Gamma\not\models\alpha$ (by 2.6 Definition). Hence if $\Gamma\models\alpha$, then $\Gamma\vdash\alpha$.

Another argument for stronger completeness theorem is the following. Suppose that $\Gamma\models\alpha$. Then $\Gamma\cup\{\neg\alpha\}$ is not satisfiable (by Lemma 5). And then by Lemma 4 it follows that $\Gamma\cup\{\neg\alpha\}$ is inconsistent. Whence, by Lemma 1 (above), $\Gamma\vdash\alpha$.

For $\Gamma=\varnothing$ the completeness of PL^{ax} follows, i.e., if $\models\alpha$, then $\vdash\alpha$.

Remark. The completeness theorem (stronger and weaker) holds for the system $PL^{ax(*)}$ too (argue!).

3.3.3. Some related results

Theorem 1. $\Gamma\vdash\alpha$ *iff* $\Gamma\models\alpha$.

Proof. From 3.3.1 Soundness Theorem (stronger form) and 3.3.2 Completeness Theorem (stronger form).

Remark. If $\Gamma=\varnothing$, then for PL^{ax} the following holds: $\vdash\alpha$ iff $\models\alpha$. This equivalence shows the co-extensivity of the two notions: one syntactic (theorem) and one semantic (valid formula) and then allows us to interchange the symbols "$\vdash$" and "$\models$" arbitrarily.

Interpolation Theorem for $\mathbf{PL^{ax}}$.[53] *If* $\vdash\alpha\supset\beta$ *and* α *and* β *have at least one variable in common, then there is a formula* γ *of* L_{PL} *all of whose variables occur in both* α *and* β *such that* $\vdash\alpha\supset\gamma$ *and* $\vdash\gamma\supset\beta$. γ *is called the interpolant for the implication* $\alpha\supset\beta$.

---

[52] Comp. Finiteness Theorem for "$\models$" (3.3.3 below).

[53] This is the *syntactic* version of the corresponding theorem for PL; comp. 2.9.

Proof. From 2.9 Interpolation Theorem and Theorem 1 above.

Theorem 2. *Let* Γ *be a set of formulas of* L_{PL}. *Then the following holds*: Γ *is consistent if and only if* Γ *is satisfiable*.

Proof. The left-right part, by Lemma 4 (3.3.2). For the right-left part, if Γ is satisfiable, then Γ is consistent, we suppose, by *reductio*, that Γ is satisfiable but it is not consistent. Hence there is a formula α of L_{PL} such that $\Gamma \vdash \alpha$ and $\Gamma \vdash \neg\alpha$ (by 3.3.2, Definition 2). And then we have $\Gamma \models \alpha$ and $\Gamma \models \neg\alpha$ (by 3.3.1, Soundness Theorem (stronger form)). But by hypothesis, Γ is satisfiable, i.e., there is an interpretation *int* such that all the formulas in Γ are true in *int*. Hence from $\Gamma \models \alpha$ we deduce that $[\alpha]^{int} = 1$ and from $\Gamma \models \neg\alpha$ it follows that $[\neg\alpha]^{int} = 1$, and this is impossible.

Finiteness Theorem for semantic consequence. $\Gamma \models \alpha$ *iff there is a finite subset* Γ_0 *of* Γ *such that* $\Gamma_0 \models \alpha$.

Proof. a) If there is a finite $\Gamma_0 \subseteq \Gamma$ such that $\Gamma_0 \models \alpha$, then since $\Gamma = \Gamma \cup \Gamma_0$ it follows that $\Gamma \models \alpha$ (by Sect. 3.1, Prop. 2*).

b) If $\Gamma \models \alpha$, then $\Gamma \vdash \alpha$, by Theorem 1, and then there is a *finite* sequence of formulas constituting the deduction of α from Γ (by 3.1, Def 3 and Def 4). Hence there is a finite $\Gamma_0 \subseteq \Gamma$ such that $\Gamma_0 \vdash \alpha$, whence, by Theorem 1, $\Gamma_0 \models \alpha$.

Theorem 3. *Let* Γ *be a set of formulas of* L_{PL}. *Then* Γ *is consistent iff every finite set* $\Delta \subseteq \Gamma$ *is consistent.*

Proof. The left-right part is simple, since if a finite set $\Delta \subseteq \Gamma$ were inconsistent, then any extension of it (and therefore Γ itself) would be inconsistent.

For the right-left part we proceed by *reductio*. Assume that every finite $\Delta \subseteq \Gamma$ is consistent and Γ is not consistent. Hence for some α $\Gamma \vdash \alpha$ and $\Gamma \vdash \neg\alpha$. But by 3.1, Def 3 and Def 4, a derivation has only *finitely* many formulas, and then from Γ only finitely many formulas occur in this derivation. Let Δ be the set of such formulas. Then $\Delta \vdash \alpha$ and $\Delta \vdash \neg\alpha$, i.e., Δ is inconsistent, contradicting our assumption.

The Compactness Theorem for PL. *Let* Γ *be a set of formulas of* L_{PL}. *Then the following holds*: Γ *is satisfiable if and only if every finite subset of* Γ *is satisfiable*.

Proof. By Theorem 2 and Theorem 3.

3.3.4. Completeness of PL^{ax} (Kalmár-style[54])

The proof of the fact that every valid formula of L_{PL} is provable in PL^{ax}, using Kalmár's method, requires a lemma asserting the existence of a deducibility relation[55] corresponding to each row of the truth table of a formula α. Let us formulate this lemma, illustrate it and prove it.

Lemma. *Let* $\alpha(p_1,...,p_k)$ *be a formula of* L_{PL} *whose only distinct variables are* $p_1,...,p_k$ $(k \geq 1)$. *Let int be an interpretation of* $p_1,...,p_k$. *Let us define* p_i^*, $1 \leq i \leq k$, *and* α^* *in the following way:*

$$p_i^* = \begin{cases} p_i, & \text{if} \quad [p_i]^{int} = 1 \\ \neg p_i, & \text{if} \quad [p_i]^{int} = 0 \end{cases} \qquad \alpha^* = \begin{cases} \alpha, & \text{if} \quad [\alpha]^{int} = 1 \\ \neg\alpha, & \text{if} \quad [\alpha]^{int} = 0 \end{cases}$$

Then: $p_1^*,...,p_k^* \vdash \alpha^*$.

Example. Let $\alpha = p_1 \supset (p_2 \supset \neg p_1)$

p_1	p_2	$\neg p_1$	$p_2 \supset \neg p_1$	$p_1 \supset (p_2 \supset \neg p_1)$
1	1	0	0	0
1	0	0	1	1
0	1	1	1	1
0	0	1	1	1

According to the lemma to each row of this truth table, that is, to each interpretation of variables of α corresponds a deducibility relation:

Int_1: $p_1, p_2 \vdash \neg\alpha$

Int_2: $p_1, \neg p_2 \vdash \alpha$

Int_3: $\neg p_1, p_2 \vdash \alpha$

Int_4: $\neg p_1, \neg p_2 \vdash \alpha$

Proof (induction on the complexity of α).

Basis. $n = 0$, i.e., the formula α is a variable p. Then if $[p]^{int} = 1$, then $p \vdash \alpha$ reduces to $p \vdash p$. And if $[p]^{int} = 0$, then $\neg p \vdash \alpha$ means $\neg p \vdash \neg p$, and both these relations hold.

Induction. Assume that lemma holds for any formula α of L_{PL} such that compl(α) < *n*, and show that lemma holds also for formulas α such that compl(α) = *n*.

---

[54] L. Kalmár [1935].

[55] As a pure syntactic fact!

Again we consider the formulas without abbreviations, hence α can have one of the following forms: $\neg\beta$ and $\beta\supset\gamma$, where compl(β) < n and compl(γ) < n.

1. $\alpha=\neg\beta$.

a) Assume $[\alpha]^{\text{int}}=1$, and then $\alpha^*=\alpha$. Hence $[\beta]^{\text{int}}=0$, and then $\beta^*=\neg\beta$. And since compl(β) < n, by inductive hypothesis we have:

(1) $p_1^*,\ldots,p_k^* \vdash \beta^*$

(2) $p_1^*,\ldots,p_k^* \vdash \neg\beta$

(3) $p_1^*,\ldots,p_k^* \vdash \alpha$

(4) $p_1^*,\ldots,p_k^* \vdash \alpha^*$

b) Assume $[\alpha]^{\text{int}}=0$, then $\alpha^*=\neg\alpha$. Hence $[\beta]^{\text{int}}=1$ and then $\beta^*=\beta$. Again, by inductive hypothesis:

(1) $p_1^*,\ldots,p_k^* \vdash \beta^*$

(2) $p_1^*,\ldots,p_k^* \vdash \beta$

(3) $p_1^*,\ldots,p_k^* \vdash \neg(\neg\beta)$; (2), 3.2.3, Th32

(4) $p_1^*,\ldots,p_k^* \vdash \neg\alpha$

(5) $p_1^*,\ldots,p_k^* \vdash \alpha^*$

2. $\alpha=\beta\supset\gamma$. By inductive hypothesis, lemma holds for β and γ, i.e., $p_1^*,\ldots,p_k^* \vdash \beta^*$ and $p_1^*,\ldots,p_k^* \vdash \gamma^*$. According to the truth table of $\supset$ we distinguish the following three subcases:

a) $[\beta]^{\text{int}}=0$. Then $[\alpha]^{\text{int}}=1$ and therefore $\beta^*=\neg\beta$ and $\alpha^*=\alpha$. Hence

(1) $p_1^*,\ldots,p_k^* \vdash \neg\beta$

(2) $p_1^*,\ldots,p_k^* \vdash \beta\supset\gamma$; (1), 3.2.3, Th2, MP

(3) $p_1^*,\ldots,p_k^* \vdash \alpha^*$

b) $[\gamma]^{\text{int}}=1$. Then $[\alpha]^{\text{int}}=1$, hence $\gamma^*=\gamma$ and $\alpha^*=\alpha$. And then

(1) $p_1^*,\ldots,p_k^* \vdash \gamma$

(2) $p_1^*,\ldots,p_k^* \vdash \beta\supset\gamma$; (1) Ax1, MP

(3) $p_1^*,\ldots,p_k^* \vdash \alpha^*$

c) $[\beta]^{\text{int}}=1$ and $[\gamma]^{\text{int}}=0$. And then $\beta^*=\beta$, $\gamma^*=\neg\gamma$ and $\alpha^*=\neg\alpha=\neg(\beta\supset\gamma)$. We have:

(1) $p_1^*,\ldots,p_k^* \vdash \beta$

(2) $p_1^*,\ldots,p_k^* \vdash \neg\gamma$

(3) $p_1^*,\ldots,p_k^* \vdash \neg(\beta\supset\gamma)$; (1) (2), 3.2.3, Th11, MP.

(4) $p_1^*,\ldots,p_k^* \vdash \alpha^*$

Completeness Theorem for PL^{ax}. *If* $\models\alpha$, *then* $\vdash\alpha$.

Proof. Assume $\models\alpha(p_1,\ldots,p_k)$. Hence for every interpretation *int* of the variables $p_1,\ldots,p_k$ we have $[\alpha]^{int}=1$, and then $\alpha^*=\alpha$. By the preceding Lemma, we have:

$$p_1^*,\ldots,p_k^* \vdash \alpha$$

and then:

a) If $[p_k]^{int}=1$ (i.e., $p_k^*=p_k$), then $p_1^*,\ldots,p_{k-1}^*,p_k \vdash \alpha$.

b) If $[p_k]^{int}=0$ (i.e., $p_k^*=\neg p_k$), then $p_1^*,\ldots,p_{k-1}^*,\neg p_k \vdash \alpha$.

Whence, by Deduction Theorem, from a) and b) we get, accordingly

(1) $p_1^*,\ldots,p_{k-1}^* \vdash p_k\supset\alpha$

(2) $p_1^*,\ldots,p_{k-1}^* \vdash \neg p_k\supset\alpha$

And by 3.2.3, Th12: $(\alpha\supset\beta)\supset((\neg\alpha\supset\beta)\supset\beta)$ and MP we get

(3) $p_1^*,\ldots,p_{k-1}^* \vdash \alpha$.

Now we repeat the preceding process, this time by taking the variable p_{k-1}, and after *k* steps we drop all the variables $p_1,\ldots,p_k$, and finally obtain: $\vdash\alpha$. But α was an arbitrary *valid* formula of L_{PL}. Therefore: if $\models\alpha$, then $\vdash\alpha$.

3.3.5. Syntactic completeness and decidability of PL^{ax}

3.3.5.1. Syntactic completeness of PL^{ax}

The idea of completeness, treated above, is usually known as the *semantic* completeness of PL^{ax}. But beside it there are several ideas of *syntactic* completeness, from which we select three as being interesting, given by the definitions below.

Definition 1. *An axiomatic system S is syntactically complete if and only if for each formula* α *of its language the following holds*: *either* $\vdash\alpha$ *or* $\vdash\neg\alpha$.

As can be argued, PL^{ax} is *not* complete in this sense, since neither p nor ¬p is provable in PL^{ax}. By 3.3.3 Theorem 1 *Remark* only valid formulas are provable, but neither p nor ¬p is valid.

Definition 2. *An axiomatic system S is syntactically complete if and only if no unprovable* ***schema*** *can be added to it without destroying consistency.*

As we saw in 3.1, the difference between an axiom and an axiom schema consists of the following: an axiom is a formula of L_{PL}, e.g., $p\supset(q\supset p)$, but an axiom schema S is a stencil, like $\alpha\supset(\beta\supset\alpha)$,[56] according to which we can construct an infinite number of axioms. Hence in such a schema containing say $S_1,...,S_k$ schematic symbols if we substitute arbitrary formulas say $\beta_1,...,\beta_k$ for $S_1,...,S_k$ respectively, what is obtained will be an axiom.

Theorem 1. PL^{ax} *is syntactically complete (in the sense of Definition 2).*

Proof. Let us suppose that S is an unprovable schema in PL^{ax}, and that S is added as a new axiom to PL^{ax} obtaining a new system $PL^{ax'}$. Since $\nvdash S$ in PL^{ax}, it follows that $\nvDash S$ (by completeness of PL^{ax}), i.e., there is an interpretation *int* of the schematic symbols in S such that $[S]^{int}=0$. Hence $PL^{ax'}\vdash S$, and therefore any formula obtained from S by substituting the formulas $\beta_1,...,\beta_k$ for $S_1,...,S_k$ respectively, will be a theorem of $PL^{ax'}$. Let us consider the following substitutions in S such that in *int* S is false: we set $p\supset p$ for each S_i, $1\leq i\leq k$, for which $[S_i]^{int}=1$ and $\neg(p\supset p)$ for each S_i such that $[S_i]^{int}=0$. If for example the unprovable schema is $S=S_1\supset S_2$ then $[S]^{int}=0$ for $[S_1]^{int}=1$ and $[S_2]^{int}=0$. Hence, we substitute $p\supset p$ for S_1 and $\neg(p\supset p)$ for S_2 and obtain the formula $(p\supset p)\supset\neg(p\supset p)$. Let us call β the so obtained formula. Being obtained in $PL^{ax'}$, the following holds: $PL^{ax'}\vdash\beta$. But β is not satisfiable, hence $\vDash\neg\beta$, and then $PL^{ax}\vdash\neg\beta$ and therefore $PL^{ax'}\vdash\neg\beta$. Hence $PL^{ax'}$ proves β and ¬β, i.e., it is inconsistent. Therefore, PL^{ax} is syntactically complete (in the sense of Def 2).

Definition 3. *An axiomatic system is syntactically complete if and only if no unprovable* ***formula*** *of its language can be added without destroying consistency.*

Theorem 2. PL^{ax} *is not syntactically complete* (*in the sense of Definition* 3).

Actually, we prove this theorem by proving the following fact: *if* $PL^{ax}\nvdash\neg\alpha$, *then* $PL^{ax}\cup\{\alpha\}$ *is consistent.*

Proof (contraposition). Suppose that $PL^{ax}\cup\{\alpha\}$is inconsistent, i.e., there is

[56] We call α, β *schematic symbols*.

a formula β of L_{PL} such that $PL^{ax}, \alpha \vdash \beta$ and $PL^{ax}, \alpha \vdash \neg\beta$. And then, by Ded. Th. $PL^{ax} \vdash \alpha \supset \beta$ and $PL^{ax} \vdash \alpha \supset \neg\beta$. But $PL^{ax} \vdash (\alpha \supset \beta) \supset ((\alpha \supset \neg\beta) \supset \neg\alpha)$ (Th. 43, Sect. 3.2.3). Whence, by MP (twice), $PL^{ax} \vdash \neg\alpha$.

Remark. If we proceed as in Ch. 2 Sect. 3.5.2, by referring to PL^{ax} as being the sets of all provable formulas in PL^{ax}, then Theorem 2 follows directly from Lemma 1 (Sect. 3.3.2).

Hence if $\neg\alpha$ is any formula not provable in PL^{ax} then by addition of α to PL^{ax} as a new axiom we obtain a consistent system.

Remark. The Theorem 2 does not hold for a system with a substitution rule as a rule of inference, for in such a case the argument for Theorem 1, with obvious changes, can be used in order to show that by adding to PL^{ax} any unprovable formula as an axiom, the resulting system is inconsistent. Hence such a system with a substitution rule is also one syntactically complete.

3.3.5.2. Decidability of PL^{ax}

Definition 1. *An axiomatic system S is decidable if and only if there is a method whose application enable us to tell of each formula (of its language) whether or not it is a theorem of S.*

Theorem. PL^{ax} *is decidable.*

Proof. By 3.3.3, Theorem 1, Remark, for any formula α of L_{PL}: $\vdash \alpha$ iff $\models \alpha$. But by 2.7 using anyone of those methods (truth table, Quine's, *reductio test* or normal forms) we can effectively determine whether or not a formula is decidable. Therefore PL^{ax} is decidable.

Definition 2. *A formula* α *is decidable in an axiomatic system if and only if the following holds*: *either* $\vdash \alpha$ *or* $\vdash \neg\alpha$.

As we know, there are formulas α of L_{PL} such that neither α not $\neg\alpha$ is provable in PL^{ax}, p, $\neg$p, $p \supset q$, $\neg q \vee r$ being some examples of these formulas.

Remark. Though PL^{ax} is a decidable system, it has undecidable formulas.

Definition 3. *An axiomatic system S has an effective proof procedure if and only if for any arbitrary theorem* α *of S there is an effective method for constructing a proof of* α *in S.*

Remark. For PL^{ax} there is such a procedure (an awkward one!) based on the proof of completeness (Kalmár-type).

Chapter 2. FIRST-ORDER LOGIC

The first-order logic (FOL) we develop in this chapter, also known as "predicate calculus"[1] or "quantification theory"[2], is a logic whose only quantified variables are individual variables and in which no place of arguments of a predicate is occupied by a predicate symbol.

1. Syntax of FOL

1.1. Alphabet, terms, formulas

A first-order language (L_{FOL}) depends on what we intend to do by using it. Some of its symbols are common to all of first order languages (the items (1)-(4) below) and some of them are different from language to language (the items (5)-(7) below).

Usually, the language of FOL contains the following symbols: *Alphabet*

(1) Symbols for propositional connectives: $\neg$, $\wedge$, $\vee$, $\supset$, $\equiv$

(2) Symbols for quantifies: $\forall$ (*for all*, the universal quantifier)
$\exists$ (*there exists*, the existential quantifier)

(3) Symbols for individual variables: x_1, x_2, x_3,..., sometimes informally denoted by x,y,z,…

(4) Auxiliary symbols:), (, sometimes informally given by], [or }, {, and , (comma)

(5) Predicate symbols: P_1, P_2, P_3,... (sometimes informally given by P, Q, R,...)

(6) Function symbols: f_1, f_2, f_3,... (sometimes informally given by f, g, h,...)

(7) Constant symbols: c_1, c_2, c_3,... (sometimes informally given by a, b, c,...)

The set of (3) is denumerable, the set of (5) is finite or denumerable but not empty, and the sets of (6) and (7) are finite or denumerable, possibly empty.

---

[1] Cf. S.C. Kleene, [1967], Ch. II.

[2] Cf. E. Mendelson, [1964], Ch. 2.

If by *Pred*, *Funct* and *Const* we denote the sets of (5), (6) and (7) respectively, then properly speaking the language of FOL (L_{FOL}) is defined in the following way:

$$L_{FOL}=Pred\cup Funct\cup Const$$

Let us define two key notions of the syntax of FOL: *term of* L_{FOL} and *formula of* L_{FOL}.

Definition 1. *The set* Term *of terms of* L_{FOL} *contains the following expressions:*

1. *any symbol of individual variable* (*Var*)
2. *any constant symbol* (*Const*)
3. *if* f *is a function symbol and* $t_1,...,t_n\in$Term, *then* $f(t_1,...,t_n)$ *is a term*.

Example. x, f(y), a, g(a,z), h(z,b,g(x))

If a term t does not contain any variable, then it is *closed*. Sometimes a function symbol f occurs with a superscript *n*, i.e., f^n, with $n\geq 1$, *n* indicating the number of accompanying arguments of f: f^1-one-place, f^2-two-place,..., f^n-*n*-place.

Definition 2. *The set* Form *of formulas of* L_{FOL} *contains strictly the following expressions:*

1. $R(t_1,...,t_n)$, *where* R *is a predicate symbol*[3] *and* $t_1,...,t_n$ *are terms. Such an expression is called atomic formula*.
2. $\neg\alpha$, *where* $\alpha\in$Form.
3. $\alpha\circ\beta$, *where* α, $\beta\in$Form, *and* "$\circ$" *is a symbol for a binary propositional connective.*
4. a) $\forall x\alpha$, *and* b) $\exists x\alpha$, *where* $\alpha\in$Form, *and* x *is a symbol for an individual variable*.

Example. P(x), Q(x,y), R(x,y,z), $\forall$xQ(x), $\exists$yR(x,y), P(x)$\wedge$Q(x,y), $\exists$x(P(x,y)$\supset$R(x)), $\exists$xS(a,x), $\forall$yT(f(x),b,c).

Let us consider the following formulas:

$$\alpha=P(x)\supset\exists yQ(x,y,z) \text{ and } \beta=\forall x(P(x)\supset\exists yQ(x,y,z)).$$

The atomic formula Q(x,y,z) in α and β, we say, is in the scope of the quantifier $\exists$y, and the formula P(x)$\supset\exists$yQ(x,y,z) in β is in the scope of the quantifier $\forall$x. As can be seen, some of the variables are in the scope of some quantifiers or are the variables accompanying them: two occurrences of y in α, three occurrences of x and two occurrences of y in β. Such

[3] Similar to function symbols, the superscript n in R^n will indicate the number of arguments of R: R^1-one-place predicate symbol,..., R^n-n-place predicate symbol. When the arguments are given, as in $R(t_1,...,t_n)$, this symbol can be omitted.

variable occurrences are called *bound*. The variable x, instead, has two *free* occurrences in α and the variable z has one free occurrence in β.

Definition 3. *An occurrence of a variable* x *is bound in* α *if it is the variable accompanying a quantifier or it is the same variable within the scope of it, otherwise* x *is free*.

Of course, a variable x can have both bound and free occurrences in α, as is the case with y in $P(x,y)\vee\exists yQ(x,y,z)$.

Definition 4. *A sentence (or closed formula) of* L_{FOL} *is a formula with no free variable*.

1.2. Substitution

Some terms and formulas may contain (individual) variables x,y,z,..., and these variables may be replaced by arbitrary terms. Such an operation is called *substitution*.

Definition 1. *A substitution is a mapping* σ *from the set of variables to the set of terms, i.e.,*

σ:*Var*$\rightarrow$Term.

Let in what follows $\sigma(t)$ and $\sigma(\alpha)$ be the notations for "the substitution in a term t" and "the substitution in a formula α", respectively.

According to the above definition, evidently, the following holds: $\sigma(c) = c$ and $\sigma(f(t_1,...,t_n)) = f(\sigma(t_1),...,\sigma(t_n))$, where each of t_i $(1 \leq i \leq n)$ are terms. Then we can say that the application of σ to a term t does generate some other term.

Example 1. Let $t = f(x,a,y)$, and σ be a substitution defined in the following way: $\sigma(x) = b$ and $\sigma(y) = c$. Let us calculate $\sigma(t)$.

$\sigma(f(x,a,y)) = f(\sigma(x), \sigma(a), \sigma(y)) = f(b,a,c)$.

Example 2. Let $t = g(x,h(y,a),z)$ and σ be defined in the following way: $\sigma(x) = w$, $\sigma(y) = b$ and $\sigma(z) = f(c)$. Then $\sigma(g(x),h(y,a),z) = $ $=g(\sigma(x),\sigma(h(y,a)),\sigma(z)) = g(\sigma(x),h(\sigma(y)),\sigma(a)),\sigma(z)) = g(w,h(b,a),f(c))$.

Similarly, as we said, the substitution can be carried out in a formula α. But in this case this operation has a restriction, given by the fact that a formula α may contain both free and bound variables, and the substitution must not affect the bound variables of α, i.e., only the *free* variables of a formula α may be substituted. This restriction is indicated by the following notation: σ_x, where the subscript "x" suggests that σ should not affect the bound occurrences of the variable x. This means that $\sigma_x(y) = \sigma(y)$ if $y \neq x$, and $\sigma_x(x) = x$ otherwise. Hence

1. $\sigma(R(t_1,...,t_n)) = R(\sigma(t_1),...,\sigma(t_n))$
2. $\sigma(\neg\alpha) = \neg\sigma(\alpha)$
3. $\sigma(\alpha \circ \beta) = \sigma(\alpha) \circ \sigma(\beta)$
4. a) $\sigma(\forall x\alpha) = \forall x\sigma_x(\alpha)$

   b) $\sigma(\exists x\alpha) = \exists x\sigma_x(\alpha)$.

Example. Let $\alpha = \exists xQ(x,y,z)$, and σ be defined by: $\sigma(x) = a$, $\sigma(y) = b$ and $\sigma(z) = c$. Then $\sigma(\exists xQ(x,y,z)) = \exists x\sigma_x(Q(x,y,z)) = \exists xQ(\sigma_x(x),\sigma_x(y),\sigma_x(z)) =$ $=\exists xQ(x,b,c)$.

A substitution σ in a formula α is *correct* only if it meets the following requirement: *the term* t *is free for* x *in the formula* α, according to the following definition.

Definition 2. *Let* α *be a formula of* L_{FOL} *and* t *be a term of* L_{FOL}. *Then* t *is free for* x *in* α *if and only if no free occurrence of* x *in* α *lies within the scope of a quantifier* $\forall y$ *or* $\exists y$ *and* y *is a variable in* t.

(Equivalent: t is free for x in α if t does not contain any variable which by substitution of t for x becomes a bound variable in the formula α).

Example. Suppose that $\alpha = \exists z\forall yQ(x,y,z)$ and $t_1 = a$, $t_2 = f(a,z)$ and $t_3 = g(y,b,c)$. According to above definition, only t_1 is free for x in α, since the terms t_2 and t_3 contain z and y, respectively, which become *bound* by substitution in α.

Remark 1. If a term t is closed (*i.e.,* it does not contain any variable), then t is free for any free variable in any formula of L_{FOL}.

Remark 2. In a formula α a term t is free for any free variable of α if no variable of t has bound occurrences in α (give an example!).

Notation. The intuitive notations used for defining the substitutions σ above, will be rendered in what follows by the more usual ones, of the following form: $b\,|\,x$, $f(a)\,|\,x$, $g(y,a)\,|\,x$, whose reading is: the substitution of the respective term (b,f(a),g(y(a)) for the variable x in a term t or in a formula α.

2. Semantics of FOL

2.1. Notions

The syntax of a language takes into account only grammatical aspects of the symbols and sequences of symbols. But a symbol has a meaning only by *interpretation*. And this is just what the semantics of a language intends to do. To interpret the language L_{FOL} means to interpret predicate, function and constant symbols. And this requires, in a way we specify below, the existence of a non-empty set called the *domain*. Both components, a domain and an interpretation function, specify a *model*. However, a formula like $P(x)\supset Q(y,z)$, can contain free variables, hence we have to specify their semantic values too, and this is given by a distinct function called *assignment*. Let us detail.

Definition 1. *A model for the language* L_{FOL} *is the ordered pair* $M=\langle D,i\rangle$ *where*

1. D *is a non-empty set* (*the domain of* M)
2. i *is the interpretation function, i.e., a mapping that associates*
   a) *to every* n-*place predicate symbol* $R^n\in \text{Pred}$ *an* n-*ary relation* $(R^n)^i\subseteq D^n$
   b) *to every* n-*place function symbol* $f^n\in \text{Funct}$ *an* n-*ary function* $(f^n)^i : D^n\to D$
   c) *to every constant symbol* $c\in \text{Const}$ *some element* $c^i\in D$.

Definition 2. *An assignment in* $M=\langle D,i\rangle$ *is a mapping* μ *from the set of variables to the set* D; *i.e.,*

$$\mu: Var\to D.$$

As can be seen if the mappings i and μ are given, then the semantic value of any term can be determinated, i.e.,

$$1.\ x^{i,\mu}=x^{\mu}\quad 2.\ c^{i,\mu}=c^{i}\quad 3.\ (f(t_1,\ldots,t_n))^{i,\mu}=f^{i}(t_1^{i,\mu},\ldots,t_n^{i,\mu}).$$

If a term t doesn't contain any variable, then its semantic value is not relative to any assignment.

Definition 3. *Let* $M=\langle D,i\rangle$ *be a model for* L_{FOL}, *let* x *be a variable and let* μ *and* ν *be two assignments in* M.[4] μ *and* ν *are called* x-*variants* (*or* x-*alternatives*) *if they assign the same value to every variable except possibly* x.

[4] This notation is borrowed from M. Fitting [1990].

The following three notions are fundamental to any semantic treatment of first-order logic: "truth of a formula α of L_{FOL} in an assignment μ in M", "truth of α in M" and "validity of α". Let us define them.

Definition 4. *Let* $M = \langle D,i \rangle$ *be a model for* L_{FOL} *and let* μ *be an assignment in* M. *Then*

1. $[R(t_1,...,t_n)]^{i,\mu} = 1$ *iff* $\langle t_1^{i,\mu},...,t_n^{i,\mu} \rangle \in R^i$
2. $[\neg\alpha]^{i,\mu} = \neg[\alpha]^{i,\mu}$
3. $[\alpha \circ \beta]^{i,\mu} = [\alpha]^{i,\mu} \circ [\beta]^{i,\mu}$
4. a) $[\forall x\alpha]^{i,\mu} = 1$ *iff* $[\alpha]^{i,\nu} = 1$ *for every assignment* ν x-*variant of* μ.

   b) $[\exists x\alpha]^{i,\mu} = 1$ *iff* $[\alpha]^{i,\nu} = 1$ *for some assignment* ν x-*variant of* μ.

Explanation. The notation $[\alpha]^{i,\mu}$ means: the truth value of α in the interpretation i of $M = \langle D,i \rangle$ and the assignment μ in M. In the case 1 we have an atomic formula, where R is an *n*-place predicate symbol and $t_1,...,t_n$ are arbitrary terms. According to Def. 1, the item 2a), $(R^n)^i$ (we'll write simply R^i) is an *n*-ary relation (predicate) $R^i \subseteq D^n$, where D^n is the set of all *n*-tuples built up from the members of D. Let us illustrate this idea.

Example. Let D = {1,2,3} be the domain of $M = \langle D,i \rangle$.[5] Then D^2 will be the set of all ordered pairs of members of D, namely

$$D^2 = \{\langle 1,1\rangle,\langle 1,2\rangle,\langle 1,3\rangle,\langle 2,1\rangle,\langle 2,2\rangle,\langle 2,3\rangle,\langle 3,1\rangle,\langle 3,2\rangle,\langle 3,3\rangle\}.$$

Now, if by R^i we understand the relation "less than", then R^i will be identified with a subset of D^2 formed by those pairs whose members lie in the relation "less than"; that is $R^i = \{\langle 1,2\rangle,\langle 1,3\rangle,\langle 2,3\rangle\}$. As can be seen, $R^i \subseteq D^2$.

Let now R(x,y) be an atomic formula of L_{FOL}, where R^i is the above relation "less than". Then

$$[R(x,y)]^{i,\mu} = 1 \text{ iff } \langle x^{i,\mu}, y^{i,\mu} \rangle \in R^i \text{ iff } \langle x^\mu, y^\mu \rangle \in R^i.$$

If $x^\mu = 1$ and $y^\mu = 3$, then $\langle 1,3\rangle \in R^i$, hence $[R(x,y)]^{i,\mu} = 1$.

If $x^\mu = 3$ and $y^\mu = 2$, then $\langle 3,2\rangle \notin R^i$, hence $[R(x,y)]^{i,\mu} = 0$.

If Q is a 2-place predicate symbol whose meaning is "equal to", then $Q^i = \{\langle 1,1\rangle,\langle 2,2\rangle,\langle 3,3\rangle\}$ etc.

---

[5] An equivalent notation: $M = \langle\{1,2,3,\},i\rangle$.

The meaning of R^n, given by $(R^n)^i \subseteq D^n$ can be rendered, equivalently, using the following notation: $(R^n)^i = \{\langle x_1,...,x_n\rangle \mid \text{Cond}\}$ where *Cond* is the condition the *n*-tuples of objects in D must satisfy. If $(R^2)^i$ is "less than" relation, then $(R^2)^i = \{\langle x, y\rangle \mid x < y\}$ and so on.

What the item 4 of Def. 1 says is the following fact: the truth value of a formula α does not depend on any assignment to the *quantified* variables (but still depends on the assignment to its free variables). Hence if a formula has no free variables (i.e., it is a sentence), then its truth value does not depend on any assignment.

Example. Let α be the formula x≤y and let $M = \langle D,i\rangle$ be a model whose domain is the set of natural numbers. It is clear, the truth value of α depends on the assignment given to *both* variables, x and y. But if β = ∀y(x≤y), then its truth value does not depend on y. Actually, the meaning of β can be expressed without any reference to y: "x is the least natural number". However, the truth of β still depends on the free variable x. Finally, if γ = ∃x∀y(x≤y), then γ is true regardless of any assignment given to x or y, since what this *closed* formula says is: "there exists a smallest natural number", a truth of number theory.

Definition 5. *A formula* α *of* L_{FOL} *is satisfiable in a model* $M = \langle D,i\rangle$ *if and only if there is some assignment* μ *in* M *such that* $[\alpha]^{i,\mu} = 1$.

Definition 6. *A formula* α *of* L_{FOL} *is satisfiable if and only if there is a model* $M = \langle D,i\rangle$ *in which* α *is satisfiable.*

Example. P(x)⊃¬P(y), P(x)⊃¬P(x) are satisfiable, but ¬(P(x)⊃P(x)) is not.

Def. 5 and Def. 6 can be accordingly extended to sets of formulas.

Definition 7. *A formula* α *of* L_{FOL} *is true in a model* $M = \langle D,i\rangle$ *if and only if* $[\alpha]^{i,\mu} = 1$ *for any assignment* μ *in* M.

Definition 8. *A formula* α *of* L_{FOL} *is false in a model* $M = \langle D,i\rangle$ *if and only if* $[\alpha]^{i,\mu} = 0$ *for any assignment* μ *in* M.

Of course, if α has at least one free variable x, then α could be neither true nor false in a model M. This is the case with α = P(x)⊃¬P(y), when the domain of $M = \langle D,i\rangle$ is the set of natural numbers and $P^i = \text{"even"}$.

As can be seen, the following equivalence holds:

(Eq) α is true in M iff ¬α is false in M.

Definition 9. *A formula* α *of* L_{FOL} *is a valid formula of* FOL *if and only if* α *is true in every model* $M = \langle D,i\rangle$ *for* L_{FOL}.

Sometimes in our proofs in FOL we'll use a relativized notion of validity: *k-validity*.[6]

Definition 10. *A formula* α *of* L_{FOL} *is k-valid if and only if* α *is true in every model* $M = \langle D,i \rangle$ *such that* $|D| = k$.[7]

If D has a finite number of members, say *k*, then the quantified expressions can be expressed, equivalently, by conjunctions and disjunctions, in the following form:

$$\forall xP(x)\equiv(P(1)\wedge P(2)\wedge...\wedge P(k))$$
$$\exists xP(x)\equiv(P(1)\vee P(2)\vee...\vee P(k)).$$

And if a formula α contains free variables, then we assign them the values from D. For example, suppose that $\alpha = \forall xP(x)\vee\neg P(x)$. If $D = \{1\}$, then $\alpha(1)$ is the formula $P(1)\vee\neg P(1)$, and it is not falsifiable, irrespective of the interpretation of P. We say that $\alpha(x)$ is 1-*valid*. But if $D = \{1,2\}$, then for the two possible assignment to the free x in α, we get two non-valid formulas of L_{FOL}: $(P(1)\wedge P(2))\vee\neg P(1)$ and $(P(1)\wedge P(2))\vee\neg P(2)$. Hence $\alpha(x)$ is not 2-valid. In a similar fashion we can show that $\beta(x) = \forall xP(x)\vee\forall x(\neg P(x)\vee Q(x))\vee\forall x(\neg P(x)\vee\neg Q(x))$ is 1- and 2-valid but not 3-valid (show that!).

Finally, let us observe that if a formula $\alpha(x)$ is *k*+1-valid, then it is also *k*-valid.[8] Since a formula α^* resulting from a formula $\alpha(x)$ applied to a domain of *k* members, results, in its turn, from a formula α^{**} resulting from $\alpha(x)$ applied to a domain of *k*+1-members 1,2,...,*k*+1, if everywhere the argument *k*+1 is replaced by 1 and if the repeating atomic formulas are kept only once. If, for example, the argument *k*+1 is replaced by 1, then the atomic formulas of the form $P(k+1)$ are replaced by $P(1)$, $Q(2,k+1)$ by $Q(2,1)$ etc. By these replacements and by elimination of repetitions, from α^{**} the formula α^* is obtained, and α^* is also valid.

Definition 11. *A formula* α *of* L_{FOL} *is k-satisfiable if there is a model* $M = \langle D,i \rangle$ *such that* $|D| = k$ *in which* α *is satisfiable*.

Evidently, the following hold:

1. α is *k*-satisfiable iff $\neg\alpha$ is not *k*-valid
2. α is not *k*-valid iff $\neg\alpha$ is *k*-satisfiable
3. α is not *k*-satisfiable iff $\neg\alpha$ is *k*-valid (from 1).

Let us illustrate the last case.

---

[6] Cf. D. Hilbert, P. Bernays [1934], 119.

[7] I.e., the cardinality of D is k.

[8] For a proof of this fact, comp. Sect. 5.1.3, Theorem 1.

Suppose $\alpha = \forall x \neg R(x,x) \wedge \forall x \forall y \forall z((R(x,y) \wedge R(y,z)) \supset R(x,z)) \wedge \forall x \exists y R(x,y)$. As can easy be argued, irrespective of the interpretation of R, for every finite *k*, α is false. Hence α is not *k*-satisfiable. Then $\neg\alpha$ is *k*-valid.

Definition 12. *A formula* β *of* L_{FOL} *is a semantical consequence of a set* Γ of formulas of L_{FOL} *if and only if for any* $M = \langle D,i \rangle$ *any assignment* μ *in* M *which satisfies every formula in* Γ *also satisfies* β.

Remark. This definition, given for FOL, does not necessarily coincide with that given for PL (comp. Ch. 1, 2.6), since the formulas we are dealing with can be arbitrary formulas of L_{FOL}, i.e., closed or open. If Γ is a set of closed formulas (sentences) of L_{FOL} and β is also closed, then the definition of semantic consequence runs as follows: β is a semantic consequence of Γ (symbolically: $\Gamma \models \beta$) if and only if for any model M the following holds: if all the formulas of Γ are true in M, then β is true in M. Evidently, with all the formulas closed, this definition does coincide with that given for PL. But if open formulas are allowed, then we have to consider either the fact that the truth of a formula α in a model M means the truth of α in *any* assignment μ in M (cf. Def. 7), or just the Def. 12, we gave above.

2.2. Proofs of validity in FOL

Example 1. Prove that the following formula of L_{FOL} is valid:
$\gamma = \forall x(\alpha \supset \beta) \supset (\alpha \supset \forall x\beta)$; x *is not free in* α.

Proof (*reductio*). Suppose that γ is not valid. By Def. 9, it follows that there is a model $M = \langle D,i \rangle$ in which γ is not true. By Def. 7, it follows that there is an assignment μ in M such that $[\gamma]^{i,\mu} = 0$, i.e., (1) $[\forall x(\alpha \supset \beta) \supset (\alpha \supset \forall x\beta)]^{i,\mu} = 0$. Then by Def. 4 (item 3) we have: (2) $[\forall x(\alpha \supset \beta)]^{i,\mu} = 1$ and (3) $[\alpha \supset \forall x\beta]^{i,\mu} = 0$. In the same way, from (3) it follows (4) $[\alpha]^{i,\mu} = 1$ and (5) $[\forall x\beta]^{i,\mu} = 0$. And from (5), by Def. 4 (item 4a) it follows (6) there is an assignment ν x-variant of μ such that $[\beta]^{i,\nu} = 0$. From (4) it follows (7) $[\alpha]^{i,\nu} = 1$, for any assignment ν x variant of μ, since x is not free in α. Then by (7) and (6) we get (8) *there is* an assignment ν x-variant of μ such that $[\alpha \supset \beta]^{i,\nu} = 0$, by Def. 4 (item 3). But from (2) it follows, by Def. 4 (item 4a), (9) $[\alpha \supset \beta]^{i,\nu} = 1$, for *any* assignment ν x-variant of μ. But (9) and (8) are contradictory.

Remark. Formula γ above is valid only with the proviso: x *is not free in* α. Let us show that this restriction is necessary. In γ the formulas α and β are arbitrary, so let us take the case $\alpha = \beta = P(x)$. In this case γ is

$$\gamma^* = \forall x(P(x) \supset P(x)) \supset (P(x) \supset \forall x P(x)),$$

where x is *free* in α, contrary to the proviso. We'll show that there is a model M in which γ^* is not true.

Firstly, let us take a model $M_1 = \langle D,i \rangle$ whose domain contains only one member a, and then a model $M_2 = \langle D,i \rangle$ whose domain has two members, a_1 and a_2. In the first case the formulas γ^* can be given equivalently in the form

$$\gamma_1^* = (P(a) \supset P(a)) \supset (P(a) \supset P(a)),$$

and this formula cannot be falsified in $M_1 = \langle D,i \rangle$, regardless of the interpretation of P.[9] In the second case, when D has two members, γ^* can be given as

$$\gamma_2^* = ((P(a_1) \supset P(a_1)) \wedge (P(a_2) \supset P(a_2)) \supset (P(a_1) \supset (P(a_1) \wedge P(a_2))).$$

In the antecedent of the consequent of γ_2^*, $P(a_1)$, we chose for x the value a_1, but equally we could take a_2. Now if $P(a_1) = 1$ and $P(a_2) = 0$, as can be seen, $\gamma_2^* = 0$. Hence the formula γ^* is not valid. The model $M_2 = \langle D,i \rangle$ can explicitly be given by defining i and by taking two objects, a_1 and a_2, such that a_1 has the property P^i and a_2 not. If $P^i =$ "prime number" and if $a_1 = 3$ and $a_2 = 4$, then in $M = \langle \{3,4\}, \text{Prime} \rangle$, γ^* is not true.

Example 2. Prove the validity of the following formula of L_{FOL}:

δ: $\forall x(\alpha(x) \supset \beta(x)) \supset (\forall x \alpha(x) \supset \forall x \beta(x))$.

Proof (*reductio*). Suppose that δ is not valid. Hence, similar to the first example, by Def. 9 and Def. 7, there is a model $M = \langle D,i \rangle$ and an assignment μ in M such that

(1) $[\forall x(\alpha(x) \supset \beta(x)) \supset (\forall x \alpha(x) \supset \forall x \beta(x))]^{i,\mu} = 0$, hence

(2) $[\forall x(\alpha(x) \supset \beta(x))]^{i,\mu} = 1$; (1) Def. 4, item 3

(3) $[\forall x \alpha(x) \supset \forall x \beta(x)]^{i,\mu} = 0$; (1) Def. 4, item 3

(4) $[\forall x \alpha(x)]^{i,\mu} = 1$; (3) Def. 4, item 3

(5) $[\forall x \beta(x)]^{i,\mu} = 0$; (3) Def. 4, item 3

(6) $[\beta(x)]^{i,\nu} = 0$, for some assignment ν x-variant of μ; (5) Def. 4, item 4a)

(7) $[\alpha(x)]^{i,\nu} = 1$; for any assignment ν x-variant of μ, (4) Def. 4, item 4a)

---

[9] I.e., γ_1^* is 1-valid.

(8) $[\alpha(x)\supset\beta(x)]^{i,\nu}=0$, for some assignment ν x-variant of μ; (7), (6) Def. 4, item 3

(9) $[\alpha(x)\supset\beta(x)]^{i,\nu}=1$ for any assignment ν x-variant of μ; (2) Def. 4, item 4a)

But (8) and (9) are contradictory.

Example 3. Prove the validity of the following formulas:

1. $\forall x\alpha(x) \equiv \neg\exists x\neg\alpha(x)$
2. $\neg\forall x\alpha(x) \equiv \exists x\neg\alpha(x)$
3. $\exists x\alpha(x) \equiv \neg\forall x\neg\alpha(x)$
4. $\neg\exists x\alpha(x) \equiv\forall x\neg\alpha(x)$.

Proof. The proof of 1 can be reduced to the proof of the two conditionals:

1(a) $\models \forall x\alpha(x)\supset\neg\exists x\neg\alpha(x)$

1(b) $\models \neg\exists x\neg\alpha(x)\supset\forall x\alpha(x)$.

1(a) (*Reductio*). Let M and μ be arbitrary. Suppose that (1) $[\forall x\alpha(x)\supset\neg\exists x\neg\alpha(x)]^{i,\mu} = 0$. I.e., (2) $[\forall x\alpha(x)]^{i,\mu} = 1$ and (3) $[\neg\exists x\neg\alpha(x)]^{i,\mu} = 0$. From (2) it follows that (4) $[\alpha(x)]^{i,\nu} = 1$ for any ν x-var. of μ, and from (3) it follows (equivalently) that (5) $[\exists x\neg\alpha(x)]^{i,\mu} = 1$, and then (6) there is a ν x-var. of μ such that $[\neg\alpha(x)]^{i,\nu} = 1$, equivalently, (7) there is a ν such that $[\alpha(x)]^{i,\nu} = 0$. But (4) and (7) are contradictory.

1(b) (*Reductio*). As above, suppose that for an assignment μ: (1) $[\neg\exists x\neg\alpha(x)\supset\forall x\alpha(x)]^{i,\mu} = 0$. And then (2) $[\neg\exists x\neg\alpha(x)]^{i,\mu} = 1$ and (3) $[\forall x\alpha(x)]^{i,\mu} = 0$. It follows that (4) $[\exists x\neg\alpha(x)]^{i,\mu} = 0$ (from (2)), and (5) there is an assignment ν: $[\alpha(x)]^{i,\nu} = 0$ (from (3)). But from (4) it follows that (6) $[\neg\alpha(x)]^{i,\nu} = 0$, for any assignment ν x-variant of μ, equivalently (7) $[\alpha(x)]^{i,\nu} = 1$ for any assignment ν x-variant of μ, contradicting (5).

Now, 2 follows from 1 by PL and 3 and 4 follow from 2 and 1 taking $\neg\alpha(x)$ instead of $\alpha(x)$ (exercise).

2.3. Basic theorems in the semantics of FOL

In this section we prove some basic theorems[10] involving the semantic notions, defined in 2.1.

Theorem 1. *Let* $M = \langle D,i\rangle$ *be a model of* L_{FOL}, *let* μ *be an assignment in* M *and let* α *be a formula of* L_{FOL}. *Then*

[10] Actually, they are *meta*-theorems, since they talk *about* what happens in the object language of FOL, though we'll call them also theorems.

a) *If* $[\alpha]^{i,\mu}=1$ *and* $[\alpha\supset\beta]^{i,\mu}=1$, *then* $[\beta]^{i,\mu}=1$.

b) *If* α *and* $\alpha\supset\beta$ *are true in* M *then* β *is true in* M.

c) *If* $\models\alpha$, *and* $\models\alpha\supset\beta$, *then* $\models\beta$.

Proof. a) (*reductio*). Suppose that $[\alpha]^{i,\mu}=1$, $[\alpha\supset\beta]^{i,\mu}=1$ and $[\beta]^{i,\mu}=0$. From the first assumption and the third it follows that $[\alpha\supset\beta]^{i,\mu}=0$, contradicting the second assumption.

b) follows from a), by 2.1, Def. 7.

c) follows from b), by 2.1, Def. 9.

Similarly, we could firstly prove c), from which b) and a) follows, *via* Def. 9 and Def. 7.

Theorem 2. $\models\forall x\alpha(x)$ *iff* $\models\alpha(x)$.

What we have to prove are the following conditionals:

a) If $\models\forall x\alpha(x)$, then $\models\alpha(x)$.

b) If $\models\alpha(x)$, then $\models\forall x\alpha(x)$ (the converse of a)).

Proof. a) Suppose M and μ arbitrary, and that (1) $[\forall x\alpha(x)]^{i,\mu}=1$. Then (2) for *any* assignment ν x-variant of μ: $[\alpha(x)]^{i,\nu}=1$; whence $[\alpha(x)]^{i,\mu}=1$. Since M and μ are arbitrary, the result a) follows. (A proof of a) can also be given using Corollary 1 of Th. 7 (below) and Th. 1c); exercise).

b) (contraposition). Suppose that $\not\models\forall x\alpha(x)$. Hence there is a model $M=\langle D,i\rangle$ and an assignment μ in M such that $[\forall x\alpha(x)]^{i,\mu}=0$, by 2.1, Def. 9 and Def. 7. So, there is an assignment ν x-variant of μ such that $[\alpha(x)]^{i,\nu}=0$. Then, by Def. 7 and Def. 9 of 2.1, $\not\models\alpha(x)$.

By this theorem it follows that a formula α of L_{FOL} is true in a model $M=\langle D,i\rangle$ if and only if its *closure*,[11] $\forall x\alpha$, is true in M.

Theorem 3. $\exists x\alpha$ *is satisfiable iff* α *is satisfiable*.

Proof. Suppose that $M=\langle D,i\rangle$ is a model of L_{FOL} and μ an assignment in M. Suppose that $[\exists x\alpha]^{i,\mu}=1$, equivalent,[12] $[\neg\forall x\neg\alpha]^{i,\mu}=1$, and then $[\forall x\neg\alpha]^{i,\mu}=0$. It follows[13] that there is an assignment ν x-variant of μ such that $[\neg\alpha]^{i,\nu}=0$; hence $[\alpha]^{i,\nu}=1$ and then α is satisfiable, by 2.1, Def. 6.

For the proof of the converse, we take its contrapositive: if

[11] The closure of α is the closed formula obtained from α by prefixing it with universal quantifier for all the free variables of α.

[12] Comp. 3.3, Th. 1, a), for the syntactic form.

[13] We do not mention all the respective definitions anymore.

$[\exists x\alpha]^{i,\mu} = 0$, then $[\alpha]^{i,\mu} = 0$. Suppose now that $[\exists x\alpha]^{i,\mu} = 0$, equivalent $[\neg\forall x\neg\alpha]^{i,\mu} = 0$ and then $[\forall x\neg\alpha]^{i,\mu} = 1$ and $[\neg\alpha]^{i,\nu} = 1$, respectively, for any assignment ν x-variant of μ. So, for any ν $[\alpha]^{i,\nu} = 0$; whence $[\alpha]^{i,\mu} = 0$.

Theorem 4. *Let* $\alpha(p_1,...,p_n)$ *be a formula of propositional logic, containing the propositional variables* $p_1,...,p_n$. *Let* φ *be a mapping from* p_i, $1 \leq i \leq n$, *to the set of formulas of* L_{FOL}. *Let* α^* *be a formula of* L_{FOL} *resulting from* α *by substituting the formulas* $\varphi(p_i)$ *of* L_{FOL} *for* p_i. *Then*

If $\models\alpha$, *then* $\models\alpha^*$.

Proof (*contraposition*). We have to prove: if $\not\models\alpha^*$, then $\not\models\alpha$. So, let us suppose that $\not\models\alpha^*$. Hence there is a model $M = \langle D,i\rangle$ and an assignment μ in M such that $[\alpha^*]^{i,\mu} = 0$. We show that in this case there is an interpretation *int* of propositional variables p_i of α such that $[\alpha]^{int} = 0$. By using the mappings i and φ we define such an interpretation, thus:

$[p_i]^{int} = [\varphi(p_i)]^{i,\mu}$, for every p_i of α.

Then the following holds: $[\alpha]^{int} = [\alpha^*]^{i,\mu}$, hence if $[\alpha^*]^{i,\mu} = 0$, then $[\alpha]^{int} = 0$. And this holds for any M. Therefore if $\models\alpha$, then $\models\alpha^*$.

Remark. The converse of this theorem does not hold generally, since, for example, $\models\alpha^* = \forall x(\beta(x)\vee\neg\beta(x))$, but there is no formula of propositional logic from which it can be obtained by substitution, in the way indicated in the theorem. But if α^* contains no quantifier, then α and φ can always be defined in such a way that to any propositional variable of α corresponds, by φ, exactly one atomic formula of L_{FOL}. And in this case, the converse of this theorem also holds. Let us see.

Theorem 4*. *Let* α^* *be a formula of* L_{FOL} *without quantifiers. Then the following holds*:

$\models\alpha$ *iff* $\models\alpha^*$.

Proof. What is still to be proved is the following thing: if $\models\alpha^*$, then $\models\alpha$, or equivalently, if $\not\models\alpha$, then $\not\models\alpha^*$. We have to show therefore that if there is an interpretation *int* of propositional variables of α such that $[\alpha]^{int} = 0$, then there is a model $M = \langle D,i\rangle$ and an assignment μ in M such that $[\alpha^*]^{i,\mu} = 0$.

Let φ be $1-1$ mapping of $P_i(t_1,...,t_n)$ to the set of propositional variables of α, such that (*) $\varphi(P_i(t_1,...,t_n)) = p_i$ for each atomic formula $P_i(t_1,...,t_n)$ of α^*.

Such a model $M = \langle D,i \rangle$ can be defined in the following way. Let D be the set of all terms occurring in α^*, with the mention that any subterm of a term t counts as a distinct term. Let μ be an assignment in M such that for any $t \in D$ holds: $t^{i,\mu} = t$. That is

1. $x^\mu = x$, for every $x \in D$.
2. $c^i = c$, for every $c \in D$.
3. If $f(t_1,...,t_n) \in D$ and $t_j^{i,\mu} = t_j$, $1 \leq j \leq n$, then $f^i(t_1,...,t_n) = f(t_1,...,t_n)$.

Then, for any atomic formula $P(t_1,...,t_n)$, if $t_1,...,t_n$ are terms in D, with $t_j^{i,\mu} = t_j$, we set $[P(t_1,...,t_n)]^{i,\mu} = [\varphi(P(t_1,...,t_n))]^{int}$. Evidently, holds: $[\alpha^*]^{i,\mu} = [\alpha]^{int}$. And then: if $\models \alpha^*$, then $\models \alpha$.

Some remarks should be inserted.

Remark 1. The above definition $t^{i,\mu} = t$, for any t in α^* is motivated by induction on the length of t, lh(t).[14] If lh(t) = 1, then we have the cases 1 and 2 above. And if lh(t)>1, then t has the form $f(t_1,...,t_n)$, with $lh(t_j)<lh(t)$, $1 \leq j \leq n$. Then $t^{i,\mu} = (f(t_1,...,t_n))^{i,\mu} = f^i(t_1^{i,\mu},...,t_n^{i,\mu}) = f^i(t_1,...,t_n)$, and then by 3 $f(t_1,...,t_n)$.

Remark 2. Theorem 4* is the base of an important result in FOL: *decidability* of FOL, for formulas not containing quantifiers in the following sense. By the equivalence of this theorem follows that α *is satisfiable iff* α^* *is satisfiable*. But since α^* does not contain quantifiers and the number of terms in α is *k*, it follows: α^* *is satisfiable iff* α^* *is k-satisfiable*. We have, therefore, a *numerical* criterion for the satisfiability of α^*. And by this last equivalence we get: $\models \alpha^*$ iff α^* is *k*-valid.

Therefore, if α^* is a formula of L_{FOL} without quantifiers, then by Theorem 4*, α^* is *effectively decidable*. *Via* Remark 2, α^* is *k*-decidable.[15]

Theorem 5. *Let* $M = \langle D,i \rangle$ *be a model for* L_{FOL}, *let* α *be a formula of* L_{FOL}, *let* μ *and* μ' *two assignment in* M *such that for every free variable* x *in* α *holds*: $x^\mu = x^{\mu'}$. *Then the following holds*:

$$[\alpha]^{i,\mu} = 1 \text{ iff } [\alpha]^{i,\mu'} = 1.$$

Proof (induction on the complexity of α[16])

[14] The length of a term t is the number of occurrences of individual, constant and function symbols in t.

[15] Comp. also Sect. 5.1.1 and 5.1.2 (below).

[16] By "complexity of α" in L_{FOL} is meant the number of occurrences of connectives and quantifies in α. Let compl(α) be this number.

Basis. $n = 0$. α is an atomic formula $R(t_1,...,t_m)$, with R an *m*-place predicate symbol and $t_1,...,t_m$ terms. We have

$$[R(t_1,...,t_m)]^{i,\mu} = 1 \text{ iff } \langle t_1^{i,\mu},...,t_m^{i,\mu}\rangle \in R^i \text{, and}$$
$$[R(t_1,...,t_m)]^{i,\mu'} = 1 \text{ iff } \langle t_1^{i,\mu'},...,t_m^{i,\mu'}\rangle \in R^i.$$

We must show that for any j: $t_j^{i,\mu} = t_j^{i,\mu'}$. We have

a) If t_j is a constant c, then, evidently, $t_j^{i,\mu} = t_j^{i,\mu'} = c^i$.

b) If t_j is a variable x, then $t_j^{i,\mu} = t_j^{i,\mu'}$, by hypothesis, since $t_j^{i,\mu} = x^\mu = x^{\mu'} = t_j^{i,\mu'}$.

c) If t_j has the form $f(t_{j_1},...,t_{j_k})$, then we have again three cases, as a term t_{j_m}, $1 \leq m \leq k$, is a constant, a variable or a function symbol. The first two cases are a) and b) above. For the last case we have to remark the following fact: since a function f^i associated by the interpretation i to the function symbol f does not depend in any way on any assignment, it follows that the values of f^i will be the same for the same arguments. And then we shall have: $t_j^{i,\mu} = t_j^{i,\mu'}$.

Induction. Suppose that the theorem holds for any formula α such that $\text{compl}(\alpha) < n$. We have to show that the theorem holds for any α such that $\text{compl}(\alpha) = n$.

We only consider the following 3 cases:[17]

a) $\alpha = \neg\beta$
b) $\alpha = \alpha\supset\beta$
c) $\alpha = \forall x\beta$.

a) $\alpha = \neg\beta$. We have $[\neg\beta]^{i,\mu} = 1$ iff $[\beta]^{i,\mu} = 0$ iff $[\beta]^{i,\mu'} = 0$ (by hypothesis of induction) iff $[\neg\beta]^{i,\mu'} = 1$.

b) $\alpha = \beta\supset\gamma$. We have $[\beta\supset\gamma]^{i,\mu} = 1$ iff $[\beta]^{i,\mu} = 0$ or $[\gamma]^{i,\mu} = 1$, iff $[\beta]^{i,\mu'} = 0$ or $[\gamma]^{i,\mu'} = 1$ (by hypothesis of induction), iff $[\beta\supset\gamma]^{i,\mu'} = 1$, by 2.1, Def. 4, item 3.

c) $\alpha = \forall x\beta$. Suppose that $[\forall x\beta]^{i,\mu} = 1$, then $[\beta]^{i,\nu} = 1$ for any ν x-variant of μ. Suppose that x is the only bound variable of α.[18] By hypothesis for any variable y of α, different from x, we have $y^\mu = y^{\mu'}$, and then $y^\nu = y^{\nu'}$,

---

[17] For the propositional logic a) and b) are sufficient, given the completeness of the set $M = \{\neg,\supset\}$; cf. Ch. 1, 2.3.

[18] The opposite case is a simple generalization of this case.

where ν and ν' are x-variants of the assignments μ and μ', respectively. But compl(β)<compl(α), and then, by hypothesis of induction, $[\beta]^{i,\nu}=1$ iff $[\beta]^{i,\nu'}=1$. As, by our supposition, we have $[\beta]^{i,\nu}=1$ for any ν x-variant of μ, it follows that $[\beta]^{i,\nu'}=1$ for any ν' x-variant of μ'. And then $[\forall x\beta]^{i,\mu'}=1$. Hence if $[\alpha]^{i,\mu}=1$, then $[\alpha]^{i,\mu'}=1$. In a similar way we can show that if $[\alpha]^{i,\mu'}=1$, then $[\alpha]^{i,\mu}=1$.

Theorem 6. *Let* t *and* u *be terms, let* t' *be the term resulting from* t *by substituting* u *for the variable* x. *Let* μ *and* μ' *be two assignments such that* $x^{\mu'}=u^{i,\mu}$. *Then* $t'^{i,\mu}=t^{i,\mu'}$.

Proof (induction on the length of t)

Basis. lh(t) = 1, i.e., t is either a variable different or not from x or a constant. The theorem holds for these cases (trivial).

Induction. Assume the theorem holds for any term t with lh(t)<*n*, and show that it holds for any t such that lh(t) = *n*.

The case *n* = 1 occurs in *Basis*, so let us suppose that *n*>1 and thus t is a term of the form $f(t_1,...,t_m)$, where $lh(t_1)<n,...,lh(t_m)<n$. According to hypotheses of the theorem, t' is obtained from t by substituting the term u for x. Hence $t' = f(t'_1,...,t'_m)$. And thus $(f(t'_1,...,t'_m))^{i,\mu} = f^i(t'^{i,\mu}_1,...,t'^{i,\mu}_m)$. But $t'^{i,\mu}_j = t^{i,\mu'}_j$, for any $1 \le j \le m$, by hypotheses of induction. And then $f^i(t'^{i,\mu}_1,...,t'^{i,\mu}_m) = f^i(t^{i,\mu'}_1,...,t^{i,\mu'}_m) = (f(t_1,...,t_m))^{i,\mu'}$, by 2.1, Def. 2.

Theorem 7. *Let* α(x) *be a formula of* L_{FOL}, *let* t *be a term free for* x *in* α(x), *let* μ *and* μ' *two assignments such that* $x^{\mu'}=t^{i,\mu}$. *Then*

$$[\alpha(t/x)]^{i,\mu}=1 \text{ iff } [\alpha(x)]^{i,\mu'}=1.$$

Proof (induction on the complexity of α)

Basis. α(x) is an atomic formula, i.e., a formula of the form $R(t_1,...,t_m)$, and contains the free variable x. Let t'_j ($1 \le j \le m$) be the term resulting from t_j by substituting t for x. By 2.1, Def. 4, item 1, we have

$$[\alpha(t/x)]^{i,\mu} = 1 \text{ iff } \langle t'^{i,\mu}_1,...,t'^{i,\mu}_m\rangle \in R^i, \text{ and}$$

$$[\alpha(x)]^{i,\mu'} = 1 \text{ iff } \langle t^{i,\mu'}_1,...,t^{i,\mu'}_m\rangle \in R^i.$$

But, by Theorem 6, if μ and μ' are such that $x^{\mu'}=t^{i,\mu}$, then $t'^{i,\mu}_j=t^{i,\mu'}_j$, for $1 \le j \le m$. And then, we have $[\alpha(t/x)]^{i,\mu}=1$ iff $[\alpha(x)]^{i,\mu'}=1$.

Induction. Suppose that the theorem holds for any formula α such that compl(α)<*n*, and show that it holds for any α such that compl(α) = *n*.

Again, we'll have 3 cases:

a) $\alpha = \neg\beta$

b) $\alpha = \beta\supset\gamma$

c) $\alpha = \forall y\beta$.

For a) the argument runs in the following way. $\alpha(t/x)$ is then the formula $\neg\beta(t/x)$. And then $[\neg\beta(t/x)]^{i,\mu} = 1$ iff $[\beta(t/x)]^{i,\mu} = 0$. But $[\beta(t/x)]^{i,\mu} = 0$ iff $[\beta(x)]^{i,\mu'} = 0$, by hypothesis of induction. But $[\beta(x)]^{i,\mu'} = 0$ iff $[\neg\beta(x)]^{i,\mu'} = 1$ iff $[\alpha(x)]^{i,\mu'} = 1$.

The argument for b) is similar to b) of Theorem 5.

The last case, c): $\alpha = \forall y\beta$. Then $\alpha(t/x) = \forall y\beta(t/x)$, with t free for x in β.

Suppose that x is free in α and then x is free in β too; hence $x \neq y$. Then $[\forall y\beta(t/x)]^{i,\mu} = 1$ iff $[\beta(t/x)]^{i,\nu} = 1$ for any assignment ν y-variant of μ iff $[\beta(x)]^{i,\nu'} = 1$ (by ind. hyp.). But the variable y does not occur in t, since t is free for x in α (by hypothesis). So, the semantic value $t^{i,\mu}$ does not depend on any assignment given to y, and then for any assignment ν y-variant of μ: $t^{i,\nu} = t^{i,\mu}$, and then $x^{\nu'} = t^{i,\mu}$. It follows that the assignment ν' such that $x^{\nu'} = t^{i,\mu}$ does coincide with any assignment y-variant of μ' such that $x^{\mu'} = t^{i,\mu}$. And, finally, since $[\beta(x)]^{i,\nu'} = 1$ for *any* ν' such that $x^{\nu'} = t^{i,\mu}$ it follows that $[\forall y\beta]^{i,\mu'} = 1$.

If x is not free in α, then the theorem follows by Theorem 5 (argue!).

Corollary 1. $\models \forall x\alpha(x)\supset\alpha(t/x)$; t *is free for* x *in* α.

Proof. Let $M = \langle D,i\rangle$ be a model for L_{FOL}, let μ be an assignment in M such that $[\forall x\alpha(x)]^{i,\mu} = 1$. It follows that $[\alpha(x)]^{i,\nu} = 1$ for *any* assignment ν x-variant of μ. Then let ν be an assignment such that $x^{\nu} = t^{i,\mu}$. By the above theorem, we have $[\alpha(t/x)]^{i,\mu} = 1$. But M and μ are arbitrary, from which the result of corollary follows.

In a similar way, we have $\models \forall x\alpha(x)\supset\alpha(x)$, this being a particular case, $t = x$, of this corollary.

Corollary 2. (Subtitution of an individual variable). *If* $\models \alpha(x)$, *then* $\models \alpha(t/x)$, t *is free for* x *in* $\alpha(x)$.

Proof. As we know (by 2.3, Theorem 2), $\models \alpha(x)$ iff $\models \forall x\alpha(x)$. And by Corollary 1, $\models \forall x\alpha(x)\supset\alpha(t/x)$. It follows that If $\models \alpha(x)$, then $\models \alpha(t/x)$.

Theorem 8. $\models \alpha(x)\supset\exists x\alpha(x)$.

Proof (*reductio*). Suppose that $\not\models \alpha(x)\supset\exists x\alpha(x)$. Then there is an $M = \langle D,i\rangle$

and an assignment ν in M such that (1) $[\alpha(x)\supset\exists x\alpha(x)]^{i,\mu}=0$ and then (2) $[\alpha(x)]^{i,\mu}=1$ and (3) $[\exists x\alpha(x)]^{i,\mu}=0$. From (3) we get (4) $[\alpha(x)]^{i,\nu}=0$ for *any* ν x-variant of μ. Let ν be such that $x^{\nu}=t^{i,\mu}(=x^{\mu})$. Then by Theorem 7, we derive (5) $[\alpha(x)]^{i,\mu}=0$, contrary to (2).

Theorem 9. $\models\forall x\alpha(x)\supset\exists x\alpha(x)$.

Proof. Theorem 7 (Corollary 1), Theorem 8.

Theorem 10. If $\models\alpha(x)\supset\beta$, then $\models\forall x\alpha(x)\supset\beta$.

Proof (*contraposition*). Suppose that (1) $\not\models\forall x\alpha(x)\supset\beta$. Then there is an $M=\langle D,i\rangle$ and an assignment μ in M such that (2) $[\forall x\alpha(x)\supset\beta]^{i,\mu}=0$, so (3) $[\forall x\alpha(x)]^{i,\mu}=1$ and (4) $[\beta]^{i,\mu}=0$. From (3) follows (5) $[\alpha(x)]^{i,\nu}=1$ for *any* assignment ν x-variant of μ. So, there is an assignment ν x-variant of μ such that $x^{\nu}=x^{\mu}$. And then we have (6) $[\alpha(x)]^{i,\mu}=1$, by Theorem 7. Therefore, (7) $[\alpha(x)\supset\beta]^{i,\mu}=0$, from (4) and (6). Hence, finally, $\not\models\alpha(x)\supset\beta$.

Remark. A proof of this theorem could be given using Corollary 1 of Theorem 7 (exercise). Some other proof of this theorem can be given using Theorem 4, Corollary 1 of Theorem 7 and *modus ponens* (exercise).

Theorem 11. *If* $\models\beta\supset\alpha(x)$, *then* $\models\beta\supset\forall x\alpha(x)$; x *is not free in* β.

Proof (*reductio*, *contraposition*) (exercise).

Theorem 12. *If* $\models\beta\supset\alpha(x)$, *then* $\models\beta\supset\exists x\alpha(x)$.

Proof (contraposition). Suppose that (1) $\not\models\beta\supset\exists x\alpha(x)$; hence there is a model $M=\langle D,i\rangle$ and an assignment μ in M such that (2) $[\beta\supset\exists x\alpha(x)]^{i,\mu}=0$, i.e., (3) $[\beta]^{i,\mu}=1$ and (4) $[\exists x\alpha(x)]^{i,\mu}=0$. From (4) it follows (5) $[\alpha(x)]^{i,\nu}=0$, for any assignment ν x-variant of μ, and then it will also be false for the assignment ν such that $x^{\nu}=x^{\mu}$, i.e., (6) $[\alpha(x)]^{i,\mu}=0$. Hence, by (3) and (6), it follows (7) $[\beta\supset\alpha(x)]^{i,\mu}=0$, and then (8) $\not\models\beta\supset\alpha(x)$.

Theorem 13. *If* $\models\alpha(x)\supset\beta$, *then* $\models\exists x\alpha(x)\supset\beta$; x *is not free in* β.

Proof (*reductio*, *contraposition*) (exercise).

Remark. Theorem 13 can be obtained directly from Theorem 11, and Theorem 12 from Theorem 10 by passing from β and α(x) to ¬β and ¬α(x), respectively, and by replacing the so obtained implications with their contrapositive.

2.4. Duality in FOL

As we saw in Ch. 1, 2.5, the *dual* α^* of a formula α of L_{PL} which contains only the Boolean connectives $(\neg,\wedge,\vee)$ can be obtained by interchanging the binary connectives "$\wedge$" and "$\vee$" in all of their occurrences in α. As can be expected, since "$\wedge$" and "$\vee$" are dual each other, then the quantifiers "$\forall$" and "$\exists$" will also be called dual. Hence to that pair $(\wedge,\vee)$ of dual connectives we add now the pair $(\forall,\exists)$ and take again the main results of that section.

Theorem.[19] *Let* $\alpha(P_1,...,P_n)$ *be a formula of* L_{FOL} *in which every* P_i, $1\leq i\leq n$, *is an atomic formula, negated or unnegated and which contains only Boolean connectives and quantifiers. Let* α^* *be the formula resulting from* α *by replacing the connectives and quantifiers with their duals and by substituting* $\neg P_i$ *for* P_i. *Then*, $\models\neg\alpha\equiv\alpha^*$.

Proof. (1) Form $\neg\alpha$, and then remove successively the negation from the front of α to its atomic components, a process by which both $\wedge$ and $\vee$ and $\forall$ and $\exists$ are interchanged, and the atomic formulas are negated.

(2) Simplify the formula so obtained by eliminating the multiple negations. The result will be just the formula α^*.

Example. Let $\alpha = \exists x(\neg P(x)\vee Q(y,z))\wedge\forall zR(z)$. By Theorem, $\alpha^* = \forall x(P(x)\wedge\neg Q(y,z))\vee\exists z\neg R(z)$. This is true, since $\neg\alpha = \neg[\exists x(\neg P(x)\vee Q(y,z))\wedge\forall zR(z)]$, equivalent $\neg\exists x(\neg P(x)\vee Q(y,z))\vee\neg\forall zR(z)$, equivalent $\forall x\neg(\neg P(x)\vee Q(y,z))\vee\exists z\neg R(z)$, equivalent $\forall x(P(x)\wedge\neg Q(y,z))\vee\exists z\neg R(z)$.

Duality Theorem. 1. $\models\alpha$ *iff* $\models\neg\alpha^\delta$

2. $\models\alpha\supset\beta$ *iff* $\models\beta^\delta\supset\alpha^\delta$

3. $\models\alpha\equiv\beta$ *iff* $\models\alpha^\delta\equiv\beta^\delta$.

Proof.[20] Suppose, as before, that besides $\forall$ and $\exists$ the only connectives of α and β are $\neg$, $\wedge$ and $\vee$.

1a) If $\models\alpha$, then $\models\neg\alpha^\delta$.

(1) Assume $\models\alpha$. Then $\models\neg\alpha^*$; by Theorem and Ch. 1, 2.1.2, Th. 3.

(2) If $\models\neg\alpha^*$, then $\models\neg\alpha^{**}$, where α^{**} is obtained from α^* by substituting $\neg P_i$ for P_i (cf. 3.2.4.2, $Subst_P$) (below).

(3) If $\models\neg\alpha^{**}$, then $\models\neg\alpha^\delta$; by eliminating multiple negations.

---

[19] This theorem is an *extension* of the corresponding theorem in PL; comp. Ch. 1, Sect. 2.5.

[20] Similar to the proof of Duality Theorem in Ch. 1, 2.5. We take again only the first two cases.

Hence (4) if $\models\alpha$, then $\models\neg\alpha^{\delta}$; (1)-(3) PL.

1b) If $\models\neg\alpha^{\delta}$, then $\models\alpha$.

(1) Assume $\models\neg\alpha^{\delta}$.

(2) $\models\neg(\neg\alpha^{\delta})^{\delta}$; (1), Th. Dual. 1a).

(3) $\models\neg(\neg\alpha^{\delta\delta})$; (2).

(4) $\models\neg\neg\alpha^{\delta\delta}$; (3).

(5) $\models\alpha$; (4) PL, Ch. 1, 2.5, Remarks (1)

1. follows from 1a) and 1b).

2a) If $\models\alpha\supset\beta$, then $\models\beta^{\delta}\supset\alpha^{\delta}$.

(1) Assume $\models\alpha\supset\beta$.

(2) $\models\neg\alpha\vee\beta$; (1) PL.

(3) $\models\neg(\neg\alpha\vee\beta)^{\delta}$; Th. Dual. 1a).

(4) $\models\neg(\neg\alpha^{\delta}\wedge\beta^{\delta})$; (3), Ch. 1, 2.5, Remarks (3)

(5) $\models\beta^{\delta}\supset\alpha^{\delta}$; (4) PL.

2b) If $\models\beta^{\delta}\supset\alpha^{\delta}$, then $\models\alpha\supset\beta$.

(1) Assume $\models\beta^{\delta}\supset\alpha^{\delta}$.

(2) $\models\neg\beta^{\delta}\vee\alpha^{\delta}$; (1) PL.

(3) $\models\neg(\neg\beta^{\delta}\vee\alpha^{\delta})^{\delta}$; (2) Th. Dualit. 1a).

(4) $\models\neg(\neg\beta^{\delta\delta}\wedge\alpha^{\delta\delta})$; (3), Ch. 1, 2.5, Remarks (3)

(5) $\models\alpha^{\delta\delta}\supset\beta^{\delta\delta}$; (4) PL.

(6) $\models\alpha\supset\beta$; (5), Ch. 1, 2.5, Remarks (1)

2 follows from 2a) and 2b).

For 3, compare Ch. 1, 2.5, Duality Theorem 3.

As evident, the Duality Theorem allows us to extend automatically the stock of valid formulas of L_{FOL}. For, since the following formulas of L_{FOL} are valid:

α_1: $\exists x(P(x)\wedge Q(x))\supset(\exists xP(x)\wedge\exists xQ(x))$

α_2: $\forall x(P(x)\wedge Q(x))\equiv(\forall xP(x)\wedge\forall xQ(x))$,

the Duality Theorem will guarantee that the following formulas will be valid as well:

α_1^*: $(\forall xP(x)\vee\forall xQ(x))\supset\forall x(P(x)\vee Q(x))$

α_2^*: $\exists x(P(x)\vee Q(x))\equiv(\exists xP(x)\vee\exists xQ(x))$.

2.5. Interpolation in FOL

Let $\alpha\supset\beta$ be a closed formula of L_{FOL}. A closed formula γ of L_{FOL} is an *interpolant* for the implication $\alpha\supset\beta$ if every predicate symbol, function symbol and constant symbol occurring in γ also occurs in both α and β and the following holds: $\models\alpha\supset\gamma$ and $\models\gamma\supset\beta$.

Interpolation Theorem (in FOL). *If* $\models\alpha\supset\beta$ (*where* α *and* β *are closed*), *then* $\alpha\supset\beta$ *has an interpolant*.

Proof (*reductio*). Similar to the proof for PL (comp. Ch. 1, Sect. 2.9), suppose that $\models\alpha\supset\beta$ and that $\alpha\supset\beta$ has no interpolant. Let $S = \{\alpha,\neg\beta\}$ with the partition $S_1 = \{\alpha\}$ and $S_2 = \{\neg\beta\}$. Now, the (closed) formula $\mathrm{Conj}(\alpha)\supset\neg\mathrm{Conj}(\neg\beta)$ has no interpolant (since otherwise it would be also an interpolant for $\alpha\supset\beta$, contra our supposition). Hence S is Craig-consistent (comp. 2.9 of Ch. 1), and therefore it has a model M[21], i.e., α and $\neg\beta$ are true in M, and then $\beta = 0$ in M. This implies that $\alpha\supset\beta = 0$ in M, and therefore $\not\models\alpha\supset\beta$, contrary to our supposition.

2.6. Beth's Definability Theorem[22]

Beth's Definability Theorem concerns the relation between two ways of defining a notion in terms of other notions, with respect to a set of closed formulas Γ.[23] If the definition of this notion is a semantic consequence of Γ, then we speak of an *explicit definition*. And if Γ uniquely determine (define) that notion, then we speak of *implicit definition*. Let us define these two ways of definition. Let Q^n be an *n*-place predicate symbol to be defined.

Definition 1. *Let* $\mathrm{FORM}(x_1,\ldots,x_n)$ *be a formula of* L_{FOL} *with the specified free variables, and not containing* Q^n. *Let* Γ *be a set of closed formulas of* L_{FOL}. $\mathrm{FORM}(x_1,\ldots,x_n)$ *is an explicit definition of* Q^n *with respect to* Γ *if the following holds:*

$$\Gamma\models\forall x_1\ldots\forall x_n(Q(x_1,\ldots,x_n)\equiv\mathrm{FORM}(x_1,\ldots,x_n)).$$

Definition 2. Q^n *is implicitly definable with respect to* Γ *if* Γ *determines* Q^n *uniquely, i.e.,*

---

[21] Cf. Sect. 3.5.4 Theorem (below).

[22] Cf. E. Beth, [1953].

[23] Or a theory *T*.

$$\Gamma \cup \Gamma^* \models \forall x_1 \ldots \forall x_n(Q(x_1,\ldots,x_n) \equiv Q^*(x_1,\ldots,x_n)),$$

where Q^* *is an n-place predicate symbol of* L_{FOL}, *different from* Q *and not occurring in* Γ, *and* Γ^* *arises from* Γ *by replacing everywhere* Q *with* Q^*.

Beth's Definability Theorem. *If* Q *is implicitly definable with respect to* Γ *then* Q *is explicitly definable with respect to* Γ.

Proof. Suppose Q is implicitly definable, and Γ, Q^*, Γ^* as in Def. 2; and then

$$\Gamma \cup \Gamma^* \models \forall x_1 \ldots \forall x_n(Q(x_1,\ldots,x_n) \equiv Q^*(x_1,\ldots,x_n)).$$

By the *finitism* of the relation "$\models$" [24] there are the finite sets Γ_0 and Γ_0^* of Γ and Γ^*, respectively, such that

(1) $\Gamma_0 \cup \Gamma_0^* \models \forall x_1 \ldots \forall x_n(Q(x_1,\ldots,x_n) \equiv Q^*(x_1,\ldots,x_n)).$

Let $\mathrm{Conj}(\Gamma_0)$ and $\mathrm{Conj}(\Gamma_0^*)$ be the conjunction of all formulas in Γ_0 and the conjunction of all formulas in Γ_0^*, respectively. Whence

(2) $\mathrm{Conj}(\Gamma_0) \wedge \mathrm{Conj}(\Gamma_0^*) \models \forall x_1 \ldots \forall x_n(Q(x_1,\ldots,x_n) \equiv Q^*(x_1,\ldots,x_n))$

(from (1) by Ch. 1, 2.6), and therefore

(3) $\models (\mathrm{Conj}(\Gamma_0) \wedge \mathrm{Conj}(\Gamma_0^*)) \supset \forall x_1 \ldots \forall x_n(Q(x_1,\ldots,x_n) \equiv Q^*(x_1,\ldots,x_n))$

from (2), by Ch. 1, 2.6, Normality Theorem, and Def. 12 Remark of Sect. 2.1 of this chapter.

Let $c_1,\ldots,c_n$ be distinct constant symbols. Then

(4) $\models (\mathrm{Conj}(\Gamma_0) \wedge \mathrm{Conj}(\Gamma_0^*)) \supset (Q(c_1,\ldots,c_n) \equiv Q^*(c_1,\ldots,c_n))$

by 2.3, Th. 7 (Corollary 1) and PL.

(5) $\models (\mathrm{Conj}(\Gamma_0) \wedge Q(c_1,\ldots,c_n) \supset (\mathrm{Conj}(\Gamma_0^*) \supset Q^*(c_1,\ldots,c_n))$

by PL.

Let $Ant = \mathrm{Conj}(\Gamma_0) \wedge Q(c_1,\ldots,c_n)$ (i.e., the antecedent of the implication in (5)), and $Conseq^* = \mathrm{Conj}(\Gamma_0^*) \supset Q^*(c_1,\ldots,c_n)$ (i.e., the consequent of the same implication).

Now, since the implication in (5) is valid, it has an interpolant γ (by Craig's Theorem) which (possibly) contain $c_1,\ldots,c_n$ (some or all of them). Let $\gamma = \mathrm{FORM}(c_1,\ldots,c_n)$. Since FORM is an interpolant, all the symbols

[24] By Sect. 2.1 Def. 12, and since Γ is a set of closed formulas of L_{FOL}, by Finitness Theorem of the semantic consequence in PL (cf. Ch. 1, Sect. 3.3.3).

occurring in it also occur in both formulas *Ant* and $Conseq^*$, and moreover

(a) $\models Ant \supset \mathrm{FORM}$, and

(b) $\models \mathrm{FORM} \supset Conseq^*$.

Observe that the predicate symbol Q^* does not occur in *Ant*, and correspondingly Q does not occur in $Conseq^*$, and then neither Q, nor Q^* is in FORM (since all the symbols in FORM must be in both *Ant* and $Conseq^*$).

(c) $\models \mathrm{FORM} \supset Conseq$; from (b), where *Conseq* is obtained from $Conseq^*$ by replacing Q^* with Q (since Q does not appear in $\mathrm{FORM} \supset Conseq^*$). In extenso, the expression (c) is just

$\models \mathrm{FORM}(c_1,...,c_n) \supset (\mathrm{Conj}(\Gamma_0) \supset Q(c_1,...,c_n))$.

(d) $\models \mathrm{Conj}(\Gamma_0) \supset (\mathrm{FORM}(c_1,...,c_n) \supset Q(c_1,...,c_n))$;

from (c) by PL (permutation of premises)

(e) $\models \mathrm{Conj}(\Gamma_0) \supset (Q(c_1,...,c_n) \supset \mathrm{FORM}(c_1,...,c_n))$;

from (a), by PL

(f) $\models \mathrm{Conj}(\Gamma_0) \supset (Q(c_1,...,c_n) \equiv \mathrm{FORM}(c_1,...,c_n))$;

from (d) and (e), by PL

(g) $\mathrm{Conj}(\Gamma_0) \models Q(c_1,...,c_n) \equiv \mathrm{FORM}(c_1,...,c_n)$;

from (f) by Normality Th. (comp. Ch. 1, 2.6)

(h) $\Gamma_0 \models Q(c_1,...,c_n) \equiv \mathrm{FORM}(c_1,...,c_n)$; (g) by Ch. 1, 2.6, Theorem

(i) $\Gamma \models Q(c_1,...,c_n) \equiv \mathrm{FORM}(c_1,...,c_n)$; (h), by Ch. 1, 2.6, Prop. 2

(j) $\Gamma \models \forall x_1 ... \forall x_n (Q(x_1,...,x_n) \equiv \mathrm{FORM}(x_1,...,x_n))$;

(The derivation of (j) from (i) is based on the following result of FOL: if Γ is a set of formulas of L_{FOL}, $\alpha(x)$ (with x free) is a formula of L_{FOL} and c is a constant symbol not occurring in the formulas of Γ and in $\alpha(x)$, then *if* $\Gamma \models \alpha(c/x)$, *then* $\Gamma \models \forall x \alpha(x)$ (easy arguable using the equivalence $\Gamma \models \alpha(c/x)$ iff $\Gamma \models \alpha(x)$ and Theorem 2 of Sect. 2.3; exercise)).

But $\mathrm{FORM}(x_1,...,x_n)$ does not contain any occurrence of the predicate symbol Q, and therefore $\mathrm{FORM}(x_1,...,x_n)$ defines explicitly Q.

3. FOL axiomatized (FOL^{ax})

3.1. An axiomatic system

An axiomatic system of the first-order logic will be a deductive construction whose starting point is a set of logical axioms and a finite set of rules of deduction (inference). If by α, β, γ we understand the formulas of L_{FOL}, then such an axiomatic construction of FOL can be expressed by the following:

Logical axioms

Ax1. $\alpha\supset(\beta\supset\alpha)$
Ax2. $(\alpha\supset(\beta\supset\gamma))\supset((\alpha\supset\beta)\supset(\alpha\supset\gamma))$
Ax3. $(\neg\beta\supset\neg\alpha)\supset(\alpha\supset\beta)$
Ax4. $\forall x\alpha(x)\supset\alpha(t/x)$, t is free for x in $\alpha(x)$
Ax5. $\forall x(\alpha\supset\beta)\supset(\alpha\supset\forall x\beta)$; x is not free in α.

Rules of deduction

Modus Ponens (MP) $\dfrac{\alpha,\alpha\supset\beta}{\beta}$; Generalization (Gen) $\dfrac{\alpha}{\forall x\alpha}$

Def ∃. $\exists x\alpha =_{df} \neg\forall x\neg\alpha$.

Definition 1. *A proof in* FOL^{ax} *is a finite sequence of formulas of* L_{FOL} *each one of which is either an axiom of* FOL^{ax} *or is an immediate consequence of an application of a deduction rule of* FOL^{ax} *to the preceding formulas of the sequence.*

Definition 2. *A formula* α *of* L_{FOL} *is a theorem of* FOL^{ax} (*symbolic*: $FOL^{ax}\vdash\alpha$[25]) *if and only if there is a proof in* FOL^{ax} *whose last formula is* α.

Definition 3. *A deduction in* FOL^{ax} *of a formula* α *from a set* Γ *of formulas is a* ***finite*** *sequence of formulas of* L_{FOL} *each one of which is either an axiom of* FOL^{ax} *or a member of* Γ *or is an immediate consequence of an application of a deduction rule of* FOL^{ax} *to the preceding formulas of the sequence.*

Definition 4. *A formula* α *of* L_{FOL} *is deducible in* FOL^{ax} *from a set* Γ *of formula* (*symbolic*: $\Gamma\vdash\alpha$), *if and only if there is a deduction in* FOL^{ax} *whose last formula is* α.

[25] Since the only axiomatic system we refer in this chapter is FOL^{ax}, often we omit this mention and we'll write simply $\vdash\alpha$.

Remark. Evidently, if $\Gamma = \varnothing$, then the *deduction* of β from Γ is passing in the *proof* of β in FOL^{ax}. That is, a formula α is a theorem of FOL^{ax} iff α is provable from zero premises. More generally, the set of theorems of FOL^{ax} does coincide with the set of formulas of L_{FOL} provable from zero premises.[26]

Definition 5. *A set* Γ *of formulas of* L_{FOL} *is* FOL^{ax}-*inconsistent* (*or simply inconsistent*) *if and only if anyone of the following holds*:

a) $\Gamma \vdash$ Form (*where* Form *is the set of all formulas of* L_{FOL}),

b) $\Gamma \vdash \beta$ *and* $\Gamma \vdash \neg\beta$, for some $\beta \in L_{FOL}$,

c) $\Gamma \vdash \beta \wedge \neg\beta$, *respectively*.

Definition 6. *A set* Γ *of formulas of* L_{FOL} *is* FOL^{ax}-*consistent* (*or simply consistent*) *if and only if* Γ *is not inconsistent. That is, if and only if anyone of the following holds*:

a) $\Gamma \nvdash$ Form (*equivalently, there is a formula* β *such that* $\Gamma \nvdash \beta$),

b) *for no* β: $\Gamma \vdash \beta$ *and* $\Gamma \vdash \neg\beta$, *or*

c) *for no* β: $\Gamma \vdash \beta \wedge \neg\beta$.[27]

3.2. Basic results on FOL^{ax}

3.2.1. $Rule_p$

As we saw in 2.3, Theorem 4, if α^* is a formula of L_{FOL} resulting from a formula α of L_{PL} by substitution of formulas of L_{FOL} for propositional variables of α, then if $\models \alpha$, then $\models \alpha^*$. Moreover, as the following theorem shows, if $\models \alpha$, then $\vdash \alpha^*$.

Theorem. *Let* $\alpha(p_1,...,p_n)$ *be a formula of* L_{PL} *containing the propositional variables* $p_1,...,p_n$. *Let* $\alpha^*(\beta_1/p_1,...,\beta_n/p_n)$ *be the formula of* L_{FOL} *coming from* α *by substitution of the formulas of* L_{FOL} $\beta_1,...,\beta_n$ *for* $p_1,...,p_n$, *respectively. Then the following holds: if* $\models \alpha$, *then* $\vdash \alpha^*$.

Proof. Assuming that $\models \alpha$, by the completeness theorem of PL^{ax} (Ch. 1, 3.3) it follows that $\vdash \alpha$. Now in the proof of α in PL^{ax} we make the same substitutions of formulas of L_{FOL} as were used in constructing α^* from α (in

---

[26] From which it follows that the consistency of the set of theorems of FOL^{ax} (let us call it *Th* until further notification) does coincide with the consistency of the empty set ($\varnothing$).

[27] If in a), b) and c) in Def. 5 and Def. 6 we set FOL^{ax} instead of Γ, this amounts to the definition of inconsistency and consistency of FOL^{ax}, respectively.

the way indicated in the theorem), and for the propositional variables in the proof which are not contained in α we substitute arbitrary formulas of L_{FOL}. What results is just a proof of α^* in FOL^{ax}, and this proof is given by using only Ax1-Ax3 and MP.

Let us refer in what follows to the applications of this theorem by $Rule_p$, expressing precisely the content of the theorem:

$$\frac{\models \alpha(p_1,\dots,p_n)}{\vdash \alpha^*(\beta_1/p_1,\dots,\beta_n/p_n)}$$

3.2.2. Consistency and soundness of FOL^{ax}

Theorem 1. FOL^{ax} *is consistent.*

Proof. All we have to show is that the set *Th* of theorems of FOL^{ax} is consistent. But it is clear that $\nvdash P(x)$, i.e., $\varnothing \nvdash P(x)$, and then the empty set $\varnothing$ is consistent, according to Def 6 above. Therefore *Th* is consistent (by Remark after Def. 4).

Theorem 2. FOL^{ax} *is sound; i.e., if* $\vdash \alpha$, *then* $\models \alpha$.

Proof. By 2.3, Theorem 4, Ax1-Ax3 are valid formulas of L_{FOL}. By 2.3, Theorem 7 (Corollary 1) Ax4 is valid. By 2.2, Example 1, Ax5 is also valid. By 2.3, Theorem 1.c) MP preserves validity, and by 2.3, Theorem 2, *Gen* also preserves validity. Therefore, every theorem of FOL^{ax} is a valid formula of L_{FOL}.

Remark. From soundness of FOL^{ax} an easy proof of consistency of FOL^{ax} follows. Since if $\vdash \alpha$, then $\models \alpha$, it follows that $\neg\alpha$ is not satisfiable, hence $\not\models \neg\alpha$, and then $\nvdash \neg\alpha$.

A still another proof of consistency of FOL^{ax} can be given by considering a model $M = \langle D,i \rangle$, whose domain D contains one member. As can be observed, in such a model all the axioms of FOL^{ax} are true and MP and *Gen* preserve validity. Hence all the theorems of FOL^{ax} are true in M. But no formula and its negation can be simultaneously true in M, hence for no formula α of L_{FOL} α and $\neg\alpha$ are both theorems of FOL^{ax}.

A variant of this argument is the following. By using the model $M = \langle D,i \rangle$ of cardinality 1, any theorem of FOL^{ax} has a propositional counterpart, a valid formula of L_{PL}, obtained by deleting all quantifiers in the theorems, all terms, commas and parenthesis and replacing each predicate symbol by a propositional variable. Now, P(x) is a formula of L_{FOL} that does not have a valid formula of L_{PL} as its counterpart. Hence it is not a theorem of FOL^{ax}, i.e., FOL^{ax} is consistent.

3.2.3. Deduction Theorem

As we saw in Ch. 1, 3.2.2 an important result concerning PL^{ax} is the deduction theorem. In FOL^{ax} we also have such a result, but here this theorem has a proviso whose meaning follows from the following considerations.

Let *Ded* = $\beta_1,...,\beta_n$, with $\beta_n = \beta$, be a deduction in FOL^{ax} of the formula β of L_{FOL} from the assumption formulas $\alpha_1,...,\alpha_m$. We say that a formula β_i ($1 \leq i \leq n$) *depends on* α_k ($1 \leq k \leq m$) in *Ded* if the following is the case: $\beta_i = \alpha_k$ or β_i is the conclusion of one application of MP or Gen from premises of which at least one depends on α_k.

A variable x is *fixed* in *Ded* for an assumption α_k if this is the case: x occurs free in α_k and *Ded* contains one application of Gen to a formula β_i depending on α_k by which x becomes bound. If in *Ded* x is fixed for an assumption, then we write $\alpha_1,...,\alpha_m \vdash_x \beta$.

Deduction Theorem.[28] *Assume that* $\alpha_1,...,\alpha_m \vdash \beta$ *and in this deduction no variable is fixed for* α_m. *Then* $\alpha_1,...,\alpha_{m-1} \vdash \alpha_m \supset \beta$.

Proof. Let *Ded* be the deduction from the hypothesis of the theorem: $\alpha_1,...,\alpha_m \vdash \beta$, obeying the proviso. We transform this deduction in a deduction *Ded** of the implication $\alpha_m \supset \beta$ from the assumptions $\alpha_1,...,\alpha_{m-1}$, by replacing all the formulas β_i ($1 \leq i \leq n$) in *Ded* with $\alpha_m \supset \beta_i$ and by inserting new formulas in the so obtained sequence of formulas,[29] let us call it *Seq*, in the following way:[30]

1. If $\beta_i = \alpha_m$, then we replace the formula $\alpha_m \supset \beta_i$ in *Seq* with the proof of $\alpha_m \supset \alpha_m$; cf. Ch. 1, 3.1, Th1.[31]

2. If $\beta_i = \alpha_k$ ($k \neq m$), then we replace the member $\alpha_m \supset \beta_i$ of *Seq* with

    $\alpha_k \supset (\alpha_m \supset \alpha_k)$; Ax1

    α_k, assumption

    $\alpha_m \supset \alpha_k$; MP.

3. If β_i is an axiom, then we replace the formula $\alpha_m \supset \beta_i$ of *Seq* with

    $\beta_i \supset (\alpha_m \supset \beta_i)$; Ax1

---

[28] Also known as Herbrand Theorem; cf. J. Herbrand [1930].

[29] In order to justify these implications.

[30] The first 4 cases of the proof are those of the corresponding proof in PL^{ax} (cf. Ch. 1, Sect. 3.2.2).

[31] Or, equivalently, before the formula $\alpha_m \supset \beta_i$ in *Seq* we set all the formulas from which $\alpha_m \supset \alpha_m$ follows; similar for the other cases; comp. Ch. 1, Sect. 3.2.2.

β_i; axiom

$\alpha_m \supset \beta_i$; MP.

4. If β_i results in *Ded* by one application of MP from formulas β_k and $\beta_k \supset \beta_i$. Hence before the formula $\alpha_m \supset \beta_i$ in *Seq* occur the formulas $\alpha_m \supset \beta_k$ and $\alpha_m \supset (\beta_k \supset \beta_i)$. Then we replace the formula $\alpha_m \supset \beta_i$ in *Seq* with

$(\alpha_m \supset (\beta_k \supset \beta_i)) \supset ((\alpha_m \supset \beta_k) \supset (\alpha_m \supset \beta_i))$; Ax2

$(\alpha_m \supset \beta_k) \supset (\alpha_m \supset \beta_i)$; MP

$\alpha_m \supset \beta_i$; MP.

5. If β_i results in *Ded* by one application of Gen from the formula β_j, then $\beta_i = \forall x \beta_j$, and we have to distinguish two cases:

a) the variable x is not free in the assumption α_m. In this case we replace the formula $\alpha_m \supset \beta_i$ in *Seq* with

$\alpha_m \supset \beta_j$

$\forall x(\alpha_m \supset \beta_j)$; Gen

$\forall x(\alpha_m \supset \beta_j) \supset (\alpha_m \supset \forall x \beta_j)$, Ax5

$\alpha_m \supset \forall x \beta_j$; MP

b) the variable x is free in the assumption α_m. In this case the formula $\alpha_m \supset \beta_i$ in *Seq* will be replaced by

β_j

$\forall x \beta_j$; Gen

$\forall x \beta_j \supset (\alpha_m \supset \forall x \beta_j)$; Ax1

$\alpha_m \supset \forall x \beta_j$; MP.

This derivation is justified by the fact that β_i occurs from β_j by Gen and in the derivation of β_i the variable x is not fixed for α_m (by hypothesis of Ded. Th.). Hence Gen has no application to a formula depending on α_m by which x (free in α_m) becomes bound. This means that the formula β_j does not depend on α_m, i.e., β_j can be derived from $\alpha_1,\ldots,\alpha_{m-1}$.

Proceeding in this way, from *Ded*, *via Seq*, we have obtained a deduction *Ded* * of the implication $\alpha_m \supset \beta$ from the assumptions $\alpha_1,\ldots,\alpha_{m-1}$.

Corollary. *Let Ded be a deduction of* β *from* $\alpha_1,\ldots,\alpha_m$, *i.e.,* $\alpha_1,\ldots,\alpha_m \vdash \beta$. *Then the following holds*:

a) *If* α_m *is closed, then* $\alpha_1,\ldots,\alpha_{m-1} \vdash \alpha_m \supset \beta$.

b) *If Ded involves no application of Gen to some free variable of* α_m, *then* $\alpha_1,\ldots,\alpha_{m-1} \vdash \alpha_m \supset \beta$.

3.2.4. Substitution Theorems

3.2.4.1. Substitution Theorem for variables

If $\alpha(x)$ is a formula of L_{FOL}, then by one application of Gen we obtain $\forall x\alpha(x)$. Hence $\alpha(x)\vdash_x\forall x\alpha(x)$. But in FOL^{ax} the following holds: If $\alpha\vdash\beta$, then: if $\vdash\alpha$, then $\vdash\beta$ (argue!). Using Ax4, $\vdash\forall x\alpha(x)\supset\alpha(t/x)$, we obtain the following theorem.

Substitution Theorem for variables ($\mathbf{Subst_x}$). *Let* $\alpha(x)$ *be a formula of* L_{FOL}, *let* $\alpha(t/x)$ *be the formula obtained from* $\alpha(x)$ *by substitution of* t *for* x (*where* t *is a term free for* x *in* $\alpha(x)$). *Then*

If $\vdash\alpha(x)$, *then* $\vdash\alpha(t/x)$ (*the syntactic counterpart of* 2.3, *Th.* 7, *Coroll.* 2).

3.2.4.2. Substitution Theorem for predicate symbols ($\mathbf{Subst_P}$)[32]

Let α be a formula of L_{FOL} containing the *n*-place predicate symbol P^n [33] (symbolic: $\alpha(P(v_1,...,v_n))$). Let $\beta(x_1,...,x_n)$ be a formula of L_{FOL} containing the free variable $x_1,...,x_n$. The other possibly free variables of β must not have any bound occurrence in α. Besides that, the variables $v_1,...,v_n$ of P must not occur bound in $\beta(x_1,...,x_n)$. Now we replace in α every atomic formula $P(v_1,...,v_n)$ by $\beta(v_1,...,v_n)$ and obtain α^* (symbolic: $\alpha^*(\beta(v_1,...,v_n)/P(v_1,...,v_n))$. Hence $\beta(v_1,...,v_n)$ comes from $\beta(x_1,...,x_n)$ by replacing x_i in all of its occurrences with v_i. Then the following theorem holds.

Substitution Theorem for predicate symbols ($\mathbf{Subst_P}$). *If* $\vdash\alpha(P(x_1,...,x_n))$, *then* $\vdash\alpha^*(\beta(v_1,...,v_n)/P(v_1,...,v_n))$.

Proof (as the proof in Ch. 1, 3.2.1.1).

Example. Let $\alpha(P(v_1,v_2)) = \exists x\forall y(P(y,y)\supset P(x,x))$.

Let $\beta(v_1,v_2) = \forall zQ(v_1,v_2,z)$.

Then $\alpha^*(\beta(v_1,v_2)/P(v_1,v_2)) = \exists x\forall y(\forall zQ(y,y,z)\supset\forall zQ(x,x,z))$.

---

[32] This is the Rule of Substitution α3) of Hilbert and Ackermann [1938]. But, as A. Church observed in his [1944], the rule is incompletely formulated, since an essential condition relative to bound variables is omitted. So is the case with Hilbert and Ackermann [1946]. In Hilbert and Ackermann [1972], Ch. 3, §5, XIII, this rule is correctly stated, as a theorem essentially in the form given here.

[33] Note that the same P^n may appear in α with different arguments. We indicate this fact by using $P(v_1,...,v_n)$, where each v_i can be an arbitrary variable.

3.2.5. Replacement Theorem

We prove, firstly, another important result in FOL^{ax}, Equivalence Theorem, from which Replacement Theorem follows as a corollary.

Equivalence Theorem. *Let* α_β *be a formula of* L_{FOL} *containing* β *as a subformula. Let* α_γ *be the formula obtained from* α_β *by replacing one or more occurrences of* β *by* γ. *Let us consider that* $x_1,...,x_n$ *are all the free variables of* β *and* γ *which are bound variables of* α. *Then*

$$\vdash \forall x_1...\forall x_n(\beta\equiv\gamma)\supset(\alpha_\beta\equiv\alpha_\gamma).$$

Proof (induction on the complexity of α).

Basis. $\beta = \alpha$. In this case by replacing β with γ, we obtain the formula of L_{FOL} $\forall x_1...\forall x_n(\beta\equiv\gamma)\supset(\beta\equiv\gamma)$, provable in FOL^{ax} (cf. 3.3, Th4, below).

Induction. Suppose that α has one of the following forms: 1. $\alpha = \neg\delta$, 2. $\alpha = \delta\supset\varepsilon$ and 3. $\alpha = \forall x\delta$, and that the theorem holds for δ and ε.

1. $\alpha = \neg\delta$; i.e., $\alpha_\beta = \neg\delta_\beta$. Then $\alpha_\gamma = \neg\delta_\gamma$. By inductive hypothesis, the theorem holds for δ. Hence $\vdash \forall x_1...\forall x_n(\beta\equiv\gamma)\supset(\delta_\beta\equiv\delta_\gamma)$. But $\vdash (\delta_\beta\equiv\delta_\gamma)\supset(\neg\delta_\beta\equiv\neg\delta_\gamma)$. Therefore $\vdash \forall x_1...\forall x_n(\beta\equiv\gamma)\supset(\neg\delta_\beta\equiv\neg\delta_\gamma)$, that is $\vdash \forall x_1...\forall x_n(\beta\equiv\gamma)\supset(\alpha_\beta\equiv\alpha_\gamma)$.

2. $\alpha = \delta\supset\varepsilon$; i.e., $\alpha_\beta = \delta_\beta\supset\varepsilon_\beta$. Then $\alpha_\gamma = \delta_\gamma\supset\varepsilon_\gamma$. By inductive hypothesis the theorem holds for δ and ε. Hence $\vdash \forall x_1...\forall x_n(\beta\equiv\gamma)\supset(\delta_\beta\equiv\delta_\gamma)$ and $\vdash \forall x_1...\forall x_n(\beta\equiv\gamma)\supset(\varepsilon_\beta\equiv\varepsilon_\gamma)$. It follows, by Rule$_p$ (of 3.2.1), that $\vdash \forall x_1...\forall x_n(\beta\equiv\gamma)\supset [(\delta_\beta\equiv\delta_\gamma)\wedge(\varepsilon_\beta\equiv\varepsilon_\gamma)]$. But $\vdash [(\delta_\beta\equiv\delta_\gamma)\wedge(\varepsilon_\beta\equiv\varepsilon_\gamma)]\supset$ $\supset[(\delta_\beta\supset\varepsilon_\beta)\equiv(\delta_\gamma\supset\varepsilon_\gamma)]$, by Rule$_p$. Hence $\vdash \forall x_1...\forall x_n(\beta\equiv\gamma)\supset[(\delta_\beta\supset\varepsilon_\beta)\equiv(\delta_\gamma\supset\varepsilon_\gamma)]$; i.e., $\vdash \forall x_1...\forall x_n(\beta\equiv\gamma)\supset(\alpha_\beta\equiv\alpha_\gamma)$.

3. $\alpha = \forall x\delta$; i.e., $\alpha_\beta = \forall x\delta_\beta$, where x is not free in $\forall x_1,...,x_n(\beta\equiv\gamma)$ (since if it were, then x would be free in β or γ, and by hypothesis it will be bound in α and then x will be one of $x_1,...,x_n$). By inductive hypothesis, $\forall x_1...\forall x_n(\beta\equiv\gamma)\supset(\delta_\beta\equiv\delta_\gamma)$, and then, using Ax5, $\forall x_1...\forall x_n(\beta\equiv\gamma)\supset\forall x(\delta_\beta\equiv\delta_\gamma)$. Finally, using 3.3, Th. 7 (below) and Rule$_p$, we obtain $\forall x_1...\forall x_n(\beta\equiv\gamma)\supset(\forall x\delta_\beta\equiv\forall x\delta_\gamma)$, i.e., $\forall x_1...\forall x_n(\beta\equiv\gamma)\supset(\alpha_\beta\equiv\alpha_\gamma)$.

Replacement Theorem (Repl$_{FOL}$). *Let* β, γ, α_β *and* α_γ *as in the preceding theorem. Then the following holds*:

If $\vdash \beta\equiv\gamma$, *then* $\vdash \alpha_\beta\equiv\alpha_\gamma$.

Proof. Follows immediately from Equivalence Theorem.

3.2.6. Relettering

Let $\alpha(x)$ be a formula containing x free, let $\alpha(y)$ be the formula obtained from $\alpha(x)$ by substituting y for x (in all occurrences of x in $\alpha(x)$).

Definition. *The formulas* $\alpha(x)$ *and* $\alpha(y)$ *are called* ***similar*** *if and only if* y *is free for* x *in* $\alpha(x)$ *and* $\alpha(x)$ *does not contain any free occurrences of* y; *and conversely*.

If $\alpha(x)$ and $\alpha(y)$ are similar then we say that one can be obtained from the other by *free relettering*.

Lemma (free relettering). *If* $\alpha(x)$ *and* $\alpha(y)$ *are similar, then*

$\vdash \alpha(x)$ *iff* $\vdash \alpha(y)$.[34]

Proof. (by Gen, Ax4 and MP).

Remark 1. The addition of "and conversely" means that x is free for y in $\alpha(y)$ and $\alpha(y)$ does not contain any free occurrences of x. Otherwise the lemma does not hold, since if, for example, $\alpha(x)$: $P(x)\lor\neg P(y)$ (a not-valid formula), then $\alpha(y)$ will be $P(y)\lor\neg P(y)$ (a valid one). If, again, in $\alpha(y)$ we substitute x for y we get a *new* formula, different from $\alpha(x)$. Intuitively, this means that $\alpha(x)$ and $\alpha(y)$ are similar if $\alpha(x)$ says of x *exactly* what $\alpha(y)$ says of y, and this means that the places where x is free in $\alpha(x)$ are *exactly* the places of y are free in $\alpha(y)$.

Lemma* (relettering of a bound variable). *If* $\alpha(x)$ *and* $\alpha(y)$ *are similar, then the following holds:*

(1) $\vdash \forall x\alpha(x)\equiv\forall y\alpha(y)$

(2) $\vdash \exists x\alpha(x)\equiv \exists y\alpha(y)$.

Proof. (Sect. 3.3, Th. 15; Sect. 3.4.1, exerc. 2).

Remark 2. According to Lemma*, we say that the formulas $\forall x\alpha(x)$ ($\exists x\alpha(x)$) and $\forall y\alpha(y)$ ($\exists y\alpha(y)$) are obtained from each other by relettering of a bound variable.

If y is not free for x in $\alpha(x)$, then from $\alpha(x)$ we can always obtain a formula $\alpha^*(x)$ in which y is free for x. We do that by simultaneous replacing all subformulas of the form $\forall y\beta(y)$ ($\exists y\beta(y)$), in which x is free, with formulas $\forall z\beta(z)$ ($\exists z\beta(z)$), where the variables z are different *each other*, different from y and do not occur in $\alpha(x)$.

---

[34] And by soundness of FOL^{ax} its *semantic* counterpart also holds: $\models \alpha(x)$ *iff* $\models \alpha(y)$ (show that!).

Example. Let $\alpha(x)$: $\forall y(P(x,y)\supset\exists y(\neg Q(x)\supset R(y)))$. As can be seen, y is not free for x in $\alpha(x)$. So we reletterate the variable y in the consequent of the implication in z and obtain $\forall y(P(x,y)\supset\exists z(\neg Q(x)\supset R(z))$, and then reletterate y in the formula so obtained in w, obtaining, finally, $\alpha^*(x)$: $\forall w(P(x,w))\supset\exists z(\neg Q(x)\supset R(z))$, in which y is free for x.
Remark 3. If $\delta(x)$ and $\delta(y)$ are similar, then *via* Lemma*, for β: $\forall x\delta(x))$ (or $\exists x\delta(x))$ and γ: $\forall y\delta(y)$ (or $\exists y\delta(y))$, by $Repl_{FOL}$ it follows that $\vdash\alpha_\beta\equiv\alpha_\gamma$ (where α_β and α_γ are as in the equivalence theorem (give the details!).

3.3. Proofs in FOL^{ax}

Th. 1. (a) $\vdash\exists x\alpha\equiv\neg\forall x\neg\alpha$

(b) $\vdash\forall x\alpha\equiv\neg\exists x\neg\alpha$

(c) $\vdash\neg\forall x\alpha\equiv\exists x\neg\alpha$

(d) $\vdash\neg\exists x\alpha\equiv\forall x\neg\alpha$.[35]

Proof (exercise; hint: use Def. $\exists$, $Rule_p$ ($\vdash\alpha\equiv\neg\neg\alpha$) and $Repl_{FOL}$).

These equivalences show us how to interchange (equivalently) the symbols $\forall$ and $\exists$ in a formula of L_{FOL}: by inserting one negation sign before and one after quantifiers and by replacing the symbols $\forall$ and $\exists$ each other. And this holds also for a compact row of more quantifiers.
Example. $\forall x\exists y\forall z\alpha\equiv\neg\exists x\neg\exists y\forall z\alpha\equiv\neg\exists x\forall y\neg\forall z\alpha\equiv\neg\exists x\forall y\exists z\neg\alpha$.
Th. 2. $\vdash\forall x\forall y\alpha\equiv\forall y\forall x\alpha$

(1) $\forall x\forall y\alpha$; hyp

(2) $\forall x\forall y\alpha\supset\forall y\alpha$; (1) Ax4

(3) $\forall y\alpha$; (1) (2) MP

(4) $\forall y\alpha\supset\alpha$; Ax4

(5) α; (3) (4) MP

(6) $\forall x\alpha$; (5) Gen

(7) $\forall y\forall x\alpha$; (6) Gen

Hence (8) $\forall x\forall y\alpha\vdash\forall y\forall x\alpha$; (1)-(7), and then

(9) $\vdash\forall x\forall y\alpha\supset\forall y\forall x\alpha$; (8) Ded. Th.

[35] The equivalences (a)-(d) of this theorem are the *syntactical* counterparts to the *semantical* ones; comp. Sect. 2.2, Example 3.

The application of Deduction Theorem is perfectly licit, since no application of Gen bounds some free variable of the assumption $\forall x\forall y\alpha$. And then the result follows by 3.2.3 Corollary b).[36]

The proof of the converse, similar.

Th. 3. $\vdash\exists x\exists y\alpha\equiv\exists y\exists x\alpha$ (exercise).

(Hint: Use Th. 2 with $\neg\alpha$ instead of α and PL).

Th. 4. $\vdash\forall x_1\ldots\forall x_n\alpha\supset\alpha$ (use Ax4) (exercise).

Th. 5. $\vdash\alpha(t/x)\supset\exists x\alpha(x)$; t is free for x in $\alpha(x)$

(1) $\forall x\neg\alpha(x)\supset\neg\alpha(t/x)$; Ax4

(2) $\alpha(t/x)\supset\neg\forall x\neg\alpha(x)$; (1), PL

(3) $\alpha(t/x)\supset\exists x\alpha(x)$; (2) Th. 1(a).

Therefore, $\alpha(t/x)\vdash\exists x\alpha(x)$. To this form of deduction we refer in what follows by Gen $\exists$ (existential generalization).

Th. 6. $\vdash\forall x(\alpha\supset\beta)\supset(\forall x\alpha\supset\forall x\beta)$

(1) $\forall x(\alpha\supset\beta)$; hyp

(2) $\forall x\alpha$; hyp

(3) $\alpha\supset\beta$; (1), Ax4, MP

(4) α; (2) Ax4, MP

(5) β; (3), (4), MP

(6) $\forall x\beta$; 5, Gen

Hence (7) $\forall x(\alpha\supset\beta), \forall x\alpha\vdash\forall x\beta$, and then

(8) $\vdash\forall x(\alpha\supset\beta)\supset(\forall x\alpha\supset\forall x\beta)$; (7), Ded. Th. (twice)

Th. 7. $\vdash\forall x(\alpha\equiv\beta)\supset(\forall x\alpha\equiv\forall x\beta)$

(1) $\forall x(\alpha\equiv\beta)$; hyp

(2) $\forall x\alpha$; hyp

(3) $\alpha\equiv\beta$; (1), Ax4, MP

(4) α; (2), Ax4, MP

(5) $\alpha\supset\beta$; (3), PL

(6) β; (4), (5), MP

(7) $\forall x\beta$; (6), Gen

Hence (8) $\forall x(\alpha\equiv\beta), \forall x\alpha\vdash\forall x\beta$, and then

(9) $\forall x(\alpha\equiv\beta)\vdash\forall x\alpha\supset\forall x\beta$; (8) Ded. Th.

(10) $\forall x(\alpha\equiv\beta)\vdash\forall x\beta\supset\forall x\alpha$; obtained in a similar way, using the hypotheses $\forall x(\alpha\equiv\beta)$ and $\forall x\beta$

[36] Where the steps of the proof are intuitively clear, they will not be explicitly mentioned anymore.

(11) $\forall x(\alpha\equiv\beta)\vdash\forall x\alpha\equiv\forall x\beta$; (9), (10), PL, Rule_p

(12) $\vdash\forall x(\alpha\equiv\beta)\supset(\forall x\alpha\equiv\forall x\beta)$; (11) Ded. Th.

Th. 8. $\vdash\forall x(\alpha\supset\beta)\supset(\exists x\alpha\supset\exists x\beta)$

(1) $\forall x(\alpha\supset\beta)$; hyp

(2) $\forall x(\neg\beta\supset\neg\alpha)$; (1), Repl_{FOL}

(3) $\forall x\neg\beta\supset\forall x\neg\alpha$; (2), Th. 6, PL

(4) $\neg\forall x\neg\alpha\supset\neg\forall x\neg\beta$; (3), PL

(5) $\exists x\alpha\supset\exists x\beta$; (4), Th. 1(a)

Hence (6) $\forall x(\alpha\supset\beta)\vdash\exists x\alpha\supset\exists x\beta$, and then

(7) $\vdash\forall x(\alpha\supset\beta)\supset(\exists x\alpha\supset\exists x\beta)$; (6) Ded. Th.

Th. 9. $\vdash\forall x(\alpha\equiv\beta)\supset(\exists x\alpha\equiv\exists x\beta)$

(1) $\forall x(\alpha\equiv\beta)$; hyp

(2) $\alpha\equiv\beta$; (1) Ax4, MP

(3) $(\alpha\equiv\beta)\supset(\alpha\supset\beta)$; PL

(4) $\alpha\supset\beta$; (2), (3), MP

(5) $\forall x(\alpha\supset\beta)$; (4) Gen

(6) $\forall x(\alpha\supset\beta)\supset(\exists x\alpha\supset\exists x\beta)$; Th. 8

(7) $\exists x\alpha\supset\exists x\beta$; (5), (6), MP

(8) $\exists x\beta\supset\exists x\alpha$; as by steps (2)-(7)

(9) $(\exists x\alpha\supset\exists x\beta)\supset[(\exists x\beta\supset\exists x\alpha)\supset(\exists x\alpha\equiv\exists x\beta)]$; Rule_p

(10) $\exists x\alpha\equiv\exists x\beta$; (7), (8), (9), MP

Hence (11) $\forall x(\alpha\equiv\beta)\vdash\exists x\alpha\equiv\exists x\beta$, and then

(12) $\vdash\forall x(\alpha\equiv\beta)\supset(\exists x\alpha\equiv\exists x\beta)$; (11) Ded. Th.

Th. 10. $\alpha(x)\supset\beta(x)\vdash_x\forall x\alpha(x)\supset\forall x\beta(x)$

(1) $\alpha(x)\supset\beta(x)$; hyp

(2) $\forall x(\alpha(x)\supset\beta(x))$; (1) Gen

(3) $\forall x\alpha(x)\supset\forall x\beta(x)$; (2), Th. 6, MP

Hence (4) $\alpha(x)\supset\beta(x)\vdash_x\forall x\alpha(x)\supset\forall x\beta(x)$

It should be observed that this deduction contains one application of Gen to the assumption $\alpha(x)\supset\beta(x)$, by which the variable x of this assumption becomes bound. Hence the symbol "$\vdash$" cannot be removed on the left, x being fixed in the above derivation, a fact expressed by the notation $\vdash_x$.

A variant of proof of Th. 10 can be given in the following way:

(1) $\alpha(x)\supset\beta(x)$; hyp

(2) $\forall x\alpha(x)$; hyp

(3) $\alpha(x)$; (2), Ax4, MP

(4) $\beta(x)$; (1), (3), MP

(5) $\forall x\beta(x)$; (4), Gen

Hence (6) $\alpha(x)\supset\beta(x), \forall x\alpha(x)\vdash_x\forall x\beta(x)$. Of course, x is fixed by one application of Gen to the formula $\beta(x)$ depending on the assumption $\alpha(x)\supset\beta(x)$. But, by Ded. Th. (Corol. a)), we can move the symbol $\vdash_x$ to the left *only once*, and obtain $\alpha(x)\supset\beta(x)\vdash_x\forall x\alpha(x)\supset x\beta(x)$.

Th. 11. $\alpha(x)\supset\beta\vdash_x\exists x\alpha(x)\supset\beta$; x is not free in β

(1) $\alpha(x)\supset\beta$; hyp

(2) $\neg\beta\supset\neg\alpha(x)$; (1), PL

(3) $\forall x(\neg\beta\supset\neg\alpha(x))$; (2) Gen

(4) $\neg\beta\supset\forall x\neg\alpha(x)$; (3), Ax5, MP

(5) $\neg\forall x\neg\alpha(x)\supset\beta$; (4), PL

(6) $\exists x\alpha(x)\supset\beta$; (5), Th. 1(a), $\mathrm{Repl}_{\mathrm{FOL}}$.

Hence (7) $\alpha(x)\supset\beta\vdash_x\exists x\alpha(x)\supset\beta$; (1)-(6).

Th. 12. $\beta\supset\alpha(x)\vdash_x\beta\supset\forall x\alpha(x)$; x is not free in β

(1) $\beta\supset\alpha(x)$; hyp

(2) $\forall x(\beta\supset\alpha(x))$; (1) Gen

(3) $\beta\supset\forall x\alpha(x)$; (2) Ax5 (since x is not free in β), MP

Hence (4) $\beta\supset\alpha(x)\vdash_x\beta\supset\forall x\alpha(x)$; (1)-(3)

Th. 13. $\vdash\forall x(\alpha\wedge\beta)\equiv(\forall x\alpha\wedge\forall x\beta)$

(1) $\forall x(\alpha\wedge\beta)$; hyp

(2) $\alpha\wedge\beta$; (1) Ax4, MP

(3) $(\alpha\wedge\beta)\supset\alpha$; $\mathrm{Rule}_{\mathrm{p}}$

(4) α; (2), (3), MP

(5) $\forall x\alpha$; (4) Gen

(6) $(\alpha\wedge\beta)\supset\beta$; $\mathrm{Rule}_{\mathrm{p}}$

(7) β; (2), (6), MP

(8) $\forall x\beta$; (7) Gen

(9) $\forall x\alpha\wedge\forall x\beta$; (5), (8), $\mathrm{Rule}_{\mathrm{p}}$, MP

Hence (10) $\forall x(\alpha\wedge\beta)\vdash\forall x\alpha\wedge\forall x\beta$, and then

(11) $\vdash\forall x(\alpha\wedge\beta)\supset(\forall x\alpha\wedge\forall x\beta)$; (1) Ded. Th.

(12) $\forall x\alpha\wedge\forall x\beta$; hyp

(13) $\forall x\alpha$; (12) $\mathrm{Rule}_{\mathrm{p}}$ MP

(14) $\forall x\alpha \supset \alpha$; Ax4
(15) α; (13), (14), MP
(16) $\forall x\beta$; (12) Rule_p MP
(17) $\forall x\beta \supset \beta$; Ax4.
(18) β; (16), (17), MP
(19) $\alpha \wedge \beta$; (15), (18), Rule_p
(20) $\forall x(\alpha \wedge \beta)$; (19) Gen

Hence (21) $\forall x \wedge \forall x\beta \vdash \forall x(\alpha \wedge \beta)$; (12)-(20), and then
(22) $\vdash (\forall x\alpha \wedge \forall x\beta) \supset \forall x(\alpha \wedge \beta)$; Ded. Th.
(23) $\forall x(\alpha \wedge \beta) \equiv (\forall x\alpha \wedge \forall x\beta)$; (11), (22), Rule_p, MP

Th. 14. $\vdash \exists x \forall y\alpha \supset \forall y \exists x\alpha$
(1) $\forall y\alpha$; hyp
(2) $\forall y\alpha \supset \alpha$; Ax4
(3) α; (1), (2), MP
(4) $\exists x\alpha$; (3), Th. 5, MP
(5) $\forall y \exists x\alpha$; (4) Gen

Hence (6) $\forall y\alpha \vdash \forall y \exists x\alpha$, and then
(7) $\vdash \forall y\alpha \supset \forall y \exists x\alpha$; (6) Ded. Th.
(8) $\vdash \exists x \forall y\alpha \supset \forall y \exists x\alpha$; (7), Th. 11, since x is not free in the consequent.

The converse of this theorem does not hold.

Th. 15. $\vdash \forall x\alpha(x) \equiv \forall y\alpha(y)$; where $\alpha(x)$ and $\alpha(y)$ are similar.[37]
(1) $\forall x\alpha(x) \supset \alpha(y)$; Ax4
(2) $\forall y(\forall x\alpha(x) \supset \alpha(y))$; (1) Gen
(3) $\forall x\alpha(x) \supset \forall y\alpha(y)$; (2) Ax5, MP
(4) $\forall y\alpha(y) \supset \alpha(x)$; Ax4
(5) $\forall x(\forall y\alpha(y) \supset \alpha(x))$; (4) Gen
(6) $\forall y\alpha(y) \supset \forall x\alpha(x)$; (5) Ax5, MP
(7) $\forall x\alpha(x) \equiv \forall y\alpha(y)$; (3), (6), Rule_p

Th. 16. $\vdash (\beta \supset \forall x\alpha(x)) \equiv \forall y(\beta \supset \alpha(y))$; y is not free in β and $\alpha(x)$ and $\alpha(y)$ are similar.
(1) $\forall x\alpha(x) \equiv \forall y\alpha(y)$; Th. 15
(2) $(\beta \supset \forall x\alpha(x)) \equiv (\beta \supset \forall y\alpha(y))$; Rule_p, MP

[37] Comp. 3.2.6 (above).

(3) $\forall y(\beta\supset\alpha(y))\supset(\beta\supset\forall y\alpha(y))$; Ax5, since y is not free in β
(4) $\beta\supset\forall y\alpha(y)$; hyp
(5) $\forall y\alpha(y)\supset\alpha(y)$; Ax4
(6) $\beta\supset\alpha(y)$; (3), (4), Rule_p
(7) $\forall y(\beta\supset\alpha(y))$; (6) Gen

Hence (8) $\beta\supset\forall y\alpha(y)\vdash\forall y(\beta\supset\alpha(y))$; (4)-(7), and then
(9) $\vdash(\beta\supset\forall y\alpha(y))\supset\forall y(\beta\supset\alpha(y))$; (8) Ded. Th.
(10) $(\beta\supset\forall y\alpha(y))\equiv\forall y(\beta\supset\alpha(y))$; (3), (9), Rule_p
(11) $(\beta\supset\forall x\alpha(x))\equiv\forall y(\beta\supset\alpha(y))$; (2), (10) Rule_p

Th. 17. $\vdash(\beta\supset\exists x\alpha(x))\equiv\exists y(\beta\supset\alpha(y))$; y is not free in β and $\alpha(x)$ and $\alpha(y)$ are similar.
(1) $\forall x\neg\alpha(x)\equiv\forall y\neg\alpha(y)$; Th. 15
(2) $\neg\forall x\neg\alpha(x)\equiv\neg\forall y\neg\alpha(y)$; (1) Rule_p MP
(3) $\exists x\alpha(x)\equiv\exists y\alpha(y)$; (2), Th. 1(a)
(4) $(\beta\supset\exists x\alpha(x))\equiv(\beta\supset\exists y\alpha(y))$; (3) Rule_p, MP
(5) $\neg(\beta\supset\exists y\alpha(y))\equiv(\beta\wedge\neg\exists y\alpha(y))$; Rule_p
(6) $(\beta\wedge\neg\exists y\alpha(y))\equiv(\beta\wedge\forall y\neg\alpha(y))$; Th. 1(d), Repl_{FOL}
(7) $(\beta\wedge\forall y\neg\alpha(y))\equiv\forall y(\beta\wedge\neg\alpha(y))$; (6) by Exerc. 2a) below
(since y is not free in β $\forall y\beta$ is just β)
(8) $\forall y(\beta\wedge\neg\alpha(y))\equiv\forall y\neg(\beta\supset\alpha(y))$; (7), Rule_p, Repl_{FOL}
(9) $\neg(\beta\supset\exists y\alpha(y))\equiv\forall y\neg(\beta\supset\alpha(y))$; (5)-(8) Rule_p
(10) $(\beta\supset\exists y\alpha(y))\equiv\neg\forall y\neg(\beta\supset\alpha(y))$; (9) Rule_p
(11) $(\beta\supset\exists y\alpha(y))\equiv\exists y(\beta\supset\alpha(y))$; (10), Th. 1(a)
(12) $(\beta\supset\exists x\alpha(x))\equiv\exists y(\beta\supset\alpha(y))$; (4), (11) Rule_p

Th. 18. $\vdash(\forall x\alpha(x)\supset\beta)\equiv\exists y(\alpha(y)\supset\beta)$; y is not free in β and $\alpha(x)$ and $\alpha(y)$ are similar.
(1) $\forall x\alpha(x)\equiv\forall y\alpha(y)$; Th. 15
(2) $(\forall x\alpha(x)\equiv\forall y\alpha(y))\supset[(\forall x\alpha(x)\supset\beta)\equiv(\forall y\alpha(y)\supset\beta)]$; Rule_p
(3) $(\forall x\alpha(x)\supset\beta)\equiv(\forall y\alpha(y)\supset\beta)$; (1), (2), MP
(4) $(\forall y\alpha(y)\supset\beta)\equiv\neg(\forall y\alpha(y)\wedge\neg\beta)$; Rule_p
(5) $\neg(\forall y\alpha(y)\wedge\neg\beta)\equiv\neg\forall y(\alpha(y)\wedge\neg\beta)$; by Exerc. 2a) (below)
(since y is not free in β), Repl_{FOL}.
(6) $(\forall y\alpha(y)\supset\beta)\equiv\neg\forall y\neg(\alpha(y)\supset\beta)$; (4)-(5) Rule_p, Repl_{FOL}.
(7) $(\forall y\alpha(y)\supset\beta)\equiv\exists y(\alpha(y)\supset\beta)$; (6) Th. 1(a)
(8) $(\forall x\alpha(x)\supset\beta)\equiv\exists y(\alpha(y)\supset\beta)$; (3), (7), Rule_p

Th. 19. $\vdash(\exists x\alpha(x)\supset\beta)\equiv\forall y(\alpha(y)\supset\beta)$; y is not free in β and $\alpha(x)$ and $\alpha(y)$ are similar.

(1) $\forall x\neg\alpha(x)\equiv\forall y\neg\alpha(y)$; Th. 15

(2) $\neg\forall x\neg\alpha(x)\equiv\neg\forall y\neg\alpha(y)$; (1) $Rule_p$, MP

(3) $\exists x\alpha(x)\equiv\exists y\alpha(y)$; (2) Th. 1(a)

(4) $(\exists x\alpha(x)\supset\beta)\equiv(\exists y\alpha(y)\supset\beta)$; (3) $Rule_p$, MP

(5) $\forall y(\alpha(y)\supset\beta)\equiv\forall y(\neg\beta\supset\neg\alpha(y))$; $Rule_p$, $Repl_{FOL}$.

(6) $\forall y(\neg\beta\supset\neg\alpha(y))\supset(\neg\beta\supset\forall y\neg\alpha(y))$; Ax5, since y is not free in β

(7) $(\neg\beta\supset\forall y\neg\alpha(y))\supset\neg\forall y\neg\alpha(y)\supset\beta$; $Rule_p$

(8) $(\neg\forall y\neg\alpha(y)\supset\beta)\supset(\exists y\alpha(y)\supset\beta)$; Th. 1(a)

(9) $\forall y(\alpha(y)\supset\beta)\supset(\exists y\alpha(y)\supset\beta)$; (5)-(8) $Rule_p$, $Repl_{FOL}$

(10) $\exists y\alpha(y)\supset\beta$; hyp

(11) $\neg\beta\supset\neg\exists y\alpha(y)$; (10) $Rule_p$, MP

(12) $\neg\beta\supset\forall y\neg\alpha(y)$; (11), Th. 1(d), $Repl_{FOL}$.

(13) $\forall y(\neg\beta\supset\neg\alpha(y))$; (12), Th. 16, MP

(14) $\forall y(\alpha(y)\supset\beta)$; (13) $Rule_p$, $Repl_{FOL}$.

Hence (15) $\exists y\alpha(y)\supset\beta\vdash\forall y(\alpha(y)\supset\beta)$; (10)-(14), and then

(16) $\vdash(\exists y\alpha(y)\supset\beta)\supset\forall y(\alpha(y)\supset\beta)$; (15) Ded. Th. (since y is not free in hyp)

(17) $(\exists y\alpha(y)\supset\beta)\equiv(\forall y(\alpha(y)\supset\beta)))$; (9), (16) $Rule_p$, MP

(18) $(\exists x\alpha(x)\supset\beta)\equiv\forall y(\alpha(y)\supset\beta)$; (4), (17), $Rule_p$, MP

Remark. The theorem Th. 1 c) and d) and the theorems Th. 16 – Th. 19 allow us to move the interior quantifiers of a formula in the front of it, playing by this an essential role in constructing normal forms in FOL, as we shall see in the next section.

Exercises

1. Show the following things:

a) $\vdash(\alpha(x)\supset\beta)\supset(\forall x\alpha(x)\supset\beta)$

b) $\vdash(\beta\supset\alpha(x))\supset(\beta\supset\exists x\alpha(x))$

c) $\vdash\exists x(\alpha\vee\beta)\equiv(\exists x\alpha\vee\exists x\beta)$.

2. Prove that if x is not free in β, then the following hold:

a) $\vdash\forall x(\alpha(x)\wedge\beta)\equiv(\forall x\alpha(x)\wedge\beta)$

b) $\vdash\forall x(\alpha(x)\vee\beta)\equiv(\forall x\alpha(x)\vee\beta)$

c) $\vdash\exists x(\alpha(x)\wedge\beta)\equiv(\exists x\alpha(x)\wedge\beta)$

d) $\vdash\exists x(\alpha(x)\vee\beta)\equiv(\exists x\alpha(x)\vee\beta)$.

Choice Rule

The *choice rule* (C-Rule,[38] in what follows) is often used in mathematical reasoning. It is based on the following fact: If we proved a formula of the form $\exists x\alpha(x)$, then we say let c be an object having the property α and write $\alpha(c)$. We unroll the whole proof and arrive, finally, at a formula not containing c.

Example. Let us prove that the following deduction holds:

$$\forall x(\alpha(x)\supset\neg\beta(x)),\exists x\alpha(x)\vdash\exists x\neg(\alpha(x)\supset\beta(x)).$$

Using C-Rule we reason as follows:

(1) $\forall x(\alpha(x)\supset\neg\beta(x))$; hyp.

(2) $\exists x\alpha(x)$; hyp.

(3) $\alpha(c)\supset\neg\beta(c)$; (1), Ax4, MP

(4) $\alpha(c)$; (2) for some c; C-Rule

(5) $(\alpha(c)\supset\neg\beta(c))\wedge\alpha(c)$; (3), (4), PL

(6) $((\alpha(c)\supset\neg\beta(c))\wedge\alpha(c))\supset\neg\beta(c)$; Rule_p (Sect. 3.2.1)

(7) $\neg\beta(c)$; (5), (6), MP

(8) $\alpha(c)\supset(\neg\beta(c)\supset\neg(\alpha(c)\supset\beta(c)))$; Rule_p

(9) $\neg(\alpha(c)\supset\beta(c))$; (4), (7), (8), MP (twice)

(10) $\exists x\neg(\alpha(x)\supset\beta(x))$; (9), Gen $\exists$

Therefore, $\forall x(\alpha(x)\supset\neg\beta(x)),\exists x\alpha(x)\vdash\exists x\neg(\alpha(x)\supset\beta(x))$.

We note that any formula provable by using C-Rule can also be proved with no use of this rule. For our example such a proof runs as follows:

(1) $\forall x(\alpha(x)\supset\neg\beta(x))$; hyp.

(2) $\neg\exists x\neg(\alpha(x)\supset\beta(x))$; hyp.

(3) $\alpha(x)\supset\neg\beta(x)$; (1), Ax4, MP

(4) $\alpha(x)\supset\beta(x)$; (2), Ax4, MP

(5) $(\alpha(x)\supset\neg\beta(x))\supset((\alpha(x)\supset\beta(x))\supset\neg\alpha(x))$; Rule_p (Sect. 3.2.1)

(6) $\neg\alpha(x)$; (3), (4), (5), MP (twice)

(7) $\forall x\neg\alpha(x)$; (6), Gen.

Therefore,

(8) $\forall x(\alpha(x)\supset\neg\beta(x)),\neg\exists x\neg(\alpha(x)\supset\beta(x))\vdash\forall x\neg\alpha(x)$; (1)-(7)

(9) $\forall x(\alpha(x)\supset\neg\beta(x)\vdash\neg\exists x\neg(\alpha(x))\supset\beta(x))\supset\forall x\neg\alpha(x)$; (8), Ded. Th.

[38] The name *Rule C* was given by B. Rosser [1953], 128, where this rule is formulated for the first time.

(10) $\forall x(\alpha(x)\supset\neg\beta(x))\vdash\neg\forall x\neg\alpha(x)\supset\exists x\neg(\alpha(x)\supset\beta(x))$; (9), *via* PL

(11) $\forall x(\alpha(x)\supset\neg\beta(x)),\exists x\alpha(x)\vdash\exists x\neg(\alpha(x)\supset\beta(x))$; (10), Th. 1(a)

3.4. Normal forms

3.4.1. Prenex normal form

Definition. *A formula* α *of* L_{FOL} *is in prenex normal form if* α *has the form* $Q_1x_1...Q_nx_n\beta$ *where* Q_ix_i, $1\leq i\leq n$, *are universal or existential quantifiers and* β *is a formula of* L_{FOL} *containing no quantifiers*.[39]

$Q_1x_1...Q_nx_n$ is called the *prefix* and β the *matrix* of α.

Theorem 1. *To every formula* α *of* L_{FOL} *there is a formula* α^* *in prenex normal form such that* $\vdash\alpha\equiv\alpha^*$.

Proof. (Algorithm of constructing α^*, described by induction on the complexity of α)

Basis. $n = 0$. In this case $\alpha^* = \alpha$.

Induction. Assume $n>0$ and for every $k < n$ the theorem holds. We have to show that the theorem also holds for $k = n$.

1. $\alpha = \neg\gamma$. By inductive hypothesis, a formula γ^* , the prenex normal form of γ, can be constructed such that $\vdash\gamma\equiv\gamma^*$. It follows, *via* $Rule_p$, that $\vdash\neg\gamma\equiv\neg\gamma^*$, and hence $\vdash\alpha\equiv\neg\gamma^*$. By applying Th. 1 (c) and (d) (Sect. 3.3) and $Repl_{FOL}$, a formula α^* in prenex normal form can be constructed such that $\vdash\neg\gamma^*\equiv\alpha^*$, and then $\vdash\alpha\equiv\alpha^*$.

2. $\alpha = \gamma\supset\delta$. By inductive hypothesis, the prenex normal forms of γ and δ can be constructed, let these be γ^* and δ^*, respectively, such that $\vdash\gamma\equiv\gamma^*$ and $\vdash\delta\equiv\delta^*$. But $\vdash(\gamma\equiv\gamma^*)\supset[(\delta\equiv\delta^*)\supset((\gamma\supset\delta)\equiv(\gamma^*\supset\delta^*))]$ (by $Rule_p$). Whence, by MP, it follows that $\vdash(\gamma\supset\delta)\equiv(\gamma^*\supset\delta^*)$. But $\gamma\supset\delta$ is just α, hence $\vdash\alpha\equiv\gamma^*\supset\delta^*$. Now, by Th. 16 - Th. 19 of Sect. 3.3 the quantifiers in the prefixes of γ^* and δ^* are gradually removed to the left of the whole formula and obtain a formula α^* in prenex normal form such that $\vdash\alpha\equiv\alpha^*$.

3. $\alpha = \forall x\gamma$. Again, by inductive hypothesis, the prenex normal form of γ can be constructed, let it be γ^*, such that $\vdash\gamma\equiv\gamma^*$. But in this case we also have $\vdash\forall x\gamma\equiv\forall x\gamma^*$, i.e., $\vdash\alpha\equiv\forall x\gamma^*$. And since γ^* is in the prenex normal form

[39] If $n = 0$, then α is already in prenex normal form.

it follows that $\alpha^* = \forall x\gamma^*$ is also in the prenex normal form; hence $\vdash \alpha \equiv \alpha^*$.

The preceding theorem has a stronger form, concerning just the form of the prefix of a prenex normal form of α. We assume that a formula α of L_{FOL}, in its prenex normal form, is closed. And this is a licit thing, since the following holds: α is provable/valid iff its closure is provable/valid.

Theorem 2. *Let* α *be a formula of* L_{FOL} *in prenex normal form. Then a formula* β *can be constructed whose prefix begins with* ∃x∀y *such that*:

$$\vdash \alpha \text{ iff } \vdash \exists x \forall y \pi\beta(x,y).$$

Proof. Suppose that α has the form πγ with arbitrary prefix π. Then

(1) $\vdash \pi\gamma \equiv \pi\gamma \wedge \exists x \forall y (P(x,y) \supset P(x,y))$, with x and y not occurring in γ (since $\vdash \exists x \forall y (P(x,y) \supset P(x,y))$)

(2) $\vdash [\pi\gamma \wedge \exists x \forall y (P(x,y) \supset P(x,y))] \equiv \exists x \forall y \pi(\gamma \wedge (P(x,y) \supset P(x,y)))$; by 3.3. Exercises 2 a) and c).

Hence (3) $\vdash \alpha \equiv \exists x \forall y \pi\beta(x,y)$. Whence $\vdash \alpha$ iff $\vdash \exists x \forall y \pi\beta(x,y)$.

Let us illustrate the construction of a prenex normal form of a formula α of L_{FOL}.

Example.[40] $\alpha = \neg\forall x \exists z (P(x) \supset Q(y,z)) \supset \forall y (P(y) \supset \neg\exists x Q(x,z))$

By one application of Th. 1(c) we remove "¬" from the front of "∀" and obtain:

(1) $\exists x \neg\exists z (P(x) \supset Q(y,z)) \supset \forall y (P(y) \supset \neg\exists x Q(x,z))$.

Now we repeat this operation, this time in order to remove "¬" from the front of "∃", in both places in (1), by applying Th. 1(d), and we get:

(2) $\exists x \forall z \neg(P(x) \supset Q(y,z)) \supset \forall y (P(y) \supset \forall x \neg Q(x,z))$.

Let us observe that this formula has the form $\exists x \alpha(x) \supset \beta$, so by applying Th. 19 we obtain:

(3) $\forall v [\forall z \neg(P(v) \supset Q(y,z)) \supset \forall y (P(y) \supset \forall x \neg Q(x,z))]$.

Now, again, the formula in brackets has the form $\forall x \alpha(x) \supset \beta$, and then by one application of Th. 18 we obtain:

(4) $\forall v \exists w [\neg(P(v) \supset Q(y,w)) \supset \forall y (P(y) \supset \forall x \neg Q(x,z))]$.

Now we bring ∀y into the prefix, using Th. 16, and then:

(5) $\forall v \exists w \forall u [\neg(P(v) \supset Q(y,w)) \supset (P(u) \supset \forall x \neg Q(x,z))]$.

By applying Th. 16 we bring ∀x into the front of the last implication and obtain:

---

[40] By the completeness of the set M = {¬,⊃} (Ch. 1, 2.3) we only consider the formulas containing these two connectives: ¬ and ⊃.

$(6)\ \forall v\exists w\forall u[\neg(P(v)\supset Q(y,w))\supset\forall s(P(u)\supset\neg Q(s,z))]$.

By one application of the same theorem, we bring, finally, $\forall$ from the consequent in the front of the whole implication, i.e.,

$(7)\ \forall v\exists w\forall u\forall t[\neg(P(v)\supset Q(y,w))\supset(P(u)\supset\neg Q(t,z))]$

and this is the prenex normal form of α.

Exercises

1. Construct the prenex normal form of the following formulas of L_{FOL}:

$\alpha = \exists x(\forall yP(x,y)\supset Q(x))\supset\neg\exists x(Q(x)\supset\forall yP(x,y))$

$\beta = \forall z(P(z)\supset(\forall y(Q(x,y)\supset\exists xQ(y,x)))$

$\gamma = \exists xR(x,y,z)\supset(P(x)\supset\forall yQ(x,y))$

$\delta = \exists x\neg(\forall yQ(x,y)\supset R(y,z))\supset\exists zQ(z)$

2. Show that if $\alpha(x)$ and $\alpha(y)$ are similar, then the following holds:

$\vdash\exists x\alpha(x)\equiv\exists y\alpha(y)$.

3.4.2. Skolem normal form

The construction of the prenex normal form of a formula α of L_{FOL} is the first step of a process by which an even simpler form can be constructed, the *Skolem normal form*.[41]

Definition 1. *A formula* α *of* L_{FOL} *is in Skolem normal form if* α *is in the prenex normal form and all the existential quantifiers precede all the universal quantifiers*.

Definition 2. *Two formulas of* L_{FOL}, α *and* β, *are said to be co-deductive* (*or deductively equivalent*[42]) *if and only if each is derivable from the other, and then the following holds*: $\vdash\alpha$ *if and only if* $\vdash\beta$.

It should be observed that the deductive equivalence is weaker than the relation $\vdash\alpha\equiv\beta$ of 3.4.1 Theorem 1, since if $\vdash\alpha\equiv\beta$, than α and β are derivable from each other using MP. But $\vdash\alpha\equiv\beta$ does not follow from co-deductivity, since $\alpha(x)$, with x free, and $\forall x\alpha(x)$ are co-deductive but not equivalent.

Theorem 1. *To every formula* α *of* L_{FOL} *there is a formula* α^* *in Skolem normal form such that* α *and* α^* *are co-deductive*.[43]

[41] Cf. T. Skolem [1920].

[42] Comp. D. Hilbert; P. Bernays, [1934], 149: "deduktionsgleich".

[43] This result holds for the so-called *pure predicate* calculus, i.e., for a first-order predicate

Proof. In proving of this theorem we confine the analysis to the formulas in prenex normal form and, moreover, to the closed ones. For, as we know, a formula $\alpha(x)$ is provable if and only if its closure is provable. The proof will be given by induction on the *rank* of α, i.e., by induction on the number of the universal quantifiers preceding the existential quantifiers in α. If the rank of α is 0, then α is already in Skolem normal form. So, we suppose that the theorem holds for formulas whose rank is less than *n* and show that it also holds for *n*.

Assume $\alpha = \exists x_1 \ldots \exists x_n \forall y \beta(x_1,\ldots,x_n,y)$, where β contains only $x_1,\ldots,x_n,y$ free.[44] Of course, if α is in prenex normal form, then β is also in prenex normal form, hence β may be a formula of the form $Q_1 z_1 \ldots Q_s z_s \gamma(x_1,\ldots,x_n,y,z_1,\ldots,z_s)$, where Q_i, $1 \leq i \leq s$, are the quantifiers $\forall$ or $\exists$ binding the variables $z_1,\ldots,z_s$ (where at least one of Q_i must be $\exists$, otherwise α would already be in Skolem normal form). Hence, written more fully, $\alpha = \exists x_1 \ldots \exists x_n \forall y Q_1 z_1 \ldots Q_s z_s \gamma(x_1,\ldots,x_n,y,z_1,\ldots,z_s)$. If $n = 0$, then $\alpha = \forall y \beta(y)$.

Now, let P^{n+1} be an n+1-place predicate symbol not occurring in α. Using it we construct the formula:

$$\delta = \exists x_1 \ldots \exists x_n [\forall y(\beta(x_1,\ldots,x_n,y) \supset P(x_1,\ldots,x_n,y)) \supset \forall y P(x_1,\ldots,x_n,y)].$$

Written more fully $\delta = \exists x_1 \ldots \exists x_n [\forall y(Q_1 z_1 \ldots Q_s z_s \gamma(x_1,\ldots,x_n,y,z_1,\ldots,z_s) \supset P(x_1,\ldots,x_n,y)) \supset \forall y P(x_1,\ldots,x_n,y)]$.

We show that α and δ are co-deductive, i.e., the following holds: $\vdash \alpha$ iff $\vdash \delta$.[45]

Assume $\vdash \delta$. By substitution of β for P in δ we obtain:

$$\delta_1 = \exists x_1 \ldots \exists x_n [\forall y(\beta(x_1,\ldots,x_n,y) \supset \beta(x_1,\ldots,x_n,y)) \supset \forall y \beta(x_1,\ldots,x_n,y)].$$

But in FOL^{ax}: $\vdash [\forall y(Q(y) \supset Q(y)) \supset \forall y Q(y)] \equiv \forall y Q(y)$. Then by substitution (for predicate symbols) and $Repl_{FOL}$ we obtain $\exists x_1 \ldots \exists x_n \forall y \beta(x_1,\ldots,x_n,y)$, i.e., the formula α.

Conversely, assume $\vdash \alpha$, i.e., $\vdash \exists x_1 \ldots \exists x_n \forall y \beta(x_1,\ldots,x_n,y)$. By $Rule_p$, from 3.3 Th. 6, for $\alpha = P(y)$ and $\beta = Q(y)$, we obtain its equivalent:

$$\forall y P(y) \supset [\forall y(P(y) \supset Q(y)) \supset \forall y Q(y)].$$

In this formula we make the following substitutions: $\beta(x_1,\ldots,x_n,y)/P(y)$ and $P(x_1,\ldots,x_n,y)/Q(y)$, and obtain:

$$\forall y \beta(x_1,\ldots,x_n,y) \supset [\forall y(\beta(x_1,\ldots,x_n,y) \supset P(x_1,\ldots,x_n,y)) \supset \forall y P(x_1,\ldots,x_n,y)].$$

---

calculus whose language does not contain functional symbols or constant symbols.

[44] Based on 3.4.1 Theorem 2 we may assume that $n > 0$, but this is not necessary.

[45] This proof is a variant of the proof in D. Hilbert and W. Ackermann, [1972], Ch. 3, §7.

From this formula, by one application of Gen with respect to the variable x_n we get:

$\forall x_n\{\forall y\beta(x_1,...,x_n,y)\supset[\forall y(\beta(x_1,...,x_n,y)\supset P(x_1,...,x_n,y))\supset\forall yP(x_1,...,x_n,y)]\}$.

But if we consider now the following form of a theorem of FOL^{ax} (cf. 3.3, Th. 8): $\forall x_n(P(x_n)\supset Q(x_n))\supset(\exists x_n P(x_n)\supset\exists x_n Q(x_n))$, in which we make the required substitutions and then apply MP, we obtain:

$\exists x_n\forall y\beta(x_1,...,x_n,y)\supset\exists x_n[\forall y(\beta(x_1,...,x_n,y)\supset P(x_1,...,x_n,y))\supset\forall yP(x_1,...,x_n,y)]$.

By repeating this process *n*-times we finally get:

$$\exists x_1...\exists x_n\forall y\beta(x_1,...,x_n,y)\supset\exists x_1...\exists x_n[\forall y(\beta(x_1,...,x_n,y)\supset P(x_1,...,x_n,y)) \supset\forall yP(x_1,...,x_n,y)].$$

The antecedent of this implication is just the formula α, assumed to be provable. Hence, by one application of MP results:

$$\delta = \exists x_1...\exists x_n[\forall y(\beta(x_1,...,x_n,y)\supset P(x_1,...,x_n,y))\supset\forall yP(x_1,...,x_n,y)],$$

and therefore $\vdash\delta$.

More fully written δ is the formula

$\exists x_1...\exists x_n[\forall y(Q_1z_1...Q_sz_s\gamma(x_1,...,x_n,y,z_1,...,z_s)\supset P(x_1,...,x_n,y))\supset\forall yP(x_1,...,x_n,y)]$

In order to construct the prenex normal form of δ we proceed as follows. By one application of Th. 18 (of 3.3) to the formula in brackets and using $Repl_{FOL}$ we derive:

$\exists x_1...\exists x_n\exists y[(Q_1z_1...Q_sz_s\gamma(x_1,...,x_n,y,z_1,...,z_s)\supset P(x_1,...,x_n,y))\supset\forall yP(x_1,...,x_n,y)]$.

Now we extend the scope of the quantifiers in the prefix π of β, i.e., $\pi = Q_1z_1...Q_sz_s$, to the whole of the first implication, in the following way.

If the first quantifier in π is $\exists z_1$, then the antecedent of the first implication is $\exists z_1Q_2z_2,...,Q_sz_s\gamma(x_1,...,x_n,y,z_1,...,z_s)$. Hence by one application of 3.3, Th. 19 it follows:

$$\forall z_1(Q_2z_2,...,Q_sz_s\gamma(x_1,...,x_n,y,z_1,...,z_s)\supset P(x_1,...,x_n,y)).$$

If the first quantifier in π is $\forall z_1$, then the antecedent of the first implication is $\forall z_1Q_2z_2,...,Q_sz_s\gamma(x_1,...,x_n,y,z_1,...,z_s)$, and then by one application of 3.3, Th. 18 we derive

$$\exists z_1(Q_2z_2,...,Q_sz_s\gamma(x_1,...,x_n,y,z_1,...,z_s)\supset P(x_1,...,x_n,y)),$$

Let us observe that if we apply this operation *s*-times, then the prefix of β passes to the front of the first implication, but the quantifiers of π interchange their places. If π^{-1} is the prefix of the first implication, then by one application of $Repl_{FOL}$, δ becomes (equivalently):

$\exists x_1...\exists x_n\exists y[\pi^{-1}(\gamma(x_1,...,x_n,y,z_1,...,z_s)\supset P(x_1,...,x_n,y))\supset\forall yP(x_1,...,x_n,y)]$.

Now we repeat the preceding operations this time in order to extend the scope of π^{-1} to the whole implication. Since $z_1,...,z_s$ have no free occurrences in $\forall yP(x_1,...,x_n,y)$, we may apply again Th. 18 and Th. 19 of 3.3, by which the quantifiers in π^{-1} interchange their places, and therefore π^{-1} becomes again π. Hence the preceding formula has now the following (equivalent) form

$$\exists x_1...\exists x_n\exists y\pi[(\gamma(x_1,...,x_n,y,z_1,...,z_s)\supset P(x_1,...,x_n,y))\supset\forall yP(x_1,...,x_n,y)].$$

Now, finally, we have to bring $\forall$ into the prefix, by one application of 3.3, Th. 16, by changing y in w (since y occurs free in γ), obtaining

$$\delta_1 = \exists x_1...\exists x_n\exists y\pi\forall w[(\gamma(x_1,...,x_n,y,z_1,...,z_s)\supset P(x_1,...,x_n,y))\supset P(x_1,...,x_n,w)],$$

where γ has no quantifiers. δ_1 is the prenex normal form of δ.

By 3.4.1, Theorem 1 we have: $\vdash\delta\equiv\delta_1$. But the rank of δ_1 is one less than the rank of α. Since $\vdash\alpha$ iff $\vdash\delta$ (as we saw above) it follows that $\vdash\alpha$ iff $\vdash\delta_1$. By inductive hypothesis a Skolem normal form of δ_1 can be constructed, let call it α^*, such that $\vdash\alpha$ iff $\vdash\alpha^*$.

Example. $\alpha = \exists x\forall y\exists zQ(x,y,z)$

As can be observed, α has the form $\exists x\forall y\beta$, where $\beta = \exists zQ(x,y,z)$, containing two free variables, x and y. Hence, we take a 2-place predicate symbol not occurring in α, say P^2, and construct the formula

$$\delta = \exists x[\forall y(\exists zQ(x,y,z)\supset P(x,y))\supset\forall yP(x,y)]$$

For the construction of its prenex normal form we have to extend the scope of the first $\forall$ to the whole implication, by one application of Th. 18, (Sect. 3.3) replacing $\forall$ with $\exists$, and obtain

$$\exists x\exists y[(\exists zQ(x,y,z)\supset P(x,y))\supset\forall yP(x,y)].$$

Now we extend the scope of $\exists z$ to the whole of the first implication, by applying Th. 19 (of 3.3), replacing $\exists$ with $\forall$, and obtain:

$$\exists x\exists y[\forall z(Q(x,y,z)\supset P(x,y))\supset\forall yP(x,y)].$$

Now we extend the scope of the first "$\forall$" to the whole implication by using Th. 18 (of 3.3), changing $\forall$ with $\exists$:

$$\exists x\exists y\exists z[(Q(x,y,z)\supset P(x,y))\supset\forall yP(x,y)].$$

Finally, by one application of Th. 16 (of 3.3) and relettering y in w we get:

$$\delta_1 = \exists x\exists y\exists z\forall w[(Q(x,y,z)\supset P(x,y))\supset P(x,w)],$$

and this is the prenex normal form of δ, hence $\vdash\delta\equiv\delta_1$, and therefore $\vdash\alpha\equiv\delta_1$. But δ_1 is also in Skolem normal form. Therefore, δ_1 is also Skolem normal form of α.

Exercise. Construct the Skolem normal forms of the formulas $\alpha - \delta$ of Exercise 1, Sect. 3.4.1.

3.5. Completeness of FOL^{ax}

3.5.1. The idea of completeness

Similar to PL^{ax} (comp. Ch. 1, 3.3.5.1), for FOL^{ax} two different sorts of completeness are definable: *syntactical* and *semantic*. An axiomatic system is syntactically complete if and only if for any formula α of its language the following holds: $\vdash \alpha$ or $\vdash \neg\alpha$. As can be seen such a meaning of completeness does not hold for FOL^{ax}, since by soundness of FOL^{ax} (cf. 3.2.2), if $\vdash \alpha$, then $\models \alpha$, equivalently, if $\not\models \alpha$, then $\nvdash \alpha$. But if α is the atomic formula P(x), then neither P(x) nor $\neg$P(x) is valid, hence, by soundness, there is a formula of L_{FOL}, P(x), such that $\nvdash$ P(x) and $\nvdash \neg$P(x).

The semantic sense of the completeness of an axiomatic system, suited for FOL^{ax}, is defined as follows.

Definition. *An axiomatic system is semantic complete if and only if for any formula* α *of its language the following holds*: if $\models \alpha$, *then* $\vdash \alpha$.

Together with the soundness of FOL^{ax} the following equivalence is obtained:

(Eq) $\models \alpha$ if and only if $\vdash \alpha$.

The completeness of FOL^{ax} and the remarkable result known as the Löwenheim-Skolem theorem are simple corollaries of a basic and more general result concerning first-order theories. Let us see.

3.5.2. First-order theories

What is usually called a first-order theory *T*[46] is a proper or improper extension of FOL^{ax}. The axioms of *T* are those of FOL^{ax}, also called *logical axioms*, and the *proper* axioms of *T* also called *non-logical* axioms of *T*. Therefore, FOL^{ax} is a first-order theory with no proper axioms.

---

[46] "First-order" is used in order to distinguish *T* from the theories in which the predicate symbols are admitted as arguments for predicate symbols and allow the quantification over predicate or functional symbols. And therefore the only syntactic entities, object to quantification, are the symbols for individual variables. Examples of first-order theories: Peano Arithmetic axiomatized, group theory, partial order theory.

Terminology (standard)

The reference to a first-order theory *T* may be given in the following (*equivalent*) ways:

(a) By specifying its axioms and rules of deduction (as in 3.1)

(b) By identifying *T* with the set of its theorems, i.e., if this is the case:

$$T \vdash \alpha \text{ (or } \vdash_T \alpha\text{) iff } \alpha \in T.$$

Definition 1. *T is called* ***axiomatizable*** (*or axiomatic*) *if there is a recursive set of sentences* Γ *such that for any sentence* α: $\alpha \in T$ *iff* $\Gamma \vdash \alpha$. *T is* ***finitely axiomatizable*** *if* Γ *is finite.*

Definition 2. *T is* ***complete*** *if and only if for any closed formula* α (*of its language*): $\alpha \in T$ *or* $\neg\alpha \in T$.

Definition 3. *T is* ***consistent*** *if there is a sentence* α *such that* $T \nvdash \alpha$.

Definition 4. *T is* ***decidable*** *if and only if T is recursive.*

Definition 5. *A first-order theory* T^* *is an* ***extension*** *of a theory T if both theories have the same symbols* (*i.e., the same language*), *and the following holds: If* $T \vdash \alpha$, *then* $T^* \vdash \alpha$.[47]

3.5.3. Lemmas

Lemma 1. *The set of expressions of a first-order theory T is denumerable.*

Let *Symb* be the set of symbols of a first-order theory *T* (or the alphabet of its syntax). An *expression* is a finite sequence of symbols of *T*.

In order to get an enumeration of all expressions of *T* to any symbol of *T* is assigned a distinct odd number.[48] If $s_1,\ldots,s_k \in Symb$, then the Gödel number of the expression Exp: $s_1s_2\ldots s_k$ is $g(\text{Exp}) = 2^{g(s_1)} \cdot 3^{g(s_2)} \cdot \ldots \cdot p_k^{g(s_k)}$, where p_k is the k*th* prime number. As can be seen the different symbols have different Gödel numbers, and (by uniqueness of the factorization of integers into prime numbers) the distinct expressions have distinct Gödel numbers. By this coding we get an enumeration of all expressions (in order of their assigned codes). Moreover, this enumeration is effective.

By proceeding similarly, we can effectively enumerate the sets of

---

[47] The notions of Def. 1 – Def. 5 will be essentially used in Sect. 5.2 (below) and further on in Ch. 3. “Recursive/decidable sets of sentences” (in Def1 and Def.4) means “recursive/decidable sets of the Gödel *numbers* of the respective sets of *sentences*”.

[48] Called its *Gödel number* or its *code*. For details, comp. Ch. 3, Sect. 4.1 (below).

formulas, closed formulas, formulas containing just one free variable, terms, closed terms. And if T is an *axiomatic* theory, then the set Th of Gödel numbers of its theorems is effectively enumerable.

Lemma 2. *Let T be a first-order theory, let α be a closed formula of its language. Then: If $T\nvdash\neg\alpha$, then $T^{*}=T\cup\{\alpha\}$ is consistent.*

Proof (*reductio*). Suppose that $T\nvdash\neg\alpha$ and T^{*} is inconsistent. Then for some β, $T^{*}\vdash\beta\wedge\neg\beta$, i.e., T, $\alpha\vdash\beta\wedge\neg\beta$. Whence, by Deduction Theorem (Corollay a)) $T\vdash\alpha\supset(\beta\wedge\neg\beta)$. Therefore, $T\vdash\neg\alpha$ (by PL), contradicting the hypothesis.

Evidently, under the same hypothesis lemma also holds in the following form: If $T\nvdash\alpha$, then $T^{*}=T\cup\{\neg\alpha\}$ is consistent.

Lemma 3. *Let T be a consistent and complete first-order theory. Then*

(1) $T\nvdash\alpha$ iff $T\vdash\neg\alpha$, i.e.,

(a) If $T\vdash\neg\alpha$, then $T\nvdash\alpha$ (by consistency).

(b) If $T\nvdash\alpha$, then $T\vdash\neg\alpha$ (by completeness).

(2) *If* $\vdash\alpha$, *then* $T\vdash\alpha$.

This result follows directly from the fact that α is a theorem of FOL^{ax} and T extends FOL^{ax}.

(3) $T\vdash\alpha\wedge\beta$ *iff* $T\vdash\alpha$ *and* $T\vdash\beta$.

(Use $\vdash(\alpha\wedge\beta)\supset\alpha$ and $\vdash(\alpha\wedge\beta)\supset\beta$ and (2)).

(4) $T\vdash\alpha\vee\beta$ *iff* $T\vdash\alpha$ *or* $T\vdash\beta$.

$T\vdash\alpha\vee\beta$ iff $T\nvdash\neg(\alpha\vee\beta)$ (by (1)) iff $T\nvdash\neg\alpha\wedge\neg\beta$ (by PL) iff $T\nvdash\neg\alpha$ or $T\nvdash\neg\beta$ (by (3) with $\neg\alpha$ and $\neg\beta$ instead of α and β and PL) iff $T\vdash\alpha$ or $T\vdash\beta$ (by compl. of T).

(5) $T\vdash\alpha\supset\beta$ *iff* $T\nvdash\alpha$ *or* $T\vdash\beta$.

$T\vdash\alpha\supset\beta$ iff $T\vdash\neg\alpha\vee\beta$ (by PL) iff $T\vdash\neg\alpha$ or $T\vdash\beta$ (by (4)) if $T\nvdash\alpha$ (by cons.) or $T\vdash\beta$.

Lindenbaum's Lemma. *If T is a consistent first-order theory, then there is a consistent and complete extension T^{*} of it.*

Proof. By Lemma 1 there is an enumeration En: $\beta_1,\beta_2,...$ of all *closed* formulas of L_T. Define an infinite sequence of first-order theories as follows:

$T_0=T$, and for any $i\geq 0$

$$T_{i+1} = \begin{cases} T_i \cup \{\beta_{i+1}\}, \ if \ \ T_i \nvdash \neg\beta_{i+1} \\ T_i, \ if \ \ T_i \vdash \neg\beta_{i+1}. \end{cases}$$

T^* = theory having as axioms all the axioms of all the theories T_i (i = 0,1,2,...).

By construction, T_{i+1} extends T_i, T^* extends all the T_i and therefore T^* extends T.

T^* is consistent.

By definition of deduction relation $\Gamma\vdash\alpha$ (i.e., by its finitism[49]) a proof of inconsistency of T^* is a proof of inconsistency in some T_i. (Since such a proof uses only a *finite* number of formulas of L_{T^*} and therefore only a *finite* number of the closed formulas in *En*. By the above enumeration each $\beta_i \in En$ has an index in *En*. So let us take the highest-indexed T_i containing the highest-indexed formula $\beta_i \in En$ such that T_i is inconsistent.) Hence if all the T_i are consistent, then T^* is also consistent. Let us show, by induction on i, that all the T_i are consistent.

Basis. T_0 is consistent (by definition of T_0, since $T_0 = T$ and T is consistent by hypothesis of the lemma).

Induction. Suppose that T_i is consistent. By definition of T_{i+1} we have: either $T_{i+1} = T_i$, and then T_{i+1} is consistent (by supposition), or $T_{i+1} \neq T_i$. In this case $T_{i+1} = T_i \cup \{\beta_{i+1}\}$ (by def. of T_{i+1}); and this means that $T_i \nvdash \neg\beta_{i+1}$, and therefore T_{i+1} is also consistent (by Lemma 2).

A similar proof, whose main ingredient is the same Lemma 2, can be given as follows. The *Basis* holds by definition, and for the induction step suppose that T_i is consistent, $T_{i+1} \neq T_i$ (i.e., $T_{i+1} = T_i \cup \{\beta_{i+1}\}$) and that T_{i+1} is inconsistent. Then $T_i \vdash \neg\beta_{i+1}$ (by Lemma 2), and then, by definition of T_{i+1}, $T_{i+1} = T_i$, contrary to our supposition ($T_{i+1} \neq T_i$).

T^* is complete

For a formula $\beta_{i+1} \in En$ we can say: either $T_i \vdash \beta_{i+1}$ or $T_i \nvdash \beta_{i+1}$. Whence, in the first case, $T^* \vdash \beta_{i+1}$, or in the second case, $T_{i+1} = T_i \cup \{\neg\beta_{i+1}\}$ (by Lemma 2), and therefore $T^* \vdash \neg\beta_{i+1}$.

Remark 1. A proof of completeness of T^* can also be given using Lemma 2 and the following *result*: if T_i is consistent and α is closed, then $T_i \cup \{\alpha\}$ is

[49] Comp. 3.1. Def. 3.

consistent or $T_i \cup \{\neg\alpha\}$ is consistent (prove this result and construct a proof for the completeness of T^* !).

Remark 2. Even if the given consistent theory is axiomatic, this fact does not guarantee by itself that T^* is also axiomatic. In order to be so, T^*, as a formal system, must have a proof procedure, i.e., a method whose application gives us an answer to the question whether or not the formula $\neg\beta_{i+1}$ is provable in T_i (i.e., an answer for deciding whether or not a formula $\beta_i \in En$ is an axiom of T^*; and this means that the set of axioms of T^* might not be decidable.

3.5.4. Completeness theorem for FOL^{ax}

As we mentioned above, two main results, completeness theorem for FOL^{ax} and Löwenheim-Skolem Theorem for first-order theories, are corollaries on a basic result on the consistent first-order theories, according to which for every such theory there exists a model of the cardinality $\aleph_0$ (i.e., a model whose domain is denumerable) in which all of its axioms/ theorems are true. In this case we also say that the theory *has a model*.

Theorem. *Every consistent first-order theory has a model of* $\aleph_0$*-cardinality.*

Proof.[50] Let T be a consistent first-order theory, let L_T be its language. Now we need a denumerable set $\{c_0,c_1,...\}$ of *new* constant symbols. Nothing is lost in generality if we suppose that from the denumerable set of constant symbols $\{a_1,a_2,...\}$ of L_T, the theory T only uses the symbols of the form a_{2i} $(i = 0,1,2,...)$. So all the other symbols, i.e., the symbols of the form a_{2i+1}, remain available as the new constant symbols. Let us refer to this denumerable set by $\{c_0,c_1,...\}$ (i.e., $c_i = a_{2i+1}$).

Let $\beta_0(x_0),\beta_1(x_1),...$ be an enumeration of all formulas containing only one free variable.[51] Let $En = \exists x_0\beta_0(x_0),\exists x_1\beta_1(x_1),...,$ be an enumeration of all *closed* formulas of the specified type, where the constant symbol c_i does not occur in $\beta_0(x_0),...,\beta_i(x_i)$, respectively.

---

[50] The first proof of the completeness of FOL^{ax} was given by Gödel [1930]. Some of the other proofs of this theorem can be found in Hilbert and Ackerman [1928] Sec. ed. §10, 6th ed., Ch. III, §8, Hilbert and Bernays [1939], §3, J. Herbrand [1930], L. Henkin [1949], S.C. Kleene [1967], §49, §52 and [1952] §§72, 73; and some proofs using topology and algebra, e.g. Rasiowa and Sikorski [1951], A. Robinson [1951], Beth [1951]. The proof given here is Henkin-type (cf. L. Henkin [1949] with a simplification given by G. Hasenjaeger [1953]; comp. also H. Scholz and G. Hasenjaeger [1961], §§108-114).

[51] Such an enumeration is possible, by 3.5.3. Lemma 1.

Let Imp_i: $\exists x_i\beta_i(x_i)\supset\beta_i(c_i/x_i)$; $i = 0,1,2,...$ or more simple Imp_i: $\exists x_i\beta_i(x_i)\supset\beta_i(c_i)$. As can be observed, the constant symbol c_i *does not occur neither in T nor* in Imp_{i-1}. Let T^* be the theory obtained from T by adding the denumerable set of implications of the form Imp_i; i.e., $T^* = T\cup\{\mathrm{Imp}_0, \mathrm{Imp}_1,...\}$. And this is equivalent to define T^* as follows: $T_0^* = T$, $T_{i+1}^* = T_i^* \cup\{\mathrm{Imp}_i\}$ and $T^* = \cup T_i^*$ $(i = 0,1,...)$.

1. T^* *is consistent.*

By the finiteness of proof-relation, the proof of the consistency of T^* is reducible to the proof that every T_i^* is consistent. And an argument by induction for this fact runs as follows.

Basis. T_0^* is consistent, since by definition $T_0^* = T$ and T is consistent by the hypothesis of the theorem.

Induction. Suppose that T_i^* is consistent and must show that T_{i+1}^* is also consistent. We prove by contraposition that if T_{i+1}^* is not consistent, then T_i^* is not consistent. Hence suppose that T_{i+1}^* is not consistent. It follows that any formula is provable in T_{i+1}^* (by PL). It follows that $T_{i+1}^* \vdash \mathrm{Non\ Imp}_i$, and therefore $T_i^*, \mathrm{Imp}_i \vdash \mathrm{Non\ Imp}_i$. But Imp_i is a closed formula, and then $T_i^* \vdash \mathrm{Imp}_i\supset\mathrm{Non\ Imp}_i$. Whence, by PL, $T_i^* \vdash \mathrm{Non\ Imp}_i$, i.e., $T_i^* \vdash \neg(\exists x_i\beta_i(x_i)\supset\beta_i(c_i))$. And then, by PL, $T_i^* \vdash \exists x_i\beta_i(x_i)$ and $T_i^* \vdash \neg\beta_i(c_i)$. If in the proof of this last formula we replace every occurrence of c_i with a variable x not occurring in the proof, then $T_i^* \vdash \neg\beta_i(x)$, and therefore, by Gen, $T_i^* \vdash \forall x\neg\beta_i(x)$. Now, since the formulas $\neg\beta_i(x)$ and $\neg\beta_i(x_i)$ are similar, it follows that $T_i^* \vdash \forall x_i\neg\beta_i(x_i)$, (by 3.2.6, Lemma*), contradicting the above result that $T_i^* \vdash \exists x_i\beta_i(x_i)$. Hence T_i^* would be inconsistent. Therefore, all the T_i^* are consistent and then T^* is also consistent. Now, since T^* is consistent, then there is a *consistent* and *complete extension* T^{cc} of T^* (by Lindenbaum's Lemma).

Definition. M = ⟨D,i⟩ *is a* ***Herbrand model*** *for a language L if*:

1. D *is the denumerable set of the* ***closed*** *terms of* L, *i.e., of terms not containing any individual variable. By 3.5.3, Lemma 1 there exists an enumeration of them.*

2. *For any closed* t: $t^i = t$, *i.e.,*

(a) $c^i = c$

(b) $[f^n(t_1 ... t_n)]^i = f^n(t_1 ... t_n)$. This follows by a simple argument by induction on the length (lh) of t. If lh(t) = 1, then as in (a) $t^i = t$. If lh(t) > 1, then we have the case (b), $t = f^n(t_1 ... t_n)$, and suppose that $t^i = t$ for all terms t_i such that $lh(t_i) < lh(t)$. Then we have:

$$[f^n(t_1 ... t_n)]^i = (f^n)^i(t_1^i ... t_n^i) = (f^n)^i(t_1 ... t_n) = f^n(t_1 ... t_n).$$

For atomic formulas $R^n(t_1 ... t_n)$ we define the truth in M of such formulas as follows:

$$[R^n(t_1 ... t_n)]^i = 1 \text{ (in M) iff } T^{cc} \vdash R^n(t_1 ... t_n).$$

All that remains to show is that M so defined is a model for *T*, and this means that we must prove that for any closed formula of L_T the following equivalence holds (where for $[\alpha]^i = 1$ we write simply $\alpha = 1$).

(EQ) $\alpha = 1$ (in M) iff $T^{cc} \vdash \alpha$,

since T^{cc} is a complete and consistent extension of *T*.

Proof. (induction on the complexity of α[52])

Basis. Compl(α) = 0. Then α is an atomic formula and (EQ) holds for it by definition.

Induction. Suppose that (EQ) holds for any formula whose compl<*k* and show that it holds for a formula α such that compl(α) = *k*. Three cases must be considered, according to the form of α: $\alpha = \neg\beta$, $\alpha = \beta \supset \gamma$ and $\alpha = \forall x\beta$, respectively.

1. $\alpha = \neg\beta$.

$\alpha = 1$ in M iff $\neg\beta = 1$ in M iff $\beta = 0$ iff $T^{cc} \nvdash \beta$ (by ind. hyp.) iff $T^{cc} \vdash \neg\beta$ (by 3.5.3, Lemma 3(1)) iff $T^{cc} \vdash \alpha$.

2. $\alpha = \beta \supset \gamma$.

$\alpha = 1$ in M iff $\beta \supset \gamma = 1$ in M iff $\beta = 0$ in M or $\gamma = 1$ in M iff $T^{cc} \nvdash \beta$ or $T^{cc} \vdash \gamma$ (by ind. hyp.) iff $T^{cc} \vdash \beta \supset \gamma$ (by 3.5.3, Lemma 3(5)) iff $T^{cc} \vdash \alpha$.

---

[52] Remember, "complexity of α" (abbr. compl(α)) is the number of all occurrences of operators (connectives and quantifiers) of α).

3. $\alpha = \forall x\beta(x)$. Let $\beta(x)$ be $\beta_k(x_k)$ in the above enumeration of all formulas containing only one free variable.

To prove (EQ) for α means to prove the following conditionals:

(a) If $\forall x_k\beta_k(x_k) = 1$ in M, then $T^{cc} \vdash \forall x_k\beta_k(x_k)$.

(b) If $T^{cc} \vdash \forall x_k\beta_k(x_k)$, then $\forall x_k\beta_k(x_k) = 1$ in M.

Proof (a). (*contraposition*).

(1) $\nvdash \forall x_k\beta_k(x_k)$[53]; hyp.

(2) $\vdash \neg\forall x_k\beta_k(x_k)$; (1), by completeness of T^{cc}.

(3) $\vdash \exists x_k\neg\beta_k(x_k)$; (2) by 3.3. Th. 1c).

(4) $\vdash \exists x_k\neg\beta_k(x_k)\supset\neg\beta_k(c_k)$; Imp_k is an axiom of T^{cc}.

(5) $\vdash \neg\beta_k(c_k)$; (3), (4), MP.

(6) $\nvdash \beta_k(c_k)$; (5) by consistency of T^{cc}.

(7) $\beta_k(c_k) = 0$; (6) by ind. hyp.

(8) $\forall x_k\beta_k(x_k) = 0$; (7); by 2.3, Th. 7 (Corollary 1).

(b). (*reductio*).

(1) $\vdash \forall x_k\beta_k(x_k)$; hyp.

(2) $\forall x_k\beta_k(x_k) = 0$ in M; hyp.

(3) $\vdash \forall x_k\beta_k(x_k)\supset\beta_k(t)$; Ax. 4 of FOL^{ax}.

(4) $\vdash \beta_k(t)$; (1), (3) for any $t \in D$ (since t is closed and then it is free for x_k in β_k).

(5) There is a $t \in D$ such that $\beta_k(t) = 0$ in M; by (2) and the fact that $\beta_k(t)$ is closed.

(6) $\nvdash \beta_k(t)$; (5) by ind. hyp.

(7) $\vdash \neg\beta_k(t)$; (6), by completeness of T^{cc}.

But (4) and (7) are contradictory.

Therefore, since card(D) = $\aleph_0$, M = $\langle D,i\rangle$ is a denumerable model for T^{cc}. But $T \subseteq T^{cc}$. And then M = $\langle D,i\rangle$ is also a denumerable model for *T*.

Remark. Even if *T* were axiomatic, T^{cc} would not be necessarily axiomatic, since the definition of T^{cc} supposes Lindenbaum's Lemma and then for each step of extension of *T* we must be able to decide whether or

---

[53] All these derivations are in T^{cc}; here we omit "T^{cc}" set in front of "$\vdash$".

not $\neg\beta_i$ is provable in T_i (comp. 3.5.3 Lindenbaum's Lemma, Remark 2).

Corollary 1. (completeness of a first-order theory). *Let T be a first-order theory. Then the following holds: If* $\models\alpha$, *then* $T\vdash\alpha$ (*where* $\alpha \in L_T$).

Proof. By 2.3 Theorem 2, a formula α is valid iff its *universal closure* is valid, and by Ax4 and Gen of FOL^{ax}, α is provable iff its universal closure is provable. So, the proof of this corollary can be reduced to the proof for closed formulas.

Suppose, by *reductio*, that α is a valid formula of L_T and α is not provable in *T*. Then, by 3.5.3 Lemma 2, $T^* = T \cup \{\neg\alpha\}$ is consistent. And then, by the above theorem, T^* has a model. Since $\neg\alpha \in T^*$ (and then $T^* \vdash \neg\alpha$) it follows that $\neg\alpha = 1$ in M. On the other hand, since α is valid (by hypothesis), α is also true in M. But, by 2.1, (Eq) (after Def. 8), α and $\neg\alpha$ cannot be simultaneously true in M. Therefore, α is provable in *T*.

Corollary 2. (completeness of FOL^{ax}). *If* $\alpha \in L_{FOL}$, *then: If* $\models\alpha$, *then* $FOL^{ax}\vdash\alpha$.

Proof (by Corollary 1 and $FOL^{ax} \subseteq T$).

Together with soundness of FOL^{ax} (comp. 3.2.2, Theorem 2), the following holds:

$\vdash\alpha$ iff $\models\alpha$,

for any $\alpha \in L_{FOL}$.

Corollary 3. (*Löwenheim-Skolem Theorem*). *Let T be a first-order theory. Then if T has a model, then T has a denumerable model.*

Proof. Suppose *T* has a model M. Then *T* is consistent (otherwise $T\vdash\alpha$ and $T\vdash\neg\alpha$, and then $\alpha = 1$ in M and $\alpha = 0$ in M (impossible, by 2.1 (Eq) after Def. 8). Whence, by the above theorem, *T* has a denumerable model.

4. First-order logic with identity (FOL^{ax}_{id})

4.1. Leibniz Principle (LP)

Let L^{id}_{FOL} be the language obtained by adding to L_{FOL} a new 2-place predicate symbol "=" called *identity*. Let $\alpha(x,x)$ be a formula of L_{FOL} and $\alpha(x,y)$ be the formula of L_{FOL} obtained from $\alpha(x,x)$ by replacing numerically arbitrary[54] free occurrences of x in α(x,x) by y, with the proviso that y is free for those occurrences of x. By a Leibniz Formula we understand a formula of L^{id}_{FOL} of the following form

$$x = y \supset (\alpha(x,x) \supset \alpha(x,y))$$

Leibniz Principle (LP). *Any Leibniz Formula is valid, i.e.,*

$$\models x = y \supset (\alpha(x,x) \supset \alpha(x,y)).$$

Proof. A simple argument is this. Let $M = \langle D,i\rangle$ be an arbitrary model for L^{id}_{FOL} and let μ be an arbitrary assignment in M. If $[x = y]^{i,\mu} = 1$ then $x^{\mu} = y^{\mu}$. Hence if $[\alpha(x,x)]^{i,\mu} = 1$, then $[\alpha(x,y)]^{i,\mu} = 1$, whence $[\alpha(x,x) \supset \alpha(x,y)]^{i,\mu} = 1$. Therefore, $[x = y \supset (\alpha(x,x) \supset \alpha(x,y))]^{i,\mu} = 1$.

The proviso "y is free for those occurrences of x in α(x,x)" in obtaining α(x,y) from α(x,x) is necessary. Otherwise, we obtain the formulas of the following form:

$$x = y \supset (\exists y\neg(x = y) \supset \exists y\neg(y = y)),$$

and such a formula is not valid (it is only 1-valid). Nevertheless, by preserving the proviso, a valid formula of L^{id}_{FOL} can be obtained:

$$x = y \supset (\exists z\neg(x = z) \supset \exists z\neg(y = z)).$$

If in **LP** we pas from α(x,x) and α(x,y) to ¬α(x,x) and ¬α(x,y), then, using Repl_{FOL}, we get $x = y \supset (\alpha(x,y) \supset \alpha(x,x))$, which with **LP** gives $x = y \supset (\alpha(x,x)\equiv\alpha(x,y))$, and, accordingly, a stronger form of Leibniz Principle

$$LP^*\colon \models x = y \supset (\alpha(x,x)\equiv\alpha(x,y))$$

4.2. First-order logic with identity (FOL^{ax}_{id})

If we add a Leibniz Formula and the formula $\forall x(x = x)$ to the axioms of FOL^{ax} and preserve the deduction rules of FOL^{ax} (MP and Gen), then we obtain a *first-order logic with identity* (FOL^{ax}_{id}). I.e., $FOL^{ax}_{id} = FOL^{ax}$ plus

Ax6. $\forall x(x = x)$; reflexivity of identity

Ax7. $x = y \supset (\alpha(x,x) \supset \alpha(x,y))$; substitutivity of identicals.

[54] In Leibniz' terms, "ubivis", not "ubique"; cf. B. Couturat [1903], 255.

4.3. Syntactical considerations

In FOL_{id}^{ax} the following holds:

1. $\vdash t = t$, for any term t (Refl)
   (Ax6, Ax4 and MP)
2. Symmetry of "=". $\vdash x = y \supset y = x$ (Symm)
   (1) $x = y \supset (x = x \supset y = x)$; as a special case of Ax7, where $\alpha(x,x)$ is $x = x$ and $\alpha(x,y)$ is $y = x$.
   (2) $x = x \supset (x = y \supset y = x)$; (1) PL
   (3) $x = y \supset y = x$; (2), 1, MP
3. Transitivity of "=". $\vdash x = y \supset (y = z \supset x = z)$ (Trans)
   (1) $y = x \supset (y = z \supset x = z)$; as a special case of Ax7, where $\alpha(y,y)$ is $y = z$ and $\alpha(y,x)$ is $x = z$
   (2) $x = y \supset (y = z \supset x = z)$; (1), 2 PL
4. $\vdash y = x \supset (z = x \supset y = z)$
   (1) $y = x \supset (x = z \supset y = z)$; Trans
   (2) $x = z \supset (y = x \supset y = z)$; (1) PL
   (3) $z = x \supset x = z$; Symm
   (4) $z = x \supset (y = x \supset y = z)$; (2) (3) PL
   (5) $y = x \supset (z = x \supset y = z)$; (4) PL
5. $\vdash x = y \supset f(z,x,w) = f(z,y,w)$.
   (1) $f(z,x,w) = f(z,x,w)$; Refl
   (2) $x = y \supset (u = f(z,x,w) \supset u = f(z,y,w))$; Ax7
   (3) $x = y \supset (f(z,x,w) = f(z,x,w) \supset f(z,x,w) = f(z,y,w))$; (2), Subst_x: $f(z,x,w)/u$ (cf. Sect. 3.2.4.1).
   (4) $f(z,x,w) = f(z,x,w) \supset (x = y \supset f(z,x,w) = f(z,y,w))$; (3) PL
   (5) $x = y \supset f(z,x,w) = f(z,y,w)$; (1) (4) MP

(A similar proof can be given using Gen with respect to u in (2), then Ax5, Ax4 and PL; exercise).

6. $\vdash x = y \supset P(z,x,w) \supset P(z,y,w)$; as a special case of Ax7

The formula in 6 is just the axiom Ax7 for atomic formulas of L_{FOL}^{id}, not containing functional symbols and in which P(z,y,w) comes from P(z,x,w) by replacing exactly one occurrence of x by y.

A stronger form of 6 is the following:

6^*. $\vdash (x_1 = y_1 \wedge ... \wedge x_k = y_k) \supset (P(x_1,...,x_k) \supset (P(y_1,...,y_k))$.

(Take $k = 2$ and detail the argument!). Whence by substitutions t_i/x_i and t'_i/y'_i ($1 \leq i \leq k$) we get:

6^{**}. $\vdash (t_1 = t'_1 \wedge ... \wedge t_k = t'_k) \supset (P(t_1,...,t_k) \supset P(t'_1,...,t'_k))$.

In a similar fashion 5 can be strengthened to

5^*. $\vdash (t_1 = t'_1 \wedge ... \wedge t_k = t'_k) \supset (f(t_1,...,t_k) = f(t'_1,...,t'_k))$

$5^*.1$. $\vdash x = y \supset t = t^*$; where t^* is obtained from t by an arbitrary replacement of x by y. $5^*.1$ is provable by induction on the complexity of t (exercise).

Let now $t_1,...,t_k$ be the argument-terms in $P(t_1,...,t_k)$ and $t'_1,...,t'_k$ be the argument-terms in $P(t'_1,...,t'_k)$, where $P(t'_1,...,t'_k)$ is resulting from $P(t_1,...,t_k)$ by arbitrary replacements of x with y. By $5^*.1$ we have

$\vdash x = y \supset t_1 = t'_1, ..., \vdash x = y \supset t_k = t'_k$

and then, by PL,

$\vdash x = y \supset (t_1 = t'_1 \wedge ... \wedge t_k = t'_k)$

Now, using 6^{**} and PL we get

7. $\vdash x = y \supset P(t_1,...,t_k) \supset P(t'_1,...,t'_k)$

Similar to 6, 7 represents Ax7 for arbitrary atomic formulas of L^{id}_{FOL}.

Remark. Using some suitable substitutions in 6 and Symm, by PL the following stronger form of 6 can be derived

6.1. $\vdash x = y \supset (P(z,x,w) \equiv P(z,y,w))$,

and then the stronger form of 7, i.e.,

7.1. $\vdash x = y \supset P(t_1,...,t_k) \equiv P(t'_1,...,t'_k)$

But if Ax7 holds for any atomic formulas of L^{id}_{FOL} (as 7 shows) it also holds for an arbitrary formula of L^{id}_{FOL}, as we'll see below (by Theorem).

8. $\vdash \exists y x = y$

(1) $x = y \supset x = y$; Rule_p (Sect. 3.2.1)

(2) $x = y \supset \exists y x = y$; (1); Sect. 2.3 (syntactic counterpart of Th. 12)

(3) $x = x \supset \exists y x = y$; (2), Subst_x (Sect. 3.2.4.1)

(4) $x = x$; Ax6

(5) $\exists y x = y$; (3), (4), MP

Evidently, from this theorem, by Gen, we also have $\vdash \forall x \exists y x = y$.

9. $\vdash \forall y(x = y \supset \alpha(y)) \supset \exists y(x = y \wedge \alpha(y))$

(1) $(x = y \supset \alpha(y)) \supset (x = y \supset (x = y \wedge \alpha(y)))$; Rule$_p$ (Sect. 3.2.1)

(2) $\forall y[(x = y \supset \alpha(y)) \supset (x = y \supset (x = y \wedge \alpha(y)))]$; (1), Gen

(3) $\forall y(x = y \supset \alpha(y)) \supset \forall y(x = y \supset (x = y \wedge \alpha(y)))$; (2), Sect. 3.3, Th.6

(4) $\forall y(x = y \supset (x = y \wedge \alpha(y))) \supset (\exists y x = y \supset \exists y(x = y \wedge \alpha(y)))$; Sect. 3.3, Th. 8

(5) $\forall y(x = y \supset \alpha(y)) \supset (\exists y x = y \supset \exists y(x = y \wedge \alpha(y)))$; (3), (4), PL

(6) $\exists y x = y \supset [\forall y(x = y \supset \alpha(y))) \supset \exists y(x = y \wedge \alpha(y))]$; (5), PL (permutation of premisses)

(7) $\forall y(x = y \supset \alpha(y)) \supset \exists y(x = y \wedge \alpha(y))$; (6), 8, MP

Lemma.[55] *For any formula* $\alpha(x) \in L_{FOL}$ *there are two equivalent formulas of* L^{id}_{FOL} *such that*

1. $\vdash \alpha(x) \equiv \exists y(x = y \wedge \alpha(y))$
2. $\vdash \alpha(x) \equiv \forall y(x = y \supset \alpha(y))$.

The result of 9 suggests how to give some simple proofs for 1 and 2. Namely, it is enough to prove the following conditionals:

(a) $\exists y(x = y \wedge \alpha(y)) \supset \alpha(x)$

(b) $\alpha(x) \supset \forall y(x = y \supset \alpha(y))$,

since by Th. 9 and (a) we derive the *converse* of (b), which together with (b) gives 2. Similarly, by 9 and (b) we get the converse of (a), which together with (a) gives 1.

Proof (a).

(1) $x = y \supset y = x$; by 2

(2) $y = x \supset (\alpha(y) \supset \alpha(x))$; Ax7

(3) $x = y \supset (\alpha(y) \supset \alpha(x))$; (1), (2), PL

(4) $(x = y \wedge \alpha(y)) \supset \alpha(x)$; (3), PL

(5) $\exists y(x = y \wedge \alpha(y)) \supset \alpha(x)$; (4), Th. 13 (Sect. 2.3) (since y is not free in $\alpha(x)$, under hypothesis that *all* occurrences of y in $\alpha(y)$ are replaced by x in $\alpha(x)$) (or by Sect.3.3, Th.11).

(b).

(1) $x = y \supset (y = x)$; by 2

[55] This result is very useful in the construction of the diagonal sentences, avoiding in this way the use of the substitution function, whose arithmetization is a relatively complicated task; comp. R. Smullyan [1992], Ch. II, §6, [1993], Ch. 0, §0.

(2) $y = x \supset (\neg\alpha(y) \supset \neg\alpha(x))$; Ax7
(3) $x = y \supset (\neg\alpha(y) \supset \neg\alpha(x))$; (1), (2), PL
(4) $(x = y \wedge \neg\alpha(y)) \supset \neg\alpha(x)$; (3), PL
(5) $\alpha(x) \supset \neg(x = y \wedge \neg\alpha(y))$, (4), PL
(6) $\alpha(x) \supset (x = y \supset \alpha(y))$; (5), PL
(7) $\alpha(x) \supset \forall y(x = y \supset \alpha(y))$; (6), Th. 11 (Sect. 2.3) (or by Sect.3.3, Th.12) (since y is not free in $\alpha(x)$)

Evidently, since 1 and 2 are theorems of FOL_{id}^{ax}, their universal closures will also be theorems of FOL_{id}^{ax}.

Theorem. *If* Ax6 *holds and* Ax7 *holds for any atomic formula of* L_{FOL}^{id}, *then* Ax7 *generally holds.*

Proof (induction on the complexity of $\alpha(x,x)$).

Basis. $\alpha(x,x)$ is atomic. Then Ax7 holds by assumption.

Induction. Assume that Ax7 holds for any formula $\alpha(x,x)$ whose complexity is less than *n*, and show that it also holds for formulas of complexity *n*.

1. $\alpha(x,x)$ is $\neg\beta(x,x)$
   (1) $y = x \supset (\beta(x,y) \supset \beta(x,x))$; by ind. hyp.
   (2) $y = x \supset (\neg\beta(x,x) \supset \neg\beta(x,y))$; (1) PL
   (3) $x = y \supset (\alpha(x,x) \supset \alpha(x,y))$; (2), Symm, PL
2. $\alpha(x,x)$ is $\beta(x,x) \supset \gamma(x,x)$
   (1) $x = y \supset y = x$; Symm.
   (2) $y = x \supset (\beta(x,y) \supset \beta(x,x))$; ind. hyp.
   (3) $x = y \supset (\beta(x,y) \supset \beta(x,x))$; (1), (2), PL
   (4) $x = y \supset (\gamma(x,x) \supset \gamma(x,y))$; ind. hyp.
   (5) $x = y \supset ((\beta(x,y) \supset \beta(x,x)) \wedge (\gamma(x,x) \supset \gamma(x,y))$; (3), (4), Rule_{P} (Sect. 3.2.1)
   via $\models (p \supset (q \supset r)) \supset ((p \supset (s \supset t)) \supset (p \supset ((q \supset r) \wedge (s \supset t))))$:
   + Subt. $x = y/p$, $\beta(x,y)/q$, $\beta(x,x)/r$, $\gamma(x,x)/s$ and $\gamma(x,y)/t$, and MP (twice)

   Let FORM be the formula in (5)

   (6) $\text{FORM} \supset [x = y \supset ((\beta(x,x) \supset \gamma(x,x)) \supset (\beta(x,y) \supset \gamma(x,y))]$ by Rule_{P}, *via* $\models [p \supset ((q \supset r) \wedge (s \supset t))] \supset [(p \supset ((r \supset s)) \supset (q \supset t))]$, + the above mentioned substitutions.
   (7) $x = y \supset ((\beta(x,x) \supset \gamma(x,x)) \supset (\beta(x,y) \supset \gamma(x,y)))$; (5), (6), MP
   i.e., $x = y \supset (\alpha(x,x) \supset \alpha(x,y))$.

3. $\alpha(x,x)$ is $\forall z\beta(x,x,z)$

(1) $x = y \supset (\beta(x,x,z) \supset \beta(x,y,z))$; by ind. hyp.

(2) $\forall z[x = y \supset (\beta(x,x,z) \supset \beta(x,y,z))]$; (1) Gen

(3) $x = y \supset \forall z(\beta(x,x,z) \supset \beta(x,y,z))$; (2) Ax5, PL

(4) $x = y \supset (\forall z\beta(x,x,z) \supset \forall z\beta(x,y,z))$; (3), Sect. 3.3, Th. 6, PL, i.e.,

(5) $x = y \supset (\alpha(x,x) \supset \alpha(x,y))$.

What shows this theorem is the following thing: if Ax6 holds and Ax7 (i.e., 7 above) holds for any atomic formulas of L^{id}_{FOL}, then

$$FOL^{ax} + Ax6 + Ax7 = FOL^{ax}_{id}.$$

Remark. Both FOL^{ax} and FOL^{ax}_{id} are examples of first-order theories, expressed in the language L_{FOL} and L^{id}_{FOL}, respectively.

Similar to FOL^{ax}, FOL^{ax}_{id} is a complete and undecidable theory.[56]

5. Undecidability of FOL

As we saw (3.2.2 and 3.5.4), by soundness and completeness of FOL^{ax}, for any formula α of L_{FOL} we have: $\models \alpha$ iff $\vdash \alpha$. On the other hand, in FOL the following holds: α is not valid iff $\neg\alpha$ is satisfiable, and $\neg\alpha$ is valid iff α is not satisfiable, respectively. Usually, in FOL to both problems, that of establishing the *validity* (*provability*) and its dual, that of *satisfiability* of a formula α of L_{FOL} we refer by using the label "the decision problem". While for PL this problem has been solved (comp. Ch. 1, 3.3.5.2), since PL^{ax} is both sound and complete and for establishing the validity of a formula we have some methods (comp. Ch. 1, 2.7), the things in FOL are more complex, given that the above equivalence $\models \alpha$ iff $\vdash \alpha$ does not yield a solution of the problem whether an arbitrary given formula α is valid/ provable or not, since in order to establish this fact we need a procedure (method) allowing us to decide which is the case.

In fact, for FOL, while for some special classes of formulas a solution of the decision problem was given, this problem as a whole remains unsolved. To begin with let us see some special cases where this problem

[56] For proofs we refer the reader to D. Hilbert and W. Ackermann [1972] (6. Aufl), Ch. 3, §9 and S.C. Kleene [1952], §73 and [1967], §52.

was solved (Sect. 5.1), and then to approach the question of the undecidability of FOL (Sect. 5.2).

5.1. Some decidability results in FOL

5.1.1. Decidability of the set of open formulas

As we saw (comp. 2.3. Th. 4 and Th. 4*) if $\alpha(p_1,...,p_n)$ is a formula of L_{PL} and $\alpha^*(\beta_1/p_1,...,\beta_n/p_n)$ is a formula of L_{FOL} resulting from α by substituting the *open* formulas $\beta_1,...,\beta_n$ for $p_1,...,p_n$ respectively, then the following holds:

(1) $\models \alpha$ iff $\vdash \alpha^*$

(2) Sat(α) iff Sat(α^*).

The equivalence (2) says: an open formula α^* of L_{FOL} is *satisfiable* if and only if α^* results from a satisfiable formula α of L_{PL} by substitution.

But if *k* is the number of terms in α^* (where each subterm of a term is counted as a distinct term), then the following equivalence also holds:

(3) Sat α^* iff $Sat_k\alpha^*$.

This equivalence passes for a *numerical* criterion for satisfiability for open formulas of L_{FOL}.

Moreover, by passing from α^* to $\neg\alpha^*$ and by (metalinguistic) negation of both members of "iff" so obtained, we derive

(4) $\models \alpha^*$ iff $\models_k \alpha^*$.

Finally, by the above equivalences, if α is an *open* formula of L_{FOL}, then we may effectively establish whether or not α is valid/ satisfiable or *k*-valid/ satisfiable. It follows that if such a formula is valid, then it is provable in FOL^{ax} using only Axioms (1)-(3) and MP (by 3.2.1). Hence for open formulas of FOL there is a decision procedure.[57]

5.1.2. Decidability of the set of k-valid and k-satisfiable formulas of L_{FOL}

Let α be an arbitrary formula of L_{FOL}. Let us take the simpler case when α does not contain functional symbols.

Theorem 1. *For any finite k, the set of k-valid formulas of* L_{FOL} *is deciable.*

Proof.

(1) Construct the prenex normal form of α, let us call it α^*. By 3.4.1

[57] Comp. also Sect. 2.3, Theorem 4 and Theorem 4*.

Theorem 1, soundness and completeness of FOL^{ax}, we have:

$\models \alpha$ iff $\models \alpha^*$.

(2) And then $\models_k \alpha$ iff $\models_k \alpha^*$; from (1) and def. of validity of α.

(3) Since $|D| = k$, all the quantifiers of the prefix of α^* can be (equivalently) eliminated (comp. 2.1. Def. 10), from α^* we obtain α^{**} such that $\models_k \alpha^* \equiv \alpha^{**}$.

(4) By (2) and (3) it follows that $\models_k \alpha$ iff $\models_k \alpha^{**}$.[58]

(5) Now, since α^{**} is an open formula, it follows that $\models \alpha^{**}$ iff $\models_k \alpha$ (by (4) of the 5.1.1). Hence, by 5.1.1, the *k*-validity of α is decidable.

Theorem 2. *For any finite k, the set of k-satisfiable formulas of* L_{FOL} *is decidable.*

Proof. By *mimicking* the preceding proof, we have:

(1) $\models \alpha \equiv \alpha^*$; by 3.4.1 Theorem 1. And then

(2) $Sat_k(\alpha \equiv \alpha^*)$; (1) (any valid formula is of course satisfiable).

(3) $Sat_k \alpha$ iff $Sat_k \alpha^*$; (2).

(4) $\models \alpha^* \equiv \alpha^{**}$; where α^{**} is an open formula (comp. 2.1. Def. 10).

(5) $Sat_k \alpha^*$ iff $Sat_k \alpha^{**}$; from (4).

(6) $Sat_k \alpha$ iff $Sat_k \alpha^{**}$; (3), (5), PL.

Now, by 5.1.1, *k*-satisfiability of α is decidable.

5.1.3. Decidability of the monadic first-order logic

Monadic first-order logic (MFOL)

The language L_M of MFOL is that of L_{FOL} with the following exceptions: it contains only monadic predicate symbols and does not contain function symbols. Hence what counts as a term of L_M are the symbols for individual variables and the symbols for constants. This does imply that what counts as atomic formula is a syntactic construction of the form $R^1(t)$, where R^1 is a monadic predicate symbol and t is a term of L_M. All the semantic notions of MFOL are that of FOL. The axiomatic system $MFOL^{ax}$ is just FOL^{ax} but with formulas in L_M. The proofs of soundness, consistency and completeness of $MFOL^{ax}$ follow from the similar proofs for FOL^{ax} (try to detail!).

[58] $\models_k \alpha^* \equiv \alpha^{**}$ does imply $\models_k \alpha^*$ iff $\models_k \alpha^{**}$ (the converse does not hold!); whence by (2) follows (4).

The formal system $MFOL^{ax}$ is *decidable*. This means that there is an effective method by which for any formula $\alpha \in L_M$ we can say whether or not it is a theorem of $MFOL^{ax}$. Let us detail.

In what follows by "α is D-valid" we understand α is *k*-valid, where $k = |D|$, i.e., α is true in any model $M = \langle D, int \rangle$,[59] where $|D| = k$.

Theorem 1. *Let* α *be a formula of* L_{FOL}. *Then if* α *is* D-*valid, then* α *is* D'-*valid, where* $D' \subseteq D$.[60]

Proof.[61] Let $\alpha(P_1,...,P_k,x_1,...,x_n)$ be a formula of L_{FOL} containing the predicate symbols $P_1,...,P_k$ (with arbitrary argument places), and $x_1,...,x_n$ be individual variables. Let $M = \langle D, int \rangle$ and $M' = \langle D', int' \rangle$ be two models such that $D' \subseteq D$. Let us consider a correspondence between D and D' in the following way: to any element $a \in D$ we associate an element $a' \in D'$. If a is an element of D' then we set $a' = a$. But if a is not an element of D', then $a' = e$ where e is an arbitrary chosen element of D'. Let μ be an assignment in M and μ' be its corresponding assignment in M'. Now, if P_i is an *r*-place predicate symbol of α ($i = 1,...,k$), then to any predicate $P_i^{int'}$ defined on D' we (uniquely) associate an *r*-place P_i^{int} defined on D, such that $[P_i(x_1,...,x_r)]^{int,\mu} = 1$ iff $[P_i(x_1,...,x_r)]^{int',\mu'} = 1$ (where $x_j^{\mu} = a_j$ and $x_j^{\mu'} = a'_j$ ($j = 1,...,r$) (i.e., both sentences have the same truth values).

Let $\alpha(P_1,...,P_k,x_1,...,x_n)$ be an arbitrary formula of L_{FOL}, let $a_1,...,a_n$ be arbitrary elements of D.

Assertion. $[\alpha(P_1,...,P_k,x_1,...,x_n)]^{int,\mu} = [\alpha(P_1,...,P_k,x_1,...,x_n)]^{int',\mu'}$.

Proof. (induction on compl(α)).

Basis. Compl(α) = 0; i.e., α is an atomic formula. Then the Assertion holds by the definition of $[P_i(x_1,...,x_n)]^{int,\mu}$.

Induction. Compl(α) > 0. Now we have, accordingly, the following three cases:

(1) $\alpha = \neg\beta = \neg\beta(P_1,...,P_k,x_1,...,x_n)$.

---

[59] $M = \langle D, int \rangle$ is just $M = \langle D, i \rangle$ (in the sense of 2.1). The use here of "int" instead of "i" is more convenient, improving the readibility of the symbolic expressions.

[60] The theorem holds for any formula of L_{FOL}, but for the argument we have in view we only consider the formulas not containing the functional terms.

[61] The proof given here is Hilbert-Ackermann-style (comp. Hilbert and Ackermann [1946], 2. verb. Aufl., Ch. III, §12; Hilbert and Ackermann [1972], 6 Aufl., Ch. 3, §11, Theorem XXVII).

(2) $\alpha = \beta \supset \gamma = \beta(P_1,...,P_k,x_1,...,x_n) \supset \gamma(P_1,...,P_k,x_1,...,x_n)$

(3) $\alpha = \forall y\beta = \forall y\beta(P_1,...,P_k,x_1,...,x_n,y)$.

In the first case we have:

$[\alpha]^{int,\mu} = 1$ iff $[\neg\beta(P_1,...,P_k,x_1,...,x_n)]^{int,\mu} = 1$ iff

iff $[\beta(P_1,...,P_k,x_1,...,x_n)]^{int,\mu} = 0$ iff

iff $[\beta(P_1,...,P_k,x_1,...,x_n)]^{int',\mu'} = 0$ (by ind. hyp.)

iff $[\neg\beta(P_1,...,P_k,x_1,...,x_n)]^{int',\mu'} = 1$

iff $[\alpha]^{int',\mu'} = 1$.

In the second case,

$[\alpha]^{int,\mu} = 1$ iff $[\beta(P_1,...,P_k,x_1,...,x_n) \supset \gamma(P_1,...,P_k,x_1,...,x_n)]^{int,\mu} = 1$

iff $[\beta(P_1,...,P_k,x_1,...,x_n)]^{int,\mu} = 0$ or

$[\gamma(P_1,...,P_k,x_1,...,x_n)]^{int,\mu} = 1$

iff $[\beta(P_1,...,P_k,x_1,...,x_n)]^{int',\mu'} = 0$ or

$[\gamma(P_1,...,P_k,x_1,...,x_n)]^{int',\mu'} = 1$ (by ind. hyp.)

iff $[\beta(P_1,...,P_k,x_1,...,x_n) \supset \gamma(P_1,...,P_k,x_1,...,x_n)]^{int',\mu'} = 1$

iff $[\alpha]^{int',\mu'} = 1$.

In the third case,

$[\alpha]^{int,\mu} = 1$ iff $[\forall y\beta(P_1,...,P_k,x_1,...,x_n,y)]^{int,\mu} = 1$

iff for any assignment ν y-variant of μ:

$[\beta(P_1,...,P_k,x_1,...,x_n,y)]^{int,\nu} = 1$ (by 2.1, Def. 4a))

iff for any assignment ν' (in D'):

$[\beta(P_1,...,P_k,x_1,...,x_n,y)]^{int',\nu'} = 1$ (by ind. hyp.)

iff $[\forall y\beta(P_1,...,P_k,x_1,...,x_n,y)]^{int',\mu'} = 1$ (by 2.1, Def. 4a))

iff $[\alpha]^{int',\mu'} = 1$.

Finally, since M and μ are arbitrary, it follows by Assertion that if $\alpha(P_1,...,P_k,x_1,...,x_n)$ is D-valid, then this formula is also D'-valid.

Corolary. *If* α *is k-satisfiable, then* α *is* k+1-*satisfiable*.

Proof.

(1) If $\models_{k+1}\neg\alpha$, then $\models_k\neg\alpha$; by Theorem 1.

(2) If $\not\models_k\neg\alpha$, then $\not\models_{k+1}\neg\alpha$; (1), by contraposition

(3) If $\mathrm{Sat}_k(\alpha)$, then $\mathrm{Sat}_{k+1}(\alpha)$; (2) by Sect. 2.1, 1 (after Def. 11)

Theorem 2. MFOL^{ax} *is decidable*.

To begin with, let us illustrate the idea of the proof of this theorem. If we have *k* properties, $\mathrm{Prop}_1,\ldots,\mathrm{Prop}_k$, and a number *n* of objects such that each object either has or has not a property Prop_i (i = 1,...,*k*), then the greatest number of classes into which these properties can classify the things is 2^k (such that each thing does belong exactly to one class). If $n < 2^k$ at least one class will be empty and if $n > 2^k$ at least one class does contain more than one object.[62] This is a *matter of logic*, since if we have three properties and *n* objects, for example, the total ways these properties can be arranged, regarding their belonging or not to an arbitrary object, is exactly $2^3 = 8$. If *a*, for example, is an object having Prop_1 and Prop_3 and not having Prop_2, then *a* belongs strictly to a class of all objects with the same properties (i.e., having Prop_1 and Prop_3 and not having Prop_2). So, no matter how great *n* is, the highest[63] number of possible classes will be 2^3.

Let $\alpha(\mathrm{P}_1,\ldots,\mathrm{P}_k)$ be an arbitrary formula of $\mathrm{L_M}$, containing the distinct monadic predicate symbols $\mathrm{P}_1,\ldots,\mathrm{P}_k$. Since 2^k-validity of a formula α can be established (cf. 5.1.2), for proving the theorem is enough to prove the following sentence: *if* α *is* 2^k*-valid, then* α *is a valid formula of* $\mathrm{L_M}$.

Moreover, supposing that this sentence holds, by contraposition we get, equivalently: if α is not valid, then α is not 2^k-valid. And in this case, *via* Theorem 1, if we want to find the domains for which α is valid, we only have to test the validity of α for domains with $1,2,\ldots,2^k-1$ elements.

Proof.[64] Suppose that $\alpha(\mathrm{P}_1,\ldots,\mathrm{P}_k)$ is 2^k-valid. Let $\mathrm{M} = \langle \mathrm{D}, \mathrm{int} \rangle$ be an arbitrary model and μ an arbitrary assignment in M.

Let $\mathrm{P}_1^{\mathrm{int}},\ldots,\mathrm{P}_k^{\mathrm{int}}$ be the respective monadic predicates (properties Prop_i, in the above illustration) defined on D. If $a \in \mathrm{D}$, then $\mathrm{P}_i^{\mathrm{int}}(a) = 1$ (i = 1,...,*k*) means that the object *a* in D has the property $\mathrm{P}_i^{\mathrm{int}}$, or, equivalently, the sentence $\mathrm{P}_i^{\mathrm{int}}(a)$ is true. Now, if we refer to *a* by $a = \mathrm{x}^{\mu}$, then we can write equivalently $[\mathrm{P}_i(\mathrm{x})]^{\mathrm{int},\,\mu} = 1$.

---

62 This is the well-known "pigeon-holes principle".

63 "The highest", since some combination of properties may happen not to belong to an object.

64 Analog to the preceding theorem, the proof given here is Hilbert-Ackermann-style; cf. Hilbert and Ackermann [1946], 2. Aufl., Ch. 3, §12, 120-1; 6. Aufl., Ch. 3, §11, 127-8.

Let us construct the (at most) 2^k classes of objects to which we refer in our illustration, in the following way. Two objects a_j and a_m are in same class iff they have the same properties, i.e., $P_i^{int}(a_j)$ iff $P_i^{int}(a_m)$, for all $i = 1,...,k$. Now, if $a \in D$, let us consider the sentences $P_1^{int}(a),...,P_k^{int}(a)$, with the corresponding truth values $v_1,...,v_k$, where $v_i = 1$ (true) or 0 (false). Let $a' = (v_1,...,v_k)$ be the k-tuple associated to a. Let D' be the set of all these k-tuples associated with the elements of D. Of course, the number of all these k-tuple is $n \leq 2^k$.

Let now $\beta(P_1,...,P_k,x_1,...,x_n)$ be a subformula[65] of $\alpha(P_1,...,P_k)$, containing the free variables $x_1,...,x_n$. Let $a_1,...,a_n$ be arbitrary objects from D, let $a'_1,...,a'_n$ be their corresponding elements in D'.[66] Let M' = ⟨D',int'⟩ be a model and μ' the assignment in M' corresponding to the assignment μ in M.

Assertion. $[\beta(P_1,...,P_k,x_1,...,x_n)]^{int,\mu} = [\beta(P_1,...,P_k,x_1,...,x_n)]^{int',\mu'}$ (i.e., *the two formulas have the same truth value*).

In order to prove the assertion, let us define the predicates $P_i^{int'}$ by the following equivalence:

$$[P_i(x)]^{int',\mu'} = 1 \text{ iff } [P_i(x)]^{int,\mu} = 1.$$

i.e., $P_i^{int'}(a') = 1$ iff $P_i^{int}(a) = 1$; with $x^{\mu'} = a'$, $x^{\mu} = a$; and this means that the predicate $P_i^{int'}$ is true of the class $a' = (v_1,...,v_k)$ (where $a' \in D'$) iff $v_i = 1$ iff the predicate P_i^{int} is true of the object $a \in D$.

The proof of *Assertion* is given by induction on the compl(β).

Basis. Suppose compl(β) = 0. Then β is an atomic formula of the form $P_i(x)$. And then $[P_i(x)]^{int',\mu'} = 1$ iff $P_i^{int'}(x^{u'}) = 1$ iff $P_i^{int'}(a'_j) = 1$ iff $P_i^{int'}(v_1,...,v_k) = 1$ iff $v_i = 1$ iff $P_i^{int}(a_j) = 1$ iff $[P_i(x)]^{int,\mu} = 1$.

Induction. Compl(β)>0. Then β is of the form ¬γ, γ⊃δ or ∀yγ, where by hypothesis the theorem holds for γ and δ.

1. $\beta = \neg\gamma$. $[\neg\gamma(P_1,...,P_k,x_1,...,x_n)]^{int,\mu} = 1$ iff

---

[65] I.e., any part of α, satisfying the definition of a formula, including α itself.

[66] If β contains constant symbols, and if $c^{int} = a_j$, for example, then a'_j is the corresponding class in D'. On the other hand, this argument also contains the case when α has no free variables.

iff $[\gamma(P_1,...,P_k,x_1,...,x_n)]^{int,\mu}=0$ iff $[\gamma(P_1,...,P_k,x_1,...,x_n)]^{int',\mu'}=0$

(ind. hyp.) iff $[\neg\gamma(P_1,...,P_k,x_1,...,x_n)]^{int',\mu'}=1$.

2. $\beta=\gamma\supset\delta$ (similar).

3. $\beta=\forall y\gamma(P_1,...,P_k,x_1,...,x_n,y)$.

$[\forall y\gamma(P_1,...,P_k,x_1,...,x_n,y)]^{int,\mu}=1$ iff for any ν y-var. of μ:

$[\gamma(P_1,...,P_k,x_1,...,x_n,y)]^{i,\nu}=1$ iff for any ν' y-var. of μ':

$[\gamma(P_1,...,P_k,x_1,...,x_n,y)]^{int',\nu'}=1$ (by ind. hyp.)

iff $[\forall y\gamma(P_1,...,P_k,x_1,...,x_n,y)]^{int',\mu'}=1$.

And then, since α is a subformula of itself, the *Assertion* follows, i.e.,

$[\alpha(P_1,...,P_k)]^{int,\mu}=[\alpha(P_1,...,P_k)]^{int',\mu'}$.

Now, since $\alpha(P_1,...,P_k)$ is 2^k-valid (by hypothesis), then it is also valid for the domain D' whose number of elements is $n\leq 2^k$ (by Theorem 1). Whence, by *Assertion*, since $M=\langle D,int\rangle$ and μ were arbitrary, it follows that $\alpha(P_1,...,P_k)$ is true in any model, and therefore α is a valid formula of L_M, and therefore it is a theorem of $MFOL^{ax}$ (by completeness theorem for $MFOL^{ax}$). Hence $MFOL^{ax}$ is decidable.

Remark. If α is a formula of L_M containing *k* predicate symbols, then the following holds:

(1) $\models_{2^k}\alpha$ iff $\models\alpha$

(2) $Sat_{2^k}\alpha$ iff $Sat\alpha$,

and therefore for such an $\alpha\in L_M$ we have

(1') α is valid, if α is D-valid, where D is finite

(2') α is satisfiable only if α is D-satisfiable, where D is finite.

These sentences do not hold for any formula containing 2-place predicate symbols.

Example.

α: $\forall x\exists yR(x,y)\wedge\neg R(x,x)\wedge\forall x\forall y\forall z((R(x,y)\wedge R(y,z))\supset R(x,z)$.

This formula is satisfiable *only* in a domain D with $|D|=\aleph_0$. And then it is not satisfiable in a domain with finite number of elements. (Take, for example, for R the relation x<y). It follows that its negation, ¬α, is D-

valid only for *finite* D.

5.2. Undecidability of FOL

In spite of some decidability results in FOL^{ax} (sect. 5.1), the FOL^{ax} *as a whole* remains an undecidable first-order theory (in a sense to be made precise below).[67]

To remember (comp. 3.5.2), a theory T is called *axiomatizable* if there is a *recursive* set Γ of sentences (closed formulas) such that $T = \{\alpha \mid \Gamma \vdash \alpha\}$.[68] If Γ is finite, then T is *finitely axiomatizable*. And T is called *decidable* iff the set *Th* of Gödel numbers of its theorems is *recursive*. T is called *recursively undecidable* if and only if *Th* is not recursive. T is *essentially recursively undecidable* if and only if every consistent extension of T (including T itself) is recursively undecidable. T is *essentially* (*recursively*) *incomplete* iff every consistent extension of it (including T itself) has an undecidable sentence.

The two notions "T is complete" and "T is decidable" are connected by the following theorem.

Theorem. *Let T be an axiomatizable theory. Then if T is complete, then T is decidable.*

Proof. If by α_n we understand (here) that n is the Gödel number of α, then $Th = \{n \mid T \vdash \alpha_n\}$. By 3.5.3 the set *Th* is recursively enumerable. We must prove that if T is complete, then *Th* is recursive. We only take the case when T is consistent. Since otherwise all formulas of L_T would be provable in T and then *Th* will be recursive.[69] Hence, by consistency: $T \nvdash \alpha$ or $T \nvdash \neg\alpha$. Since T is complete (by hypothesis), for any closed formula α: $T \vdash \alpha$ or $T \vdash \neg\alpha$. Together, consistency and completeness do imply $T \nvdash \alpha$ iff $T \vdash \neg\alpha$. Therefore, the set $\widetilde{Th}$ (the complement of *Th*) is the union of two sets: a set S_1 of Gödel numbers of those expressions that are not closed formulas at all, and S_2 of Gödel numbers of those closed formulas whose negations are

[67] This section requires Ch. 3, Sect. 3 and 4.2.5, and could be skipped at a first reading.
[68] To these theories we also refer by using the expression "formal systems". If T is an axiomatizable theory whose language is L_{PA}, then it is also called "the formal system of arithmetic".
[69] This follows from the fact that intuitively we can effectively decide whether a sequence of symbols is a formula/closed formula. And then, *via* Church's Thesis, that such sets are recursive. Or, as it happens with L_{PA}, for recursiveness of "formula" comp. Ch. 3, Sect. 4.1.

provable in T, i.e., $S_2 = \{n \mid neg(n) \in Th\}$. S_1 is a recursive set, since the set of all closed formulas of T is recursive (by the fact that if a set is recursive, then its complement is also recursive). The set S_2 is recursively enumerable since it is obtainable by substituting the negation function *neg* in the recursive enumerable set *Th*. So $\widetilde{Th}$ is recursively enumerable. Now, since both sets, *Th* and $\widetilde{Th}$ are recursively enumerable, it follows that *Th* is recursive.

Let T_1 and T_2 be two first-order theories in the same language L.[70] T_2 is an *extension* of T_1 if all theorems of T_1 are also theorems of T_2. If this is the case, we also say that T_1 is a *subtheory* or *subsystem* of T_2. T_2 is a *finite extension* of T_1 if and only if the following holds: $T_1 = \{\alpha \mid Ax_{T_1} \vdash \alpha\}$ and $T_2 = \{\beta \mid Ax_{T_1} \cup \{\alpha_1,...,\alpha_n\} \vdash \beta\}$, where Ax_{T_1} is the set of axioms of T_1 and $\{\alpha_1,...,\alpha_n\}$ is a *finite* set of formulas which are not axioms of T_1. $T_1 \cup T_2$ is a *consistent* theory if the union of their axioms is a consistent set.

5.2.1. Two axiomatizable subsystems of PA^{ax}

The first subsystem of PA^{ax}, relevant in our considerations (in this book) is the formal system Q (due to Raphael Robinson). It is a finitely axiomatizable system, i.e., it has only a *finite* number of proper axioms (given below, where " ' " is the successor function).[71]

(Q_1) $\neg(x' = 0)$

(Q_2) $x' = y' \supset x = y$

(Q_3) $x + 0 = x$

(Q_4) $x + y' = (x + y)'$

(Q_5) $x \cdot 0 = 0$

(Q_6) $x \cdot y' = x \cdot y + x$

---

[70] This means that all their symbols are common.

[71] In fact, the label "Robinson system Q" is referring to a variety of formal systems of the same kind (i.e., finitely axiomatizable). The system Q just presented by (Q_1)-(Q_9) is that from Boolos, Burgess and Jeffrey [2002], §16.2. But what is usually called "Robinson system Q" is the system formed from the axioms (Q_1)-(Q_6) (from the above list) plus the axiom Q*: $x = 0 \vee \exists y x = y'$ (cf. A. Tarkis, A. Mostowski and R.M. Robinson [1953], 51; comp. also G. Boolos and R. Jeffrey [1991], 158). For the relationship between these two axiomatic constructions, comp. Boolos, Burgess and Jeffrey [2002], §16.4. For other references to "Robinson system Q", comp. *inter alia* R.M. Smullyan [1993], Ch. 0, §7, E. Mendelson [1964], 152 and S.C. Kleene [1952], §49, Lemma 18b.

(Q_7) $\neg(x < 0)$

(Q_8) $x < y' \equiv (x < y \vee x = y)$

(Q_9) $x < y \vee x = y \vee y < x$.

The subsystem of PA^{ax}, with these nine proper axioms, is also called "minimal arithmetic",[72] i.e., the theory consisting in all the theorems provable from (Q_1)-(Q_9), and therefore true in all models of these axioms.

The other formal system (also due to Raphael Robinson) is the system R, whose proper axioms are all formulas of the following forms (where *m* and *n* are arbitrary natural numbers):

Ω_1: $\overline{m} + \overline{n} = \overline{k}$, where $m + n = k$

Ω_2: $\overline{m} \times \overline{n} = \overline{k}$, where $m \cdot n = k$

Ω_3: $\overline{m} \neq \overline{n}$, where $m \neq n$

Ω_4: $x \leq \overline{n} \equiv (x = 0 \vee \ldots \vee x = \overline{n})$

Ω_5: $x \leq \overline{n} \vee \overline{n} \leq x$.

Since Ω_1-Ω_5 are axioms schemes, the system R is not finitely axiomatizable.[73]

Some facts about Q[74]

In what follows let $T \supseteq$Q be any first-order extension of Q. Let *Th* be (as above) the set of Gödel numbers of the theorems of *T*. We say that a set S is *formally expressible* in *T*[75] iff there is a formula $\alpha(x)$ of L_T (with x free) such that for any *n* the following hold:

(a) If $n \in$S, then $T \vdash \alpha(\overline{n})$.

(b) If $n \notin$S, then $T \vdash \neg\alpha(\overline{n})$.

Lemma. *If $T \supseteq$Q is consistent, then Th is not expressible in T.*

Proof (*reductio*). Suppose that $T \supseteq$Q is consistent and *Th* is expressible in *T* by, say, a formula TH(x) (with x free). Let $\neg$TH(x) be a formula of L_T. By Diagonal Lemma (Ch. 3, 4.2.2.1) there is a closed formula G of L_T such that

(*) $T \vdash G \equiv \neg TH(\overline{g})$, where *g* is the Gödel number of G.

A short argument shows us that G is a theorem of *T*, i.e., $T \vdash$G. For suppose that $T \nvdash$G. Then, by (b), it follows that $T \vdash \neg TH(\overline{g})$, whence, by (*), we have $T \vdash$G, and therefore G is a theorem of *T*. But then $g \in Th$,

[72] Boolos, Burgess and Jeffrey [2002], 208.

[73] This system is not further on used in our considerations.

[74] These facts hold of any consistent extension *T* of Q.

[75] Comp. Ch. 3, Sect. 2.

whence, by (a), $T \vdash \mathrm{TH}(\bar{g})$. And therefore, by (*) and PL, $T \vdash \neg$G. From which it follows that *T* is inconsistent, contrary to the hypothesis.
Hence if *T* is Q itself, it follows that the set *Th* is not expressible in Q.[76]

Theorem 1. *The system* Q *is essentially recursively undecidable.*

Proof. By Lemma, the set *Th* is not expressible in *T*. Given that every recursive set is expressible in *T* (comp. Ch. 3, Sect. 4.1 Eq2), follows that if *Th* is not expressible in *T*, then *Th* is not recursive (by contraposition).

Theorem 2. *Any consistent axiomatizable extension of* Q *is incomplete.*

Proof. By Theorem (above), any complete axiomatizable theory *T* is decidable. But, by Theorem 1, no consistent extension of Q is decidable. Hence any such extension is incomplete (i.e., it has an undecidable sentence, and then Q is essentially recursively incomplete).

Theorem 3. (Church's Theorem). *The set Val of Gödel numbers of the valid sentences of* Q *is not decidable.*

Proof. Let Conj(Q) be the conjunction of all axioms of Q. Let α be a sentence provable in Q. Then $\mathrm{Q} \vdash \alpha$ iff $\mathrm{Conj(Q)} \models \alpha$ iff $\models \mathrm{Conj(Q)} \supset \alpha$. Now, let *n* be the Gödel number of α and *c* be the Gödel number of the formula Conj(Q). Then $f(n)$ will be the Gödel number of the formula $\mathrm{Conj(Q)} \supset \alpha$, and this number is $c * 2^{11} * n$, where "$*$" is the primitive recursive function called juxtaposition (or concatenation).[77] Since *f* is recursive, it follows that for any *n*:

$$n \in Th \text{ iff } f(n) \in Val.$$

And then if *Val* were recursive, then *Th* would be recursive, since *Th* is obtainable from *Val* by substitution of the recursive function *f*. And then Q would be recursively decidable, which it is not (by Theorem 1).

By Theorem 1 and Theorem 3 the sets of theorems and of valid sentences (with the respective Gödel numbers) of any axiomatizable theory are not recursive, equivalently (by Church's Thesis)[78] they are not effectively decidable. But these sets, *Th* and *Val*, are recursively enumerable (comp. 3.5.3).[79]

---

[76] Since Q is consistent, having a model.

[77] Comp. Ch. 3, Sect. 3.2 (example 4); where 11 is the Gödel number of "$\supset$", in the arithmetization of Sect. 4.1 (of Ch. 3).

[78] Note the distinction between "Church's Theorem" and "Church's Thesis". The former, in a variety of its forms, *is* provable, the latter is *not*. By Church's Thesis the intuitive (imprecise) notion "effectively computable function" ("effectively decidable predicate") is coextensive with the formal (precise) one, that of "recursive".

[79] More about recursive enumerable relations, cf. Ch. 3, Sect. 4.2.5.

Intuitively, *Th* and *Val* are only *positively* effectively decidable in the following sense: since the notion of proof is recursive (comp. Ch. 3, Sect. 4.1) and the set of axioms is also recursive (*T* being axiomatizable), by searching through all proofs, *if a formula* α *is provable/valid* one may eventually find out this formula. But since *Th* and *Val* are not recursive, there is no effective procedure (method) for telling whether the Gödel number of any given formula α is in *Th* or *Val*.

On the other hand, if a set S and its complement $\widetilde{\mathrm{S}}$ are both effectively recursively enumerable, then S is a decidable set, since if we have a procedure for establishing when a number $n \in \mathrm{S}$, but this procedure doesn't tell us anything if $n \notin \mathrm{S}$, and we also have a procedure for establishing when $n \notin \mathrm{S}$, without saying anything when $n \in \mathrm{S}$, then by applying alternatively these procedures we may establish if *n is or is not* in S.

5.2.2 Undecidability of FOL

Theorem 1. *Let* T_1 *and* T_2 *be two first-order theories in the language* $\mathrm{L_{PA}}$. *If* T_2 *is a finite extension of* T_1 *and* T_2 *is recursively undecidable, then* T_1 *is recursively undecidable*.

Proof.[80] Let T_1 and T_2 be as in the hypothesis, i.e., $T_1 = \{\alpha \mid \mathrm{Ax}_{T_1} \vdash \alpha\}$ and $T_2 = \{\beta \mid \mathrm{Ax}_{T_1} \cup \{\alpha_1,...,\alpha_n\} \vdash \beta\}$. Let us take $\alpha_1,...,\alpha_n$ be closed formulas. Then we have $T_2 \vdash \beta$ iff $T_1 \cup \{\alpha_1,...,\alpha_n\} \vdash \beta$ iff T_1, $\alpha_1 \wedge ... \wedge \alpha_n \vdash \beta$ (by PL) iff $T_1 \vdash (\alpha_1 \wedge ... \wedge \alpha_n) \supset \beta$ (by Ded. Th.). Let *Conj* be the Gödel number of the expression $(\alpha_1 \wedge ... \wedge \alpha_n)$ and *b* the Gödel number of β. Then the Gödel number of $(\alpha_1 \wedge ... \wedge \alpha_n) \supset \beta$ will be $Conj * 2^{11} * b$.[81] Hence, from the preceding equivalences, it follows that for any formula β the following holds: $b \in Th_{T_2}$ iff $Conj * 2^{11} * b \in Th_{T_1}$. Now, similar to the argument from Theorem 3 (Sect. 5.2.1), if the set Th_{T_1} were recursive and since Th_{T_2} is obtained from Th_{T_1} by substitution, it follows that Th_{T_2} would be recursive, contra hypothesis of recursive undecidability of T_2.

Theorem 2. *Let T be a first-order theory in the language* $\mathrm{L_{PA}}$. *If* $T \cup \mathrm{Q}$ *is consistent, then T is recursively undecidable.*

[80] With some modifications the proof of Th. 1-3 are those of Mendelson [1964], 154-155.

[81] According to the Gödel numbering of Ch. 3, Sect. 4.1.

Proof. Since $T \cup \mathrm{Q}$ is consistent (by hyp.), T is a consistent extension of Q. And since Q is essentially recursively undecidable (by 5.2.1, Theorem 1) it follows that $T \cup \mathrm{Q}$ is recursively undecidable. But $T \cup \mathrm{Q}$ is a finite extension of T, and then T is recursively undecidable (by Theorem 1).

Corollary. *Let* PRED *be a predicate calculus whose language is* $\mathrm{L_{PA}}$. *Then* PRED *is recursively undecidable.*

Proof. Evidently, $\mathrm{PRED} \cup \mathrm{Q} = \mathrm{Q}$. Since Q is consistent (since it has the standard model of $\mathrm{L_{PA}}$) it follows that $\mathrm{PRED} \cup \mathrm{Q}$ is consistent. And then, by Theorem 2, PRED is recursively undecidable.

Theorem 3. (Church's Theorem). $\mathrm{FOL^{ax}}$ *is recursively undecidable*.[82]

Proof. Let us observe that the language of PRED is just $\mathrm{L_{PA}}$, i.e., it only contains a predicate symbol (for identity), a constant symbol (for zero), and three functional symbols (for successor, addition and multiplication). Instead, the language of FOL contains the possible denumerable sets of predicate symbols, functional symbols and constant symbols. Therefore, PRED is a first-order logic but in the language of PA.

Now, by soundness and Gödel's completeness theorems we have: $\mathrm{PRED} \vdash \alpha$ iff α is a valid formula. Similarly, $\mathrm{FOL^{ax}} \vdash \alpha$ iff α is valid. Therefore, $\mathrm{PRED} \vdash \alpha$ iff $\mathrm{FOL^{ax}} \vdash \alpha$. Let $Form_{\mathrm{PRED}}$ be the set of Gödel numbers of the formulas of PRED. Let Th_{PRED} and $Th_{\mathrm{FOL^{ax}}}$ be the set of Gödel numbers of the theorems of PRED and the set of Gödel numbers of the theorems of $\mathrm{FOL^{ax}}$, respectively. As can be observed, $Th_{\mathrm{PRED}} = Th_{\mathrm{FOL^{ax}}} \cap Form_{\mathrm{PRED}}$. Now, if $Th_{\mathrm{FOL^{ax}}}$ were recursive, then since $Form_{\mathrm{PRED}}$ is recursive it follows that Th_{PRED} would be recursive, contrary to the corollary of Theorem 2. Hence $Th_{\mathrm{FOL^{ax}}}$ is not recursive, and this means that $\mathrm{FOL^{ax}}$ is recursively undecidable.

Remark. What is usually called "the pure first-order predicate calculus" ($\mathrm{PFOL^{ax}}$) is $\mathrm{FOL^{ax}}$ with the language not containing functional symbols and constant symbols. The recursive undecidability of this calculus follows from the recursive undecidability of $\mathrm{FOL^{ax}}$ by a "translation" of any formula α of $\mathrm{L_{FOL}}$ in a formula α^* of the language of pure predicate logic $\mathrm{L_{PFOL}}$ such that $\mathrm{FOL^{ax}} \vdash \alpha$ iff $\mathrm{PFOL^{ax}} \vdash \alpha^*$. Since there is a recursive function $f(x)$ such that if n is the Gödel number of α, then $f(n)$ is the Gödel number of α^*, the

[82] A. Church [1936(a)].

recursiveness of the set $Th_{\mathrm{PFOL}^{\mathrm{ax}}}$ would imply the recursiveness of the set $Th_{\mathrm{FOL}^{\mathrm{ax}}}$, contrary to Theorem 3.[83]

[83] For details, comp. E. Mendelson [1964], 155-156.

Chapter 3. FORMAL NUMBER THEORY

1. Peano Arithmetic axiomatized (PA^{ax})

1.1. Language of PA (L_{PA})

An axiomatic system of PA^{ax} is a first-order theory based on Peano's postulates. Its language, L_{PA}, contains a single predicate symbol, "=", two binary function symbols, "+" and "·", a unary function symbol, " ' ", and one constant symbol, "0". Symbolically, $L_{PA}: \langle =, 0, ', +, \cdot \rangle$. These symbols of L_{PA} are used in building up terms and formulas.

Terms of L_{PA}. We define this notion recursively, using the metamathematical symbols t_1 and t_2, in the following way:

1. 0 is a term.
2. Any symbol for a variable x_1, x_2, x_3,... is a term.
3. If t_1 and t_2 are terms, then t_1+t_2 is a term.
4. If t_1 and t_2 are terms, then $t_1 \cdot t_2$ is a term.
5. If t is a term, then t' is a term.

An expression is a term only if it can be so qualified using the clauses 1-5.

Example of terms. By 1 and 5, 0, 0', 0'', 0''',... are terms; they are usually called *numerals* and abbreviated by $0, \overline{1}, \overline{2}, \overline{3}, \ldots$ In general, if *n* is a natural number, then $\overline{n}$ is the corresponding numeral (i.e., 0 followed by *n* strokes). If we want to define numerals recursively, we say: 0 is a numeral and if t is a numeral, then t' is a numeral. By 2 and 5, x'_3 and $\overline{2}$ are terms, and by 3 and 4, $x_1 + \overline{5}$ and $x_2 \cdot \overline{3}$ are terms.

Formulas of L_{PA}. This notion is that defined recursively in Ch. 2,1.1, but this time using only the symbols of L_{PA}, respectively.

1. If t_1 and t_2 are terms, then $t_1=t_1$ is a formula (an atomic one).
2. If α is a formula, then $\neg\alpha$ is a formula.
3. If α and β are formulas, then $\alpha \circ \beta$ is a formula, where "$\circ$" denotes anyone of the binary propositional connectives.[1]

[1] Comp. Ch. 1,1.

4. If α is a formula and x is a variable, then $\forall x\alpha$ and $\exists x\alpha$ are formulas.

An expression is a formula only if can be defined on the basis of clauses 1-4.

Examples of formulas. $x_1 = x_2$, $x_1 + x'_2 = x_4$, $x_2 \cdot x_3 = x'_1$, $\neg(x_1 = x_3)$, $\exists x_3(x_3 + x_1 = x_2)$, $(x_1 + x'_2 = x_4) \supset x_1 = x_5$, $x_2 + \overline{3} = x_5$, $\forall x_1(x_1 = x_2)$.

These definitions of term and formula, given by recursion, show that any term or formula can be built up from 0 and variable symbols by one or more steps, each one of which corresponding to a clause of a definition.

As we saw (Ch. 2, Sect. 2.1), a model for the language of FOL consists of a (non-empty) set D, called the domain, and an interpretation function i, by which the symbols of L_{FOL} get the respective meanings. By a model of L_{PA} we understand the following: the set **N** of natural numbers (its domain) and the symbols of L_{PA} interpreted as follows: "=" is the identity (or equality), "0" is the first natural number, " ' " is the operation "plus 1", and "+", "·" are the usual operation of addition and multiplication, respectively. Simbolically, $M = \langle \mathbf{N},=,0,',+,\cdot \rangle$. It is called the *standard model* for L_{PA}.

1.2. PA^{ax}

The theory we are concerned here, Peano Arithmetic axiomatized, PA^{ax}, is a *proper* extension of the predicate logic. Hence the axioms of PA^{ax} are of two sorts: *logical* axioms and *proper* (non-logical) axioms.

The logical axioms of PA^{ax} are those of FOL^{ax} [2], i.e.,

Ax1. $\alpha \supset (\beta \supset \alpha)$

Ax2. $(\alpha \supset (\beta \supset \gamma)) \supset ((\alpha \supset \beta) \supset (\alpha \supset \gamma))$

Ax3. $(\neg\beta \supset \neg\alpha) \supset (\alpha \supset \beta)$

Ax4. $\forall x\alpha(x) \supset \alpha(t/x)$; where t is a term of L_{PA}, free for x in $\alpha(x)$

Ax5. $\forall x(\alpha \supset \beta) \supset (\alpha \supset \forall x\beta)$; where α does not contain x free.

In these axioms, α, β, and γ are arbitrary formulas of L_{PA}.

The proper axioms of PA^{ax} are:

PA1. $x_1 = x_2 \supset (x_1 = x_3 \supset x_2 = x_3)$

PA2. $x_1 = x_2 \supset x'_1 = x'_2$

[2] Comp. Ch. 2,3.1.

PA3. $\neg(0 = x'_1)$

PA4. $x'_1 = x'_2 \supset x_1 = x_2$

PA5. $x_1 + 0 = x_1$

PA6. $x_1 + x'_2 = (x_1 + x_2)'$

PA7. $x_1 \cdot 0 = 0$

PA8. $x_1 \cdot x'_2 = x_1 \cdot x_2 + x_1$

PA9. $\alpha(0) \supset (\forall x(\alpha(x) \supset \alpha(x')) \supset \forall x\alpha(x))$, where $\alpha(x)$ is any formula of L_{PA}.

The rules of inference of PA^{ax} are those of FOL^{ax}, i.e.,

Modus Ponens (MP) $\dfrac{\alpha, \alpha \supset \beta}{\beta}$

Generalization (Gen) $\dfrac{\alpha}{\forall x\alpha}$

As can be seen, properly speaking, only PA1-PA8 are axioms of PA^{ax}. The other, Ax1-Ax5, PA9 are axioms *schemas*, generating axioms of the respective forms for any specification of the names α, β, γ by the formulas of L_{PA}. The rules of inference also have a schematic formulation.

PA9 can be reshaped, using MP, in the following induction rule:

Ind Rule. From $\alpha(0)$ and $\forall x(\alpha(x) \supset \alpha(x'))$, the formula $\forall x\alpha(x)$ can be inferred.

1.3. Consequences of PA1-PA8

For any terms t_1, t_2, t_3[3] of L_{PA} the following formulas are theorems of PA^{ax}. Actually, they are immediate consequences of the respective axioms of PA^{ax}.

$Conseq_{PA1}$. $t_1 = t_2 \supset (t_1 = t_3 \supset t_2 = t_3)$

$Conseq_{PA2}$. $t_1 = t_2 \supset t'_1 = t'_2$

$Conseq_{PA3}$. $0 \neq t'_1$ (where $\neq$ is the negation of $=$)

$Conseq_{PA4}$. $t'_1 = t'_2 \supset t_1 = t_2$

$Conseq_{PA5}$. $t_1 + 0 = t_1$

$Conseq_{PA6}$. $t_1 + t'_2 = (t_1 + t_2)'$

[3] I.e. t_1, t_2, t_3, t, r, s are **metaterms**; they denote arbitrary terms.

Conseq$_{PA7}$. $t_1 \cdot 0 = 0$
Conseq$_{PA8}$. $t_1 \cdot t'_2 = t_1 \cdot t_2 + t_1$
Proof.
Conseq$_{PA1}$. $t_1 = t_2 \supset (t_1 = t_3 \supset t_2 = t_3)$
(1) $x_1 = x_2 \supset (x_1 = x_3 \supset x_2 = x_3)$; PA1
(2) $\forall x_1 \forall x_2 \forall x_3 \ (x_1 = x_2 \supset (x_1 = x_3 \supset x_2 = x_3))$; (1) Gen
(3) $\forall x_2 \forall x_3 \ (t_1 = x_2 \supset (t_1 = x_3 \supset x_2 = x_3))$; (2) Ax4, MP
(with t_1 subject to the proviso of Ax4)
(4) $\forall x_3 \ (t_1 = t_2 \supset (t_1 = x_3 \supset t_2 = x_3))$; (3) Ax4, MP
(with t_2 subject to the same proviso)
(5) $t_1 = t_2 \supset (t_1 = t_3 \supset t_2 = t_3)$; (4) Ax4, MP

In a similar fashion, using Gen, Ax4 and MP all the other consequences can be proved.

1.4. Theorems

Theorem I. *For any terms* t, t_1, t_2, t_3 *the following formulas are theorems of* PA^{ax}.

1. $t = t$; Reflexivity (Refl)
   (1) $t + 0 = t$; Conseq$_{PA5}$
   (2) $t + 0 = t \supset (t + 0 = t \supset t = t)$; Conseq$_{PA1}$
   (3) $t + 0 = t \supset t = t$; (1), (2), MP
   (4) $t = t$; (1), (3), MP
2. $t_1 = t_2 \supset t_2 = t_1$; Symmetry (Sym)
   (1) $t_1 = t_2 \supset (t_1 = t_1 \supset t_2 = t_1)$; Conseq$_{PA1}$
   (2) $t_1 = t_1 \supset (t_1 = t_2 \supset t_2 = t_1)$; PL (cf. Ch. 1, Sect. 2.2,42)
   (3) $t_1 = t_2 \supset t_2 = t_1$; 1, (2), MP
3. $t_1 = t_2 \supset (t_2 = t_3 \supset t_1 = t_3)$; Transitivity (Trans)
   (1) $t_2 = t_1 \supset (t_2 = t_3 \supset t_1 = t_3)$; Conseq$_{PA1}$
   (2) $t_1 = t_2 \supset t_2 = t_1$; 2
   (3) $t_1 = t_2 \supset (t_2 = t_3 \supset t_1 = t_3)$; (1), (2), PL

Refl, Sym and Trans of "=" show that the identity is an equivalence relation.[4]

4. $t_1 = t_2 \supset (t_3 = t_2 \supset t_1 = t_3)$ (exerc.)

5. $t'_1 + t_2 = (t_1 + t_2)'$

Induction on y in $\alpha(y)$: $x'+y = (x+y)'$

Basis. (1) $x'+0 = x'$; $\text{Conseq}_{\text{PA5}}$

(2) $x + 0 = x$; $\text{Conseq}_{\text{PA5}}$

(3) $(x+0)' = x'$; (2) $\text{Conseq}_{\text{PA2}}$, MP

(4) $x'+0 = (x+0)'$; (1), (3), 4, MP

Hence $\vdash \alpha(0)$

Ind. (1) $x'+y = (x+y)'$; hyp.

(2) $x'+y' = (x'+y)'$; $\text{Conseq}_{\text{PA6}}$

(3) $(x'+y)' = (x+y)''$; (1) $\text{Conseq}_{\text{PA2}}$, MP

(4) $x'+y' = (x+y)''$, (2), (3), 3, MP

(5) $x + y' = (x+y)'$; $\text{Conseq}_{\text{PA6}}$

(6) $(x+y')' = (x+y)''$; (5) $\text{Conseq}_{\text{PA2}}$, MP

(7) $x'+y' = (x+y')'$; (4), (6), 4, MP

Therefore

(8) $x'+y = (x+y)' \vdash x'+y' = (x+y')'$; (1)-(7)

And hence

(9) $(x'+y = (x+y)') \supset (x'+y' = (x+y')')$; (8), Ded. Th.

I.e., $\vdash \alpha(y) \supset \alpha(y')$

Whence, by *Basis* and *Ind*: $\vdash \forall y\alpha(y)$, i.e., $\vdash \forall y(x'+y=(x+y)')$, using Gen, Ax4 and MP it follows $t'_1 + t_2 = (t_1 + t_2)'$.

6. $t = 0 + t$; for any term t

(Induction on x in $\alpha(x)$: $x = 0 + x$; exerc.).

7. $t_1 = t_2 \supset (t_1 + t_3 = t_2 + t_3)$

(Induction on z in $\alpha(z)$: $x = y \supset (x + z = y + z)$; exerc.)

8. $t_1 + t_2 = t_2 + t_1$;

(Induction on y in $\alpha(y)$: $x + y = y + x$)

9. $t_1 = t_2 \supset (t_3 + t_1 = t_3 + t_2)$; exerc.

[4] They are also provable in what is usually called first-order logic with identity; comp. Ch. 2, Sect. 4.2.

10. $(t_1 + t_2) + t_3 = t_1 + (t_2 + t_3)$

(Induction on z in $\alpha(z)$: $(x + y) + z = x + (y + z)$; exerc.).

11. $t_1 = t_2 \supset t_1 \cdot t_3 = t_2 \cdot t_3$

(Induction on z in $\alpha(z)$: $x = y \supset (x \cdot z = y \cdot z)$).

Basis.. (1) $x = y$; hyp

(2) $x \cdot 0 = 0$; Conseq$_{PA7}$

(3) $y \cdot 0 = 0$; Conseq$_{PA7}$

(4) $x \cdot 0 = y \cdot 0$; (2), (3), 4

Hence

(5) $x = y \vdash x \cdot 0 = y \cdot 0$; (1)-(4), and then

(6) $x = y \supset (x \cdot 0 = y \cdot 0)$; (5) Ded. Th.

I.e., $\vdash \alpha(0)$

Ind. (1) $x = y \supset (x \cdot z = y \cdot z)$; hyp

(2) $x = y$; hyp

(3) $x \cdot z = y \cdot z$; (1), (2), MP

(4) $x \cdot z + x = y \cdot z + x$; (3), 7, MP

(5) $x = y \supset (y \cdot z + x = y \cdot z + y)$; 9

(6) $y \cdot z + x = y \cdot z + y$; (2), (5), MP

(7) $x \cdot z + x = y \cdot z + y$; (4), (6), 3

(8) $x \cdot z' = x \cdot z + x$; Conseq$_{PA8}$

(9) $x \cdot z' = y \cdot z + y$; (8), (7), 3, MP

(10) $y \cdot z' = y \cdot z + y$; Conseq$_{PA8}$

(11) $x \cdot z' = y \cdot z'$; (9), (10), 4, MP

Hence

(12) $x = y \supset x \cdot z = y \cdot z \vdash x = y \supset x \cdot z' = y \cdot z'$; (1)-(11)

and then

(13) $(x = y \supset x \cdot z = y \cdot z) \supset (x = y \supset x \cdot z' = y \cdot z')$; (12), Ded. Th.

I.e., $\vdash \alpha(z) \supset \alpha(z')$

Whence, by *Basis* and *Ind*, Gen, Ax4 and MP:

$t_1 = t_2 \supset t_1 \cdot t_3 = t_2 \cdot t_3$

12. $0 \cdot t = 0$; for any term t; exercise

13. $t'_1 \cdot t_2 = t_1 \cdot t_2 + t_2$ (Induction on y in $\alpha(y)$: $x' \cdot y = x \cdot y + y$)

14. $t_1 \cdot t_2 = t_2 \cdot t_1$ (Induction on y in $\alpha(y)$: $x \cdot y = y \cdot x$)

Basis. (1) $x \cdot 0 = 0$; Conseq$_{PA7}$
(2) $0 \cdot x = 0$; 12
(3) $x \cdot 0 = 0 \cdot x$; (1), (2), 4, MP
I.e., $\vdash \alpha(0)$

Ind. (1) $x \cdot y = y \cdot x$; hyp
(2) $x \cdot y + x = y \cdot x + x$; (1), 7, MP
(3) $x \cdot y' = x \cdot y + x$; Conseq$_{PA8}$
(4) $x \cdot y' = y \cdot x + x$; (3), (2), 3, MP
(5) $y' \cdot x = y \cdot x + x$; 13
(6) $x \cdot y' = y' \cdot x$; (4), (5), 4, MP
(7) $x \cdot y = y \cdot x \vdash x \cdot y' = y' \cdot x$; (1)-(6)
(8) $x \cdot y = y \cdot x \supset x \cdot y' = y' \cdot x$; (7) Ded. Th.
I.e., $\vdash \alpha(y) \supset \alpha(y')$

Whence, by *Basis* and *Ind*, Gen, Ax4 and MP:

$$t_1 \cdot t_2 = t_2 \cdot t_1$$

15. $t_1 = t_2 \supset t_3 \cdot t_1 = t_3 \cdot t_2$
(1) $t_1 = t_2 \supset t_1 \cdot t_3 = t_2 \cdot t_3$; 11
(2) $t_1 = t_2$; hyp
(3) $t_1 \cdot t_3 = t_2 \cdot t_3$; (1), (2), MP
(4) $t_3 \cdot t_1 = t_1 \cdot t_3$; 14
(5) $t_3 \cdot t_1 = t_2 \cdot t_3$; (3), (4), 3
(6) $t_2 \cdot t_3 = t_3 \cdot t_2$; 14
(7) $t_3 \cdot t_1 = t_3 \cdot t_2$; (5), (6), 3, MP

Hence
(8) $t_1 = t_2 \vdash t_3 \cdot t_1 = t_3 \cdot t_2$; (1)-(7), and then
(9) $t_1 = t_2 \supset t_3 \cdot t_1 = t_3 \cdot t_2$; (8) Ded. Th.

Remark. A simple consequence of the above results is the following: PA *is a first-order theory with identity*. By Chapter 2, 4.2 and the final remark, in such a theory Ax6: $\forall x(x = x)$ and Ax7: $x = y \supset (\alpha(x, x) \supset \alpha(x, y))$, where the latter expression is a Leibniz Formula, hold. That this is also the case for PA^{ax} follows from some of the above theorems: 1, 2, 3, 4, 7, 9, 11 and 15 and Chapter 2, 4.3 (Theorem).

16. $t_1 \cdot (t_2 + t_3) = t_1 \cdot t_2 + t_1 \cdot t_3$ (distributivity)

(Induction on z in $\alpha(z)$: $x \cdot (y+z) = x \cdot y + x \cdot z$)

Basis. (1) $y + 0 = y$; Conseq$_{PA5}$

(2) $x \cdot (y+0) = x \cdot y$; (1), 15

(3) $x \cdot 0 = 0$; Conseq$_{PA7}$

(4) $x \cdot 0 = 0 \supset x \cdot y + x \cdot 0 = x \cdot y + 0$; 9

(5) $x \cdot y + x \cdot 0 = x \cdot y + 0$; (3), (4), MP

(6) $x \cdot y + 0 = x \cdot y$; Conseq$_{PA5}$

(7) $x \cdot y = x \cdot y + 0$; (6), 2

(8) $x \cdot y = x \cdot y + x \cdot 0$; (7), (5), 4, MP

(9) $x \cdot (y+0) = x \cdot y + x \cdot 0$; (2), (8), 3, MP

Hence $\vdash \alpha(0)$

Ind. (1) $x \cdot (y+z) = x \cdot y + x \cdot z$; hyp

(2) $x \cdot (y+z) + x = (x \cdot y + x \cdot z) + x$; (1), 7

(3) $x \cdot (y+z) + x = x \cdot (y+z)'$; Conseq$_{PA8}$, 2

(4) $(y+z)' = y + z'$; Conseq$_{PA6}$; 2

(5) $x \cdot (y+z)' = x \cdot (y+z')$; (4), 15, MP

(6) $x \cdot (y+z) + x = x \cdot (y+z')$; (3), (5), 3, MP

(7) $(x \cdot y + x \cdot z) + x = x \cdot y + (x \cdot z + x)$; 10

(8) $x \cdot z + x = x \cdot z'$; Conseq$_{PA8}$, 2

(9) $x \cdot y + (x \cdot z + x) = x \cdot y + x \cdot z'$; (8), 9

(10) $(x \cdot y + x \cdot z) + x = x \cdot y + x \cdot z'$; (7), (9), 3

(11) $(x \cdot y + x \cdot z) + x = x \cdot (y + z')$; (2), (6) Conseq$_{PA1}$

(12) $x \cdot (y+z') = x \cdot y + x \cdot z'$; (11), (10); Conseq$_{PA1}$

Hence

(13) $x \cdot (y+z) = x \cdot y + x \cdot z \vdash x \cdot (y+z') = x \cdot y + x \cdot z'$; (1)-(12)

(14) $(x \cdot (y+z) = x \cdot y + x \cdot z) \supset (x \cdot (y+z') = x \cdot y + x \cdot z')$; (13) Ded. Th.

I.e., $\vdash \alpha(z) \supset \alpha(z')$

And then, by *Basis* and *Ind*, Gen, Ax4, MP 16 follows.

17. $(t_2 + t_3) \cdot t_1 = t_2 \cdot t_1 + t_3 \cdot t_1$ (distributivity)

(1) $(t_2 + t_3) \cdot t_1 = t_1 \cdot (t_2 + t_3)$; 14

(2) $t_1 \cdot (t_2 + t_3) = t_1 \cdot t_2 + t_1 \cdot t_3$; 16

(3) $(t_2 + t_3) \cdot t_1 = t_1 \cdot t_2 + t_1 \cdot t_3$; (1), (2), 3

(4) $t_1 \cdot t_2 = t_2 \cdot t_1$; 14

(5) $t_1 \cdot t_2 + t_1 \cdot t_3 = t_2 \cdot t_1 + t_1 \cdot t_3$; (4), 7, MP

(6) $t_1 \cdot t_3 = t_3 \cdot t_1$; 14

(7) $t_2 \cdot t_1 + t_1 \cdot t_3 = t_2 \cdot t_1 + t_3 \cdot t_1$; (6), 9, MP

(8) $t_1 \cdot t_2 + t_1 \cdot t_3 = t_2 \cdot t_1 + t_3 \cdot t_1$; (5), (7), 3, MP

(9) $(t_2 + t_3) \cdot t_1 = t_2 \cdot t_1 + t_3 \cdot t_1$; (3), (8), 3, MP

18. $(t_1 \cdot t_2) \cdot t_3 = t_1 \cdot (t_2 \cdot t_3)$ (associativity of "$\cdot$")

(Induction on z in $\alpha(z)$: $(x \cdot y) \cdot z = x \cdot (y \cdot z)$)

Basis. (1) $(x \cdot y) \cdot 0 = 0$; Conseq$_{PA7}$

(2) $y \cdot 0 = 0$; Conseq$_{PA7}$

(3) $x \cdot (y \cdot 0) = x \cdot 0$; (2), 15

(4) $x \cdot 0 = 0$; Conseq$_{PA7}$

(5) $x \cdot (y \cdot 0) = 0$; (3), (4), 3, MP

(6) $(x \cdot y) \cdot 0 = x \cdot (y \cdot 0)$; (1), (5), 4, MP

Ind. (1) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$; hyp

(2) $(x \cdot y) \cdot z + x \cdot y = x \cdot (y \cdot z) + x \cdot y$; (1), 7, MP

(3) $x \cdot (y \cdot z + y) = x \cdot (y \cdot z) + x \cdot y$; 16

(4) $(x \cdot y) \cdot z + x \cdot y = x \cdot (y \cdot z + y)$; (2), (3), 4, MP

(5) $y \cdot z' = y \cdot z + y$; Conseq$_{PA8}$

(6) $x \cdot (y \cdot z') = x \cdot (y \cdot z + y)$; (5), 15, MP

(7) $(x \cdot y) \cdot z + x \cdot y = x \cdot (y \cdot z')$; (4), (6), 4, MP

(8) $(x \cdot y) \cdot z' = (x \cdot y) \cdot z + x \cdot y$; Conseq$_{PA8}$

(9) $(x \cdot y) \cdot z' = x \cdot (y \cdot z')$; (8), (7), 3, MP

Hence

(10) $(x \cdot y) \cdot z = x \cdot (y \cdot z) \supset (x \cdot y) \cdot z' = x \cdot (y \cdot z')$; (1)-(9), Ded. Th.

I.e., $\vdash \alpha(z) \supset \alpha(z')$

Whence, by *Basis* and *Ind*, Gen, Ax4, MP

19. $t_1 + t_3 = t_2 + t_3 \supset t_1 = t_2$

(Induction on z in $\alpha(z)$: $x + z = y + z \supset x = y$)

Basis. (1) $x + 0 = y + 0$; hyp

(2) $(x + 0)' = (y + 0)'$; (1), Conseq$_{PA2}$, MP

(3) $x + 0 = x$; $Conseq_{PA5}$

(4) $(x + 0)' = x'$; (3) $Conseq_{PA2}$, MP

(5) $y + 0 = y$; $Conseq_{PA5}$

(6) $(y + 0)' = y'$; (5) $Conseq_{PA2}$, MP

(7) $x' = (y + 0)'$; (4), (2), $Conseq_{PA1}$, MP

(8) $x' = y'$; (7), (6), 3

(9) $x = y$; (8), $Conseq_{PA4}$, MP

Hence

(10) $x + 0 = y + 0 \supset x = y$; (1)-(9), Ded. Th.

I.e., $\vdash \alpha(0)$

Ind. (1) $x + z = y + z \supset x = y$; hyp

(2) $x + z' = y + z'$; hyp

(3) $x + z' = (x + z)'$; $Conseq_{PA6}$

(4) $(x + z)' = y + z'$; (3), (2), $Conseq_{PA1}$

(5) $y + z' = (y + z)'$; $Conseq_{PA6}$

(6) $(x + z)' = (y + z)'$; (4), (5), 3, MP

(7) $(x + z)' = (y + z)' \supset x + z = y + z$; $Conseq_{PA4}$

(8) $x + z' = y + z' \supset x = y$; (7) (*via* (3) and (5)), (1), 3, MP

And then

(9) $x + z = y + z \supset x = y \vdash x + z' = y + z' \supset x = y$; (1)-(8)

Hence

(10) $(x + z = y + z \supset x = y) \supset (x + z' = y + z' \supset x = y)$; (9) Ded. Th.

I.e., $\vdash \alpha(z) \supset \alpha(z')$

Whence, by *Basis* and *Ind*, Gen, Ax4, MP:

$t_1 + t_3 = t_2 + t_3 \supset t_1 = t_2$

As we know, the terms $0, 0', 0'', 0''', \ldots$ are called *numerals* and are usually abbreviated (Abbrev.) by $0, \overline{1}, \overline{2}, \overline{3}, \ldots$

Theorem II.

1. $t + \overline{1} = t'$

(1) $t + 0' = (t + 0)'$; $Conseq_{PA6}$

(2) $t + 0 = t$; $Conseq_{PA5}$

(3) $(t + 0)' = t'$; (2) $Conseq_{PA2}$

(4) $t + 0' = t'$; (1), (3), Th. I.3[5] MP

(5) $t + \overline{1} = t'$; (4) Abbrev.

2. $t \cdot \overline{1} = t$; exercise

$t \cdot \overline{2} = t + t$; exercise

and so on

3. $t_1 + t_2 = 0 \supset t_1 = 0 \wedge t_2 = 0$, for any terms t_1, t_2

(Induction on y in $\alpha(y)$: $x + y = 0 \supset (x = 0 \wedge y = 0)$)

Basis. (1) $x + 0 = 0$; hyp

(2) $x + 0 = x$; Conseq_{PA5}

(3) $x + 0 = x \supset (x + 0 = 0 \supset x = 0)$; Conseq_{PA1}

(4) $x = 0$; (1), (2), (3), MP (twice)

(5) $0 = 0$; Th I.1

(6) $x = 0 \wedge 0 = 0$; (4), (5), PL

(7) $x + 0 = 0 \supset (x = 0 \wedge 0 = 0)$; (1)-(6) Ded. Th.

I.e., $\vdash \alpha(0)$

Ind. (1) $x + y' = (x + y)'$, Conseq_{PA6}

(2) $\neg((x + y)' = 0)$; by Conseq_{PA3}

(3) $x + y' = (x + y)' \supset (x + y' = 0 \supset (x + y)' = 0)$; Conseq_{PA1}

(4) $x + y' = (x + y)' \supset (\neg((x + y)' = 0) \supset \neg(x + y' = 0))$; (3), PL

(5) $\neg(x + y') = 0$; (1), (2), (4), MP (twice)

Hence

(6) $x + y' = 0 \supset x = 0 \wedge y' = 0$; (5) by $\neg p \supset (p \supset q)$

I.e., $\vdash \alpha(y')$, and then $\vdash \alpha(y) \supset \alpha(y')$; by $q \supset (p \supset q)$

Whence, by *Basis* and *Ind*, Gen, Ax4 and MP, follows

$t_1 + t_2 = 0 \supset (t_1 = 0 \wedge t_2 = 0)$

4. $t_1 + t_2 = \overline{1} \supset ((t_1 = 0 \wedge t_2 = \overline{1}) \vee (t_1 = \overline{1} \wedge t_2 = 0))$;

(Induction on y in $\alpha(y)$:

$x + y = \overline{1} \supset ((x = 0 \wedge y = \overline{1}) \vee (x = \overline{1} \wedge y = 0))$)

Basis. (1) $x + 0 = \overline{1}$; hyp

(2) $x + 0 = x$; Conseq_{PA5}

(3) $x = \overline{1}$; (1), (2), Conseq_{PA1}

(4) $0 = 0$; Th. I.1

[5] "Th. I.3" stands for part (formula) 3 of Theorem I; similarly for "Th. II", "Th. III" etc.

(5) $(x = \overline{1} \wedge 0 = 0)$; (3), (4), PL

(6) $x + 0 = \overline{1} \vdash x = \overline{1} \wedge 0 = 0$; (1)-(5)

(7) $x + 0 = \overline{1} \supset (x = \overline{1} \wedge 0 = 0)$; (6) Ded. Th.

(8) $x + 0 = \overline{1} \supset (x = \overline{1} \wedge 0 = 0) \vee (x = 0 \wedge 0 = \overline{1})$; (7) PL

I.e., $\vdash \alpha(0)$

Ind. (1) $x + y' = \overline{1}$; hyp

(2) $x + y' = (x + y)'$; Conseq_{PA6}

(3) $(x + y)' = 0'$; (1), (2), Conseq_{PA1}, Abbrev

(4) $x + y = 0$; (3), Conseq_{PA4}

(5) $x = 0 \wedge y = 0$; (4), Th II.3, MP

(6) $(x = 0 \wedge y = 0) \supset x = 0$; PL

(7) $x = 0$; (5), (6), MP

(8) $(x = 0 \wedge y = 0) \supset y = 0$; PL

(9) $y = 0$; (5), (8), MP

(10) $y' = \overline{1}$; (9), Conseq_{PA2}, MP, Abbrev.

(11) $x = 0 \wedge y' = \overline{1}$; (7), (10), PL

(12) $(x = 0 \wedge y' = \overline{1}) \supset ((x = 0 \wedge y' = 1) \vee (x = \overline{1} \wedge y' = 0)$; PL

(13) $(x = 0 \wedge y' = \overline{1}) \vee (x = \overline{1} \wedge y' = 0)$; (11), (12), MP

Hence

(14) $(x + y' = \overline{1}) \supset ((x = 0 \wedge y' = \overline{1}) \vee (x = \overline{1} \wedge y' = 0))$;

(1)-(13), Ded. Th.

I.e., $\vdash \alpha(y')$, and then, by PL

$\vdash \alpha(y) \supset \alpha(y')$

Whence, by *Basis* and *Ind*, Gen, Ax4, MP it follows that

$t_1 + t_2 = \overline{1} \supset ((t_1 = 0 \wedge t_2 = \overline{1}) \vee (t_1 = \overline{1} \vee t_2 = 0))$

5. $t_1 \cdot t_2 = \overline{1} \supset (t_1 = \overline{1} \wedge t_2 = \overline{1})$

(Induction on y in $\alpha(y)$: $x \cdot y = \overline{1} \supset (x = \overline{1} \wedge y = \overline{1})$); (exerc.).

6. $t_1 \neq 0 \supset (t_2 \cdot t_1 = 0 \supset t_2 = 0)$

Let FORM: $y \neq 0 \supset (x \cdot y = 0 \supset x = 0)$.

Let $\alpha(x)$: $x \cdot y = 0 \supset (x = 0)$.

Assume: $y \neq 0$.

We show that under the given assumption, $\alpha(x)$ is deducible.

Basis. (1) $0 \cdot y = 0$; Th. I.12

(2) $0 = 0$; Th. I.1

(3) $0 \cdot y = 0 \supset 0 = 0$; (1), (2), PL

I.e., $\alpha(0)$

Ind. (1) $\alpha(x)$: $x \cdot y = 0 \supset x = 0$; hyp.

(2) $x' \cdot y = 0$; hyp.

(3) $x' \cdot y \neq 0$; since $x' \neq 0$ ($\text{Conseq}_{\text{PA3}}$) and $y \neq 0$ (by assumption)

(4) $x' \cdot y = 0 \supset x' = 0$; (3), PL

I.e., $\alpha(x')$, and then

(5) $\alpha(x) \supset \alpha(x')$; by Ded. Th.

So, by *Basis* and *Ind*, with the condition $y \neq 0$ follows that $\forall x \alpha(x)$ holds; i.e.,

If $y \neq 0$, then $\forall x \alpha(x)$, i.e., $y \neq 0 \vdash \forall x \alpha(x)$, and therefore $\vdash y \neq 0 \supset \forall x \alpha(x)$. Whence $\vdash \forall x (y \neq 0 \supset \alpha(x))$; by Ch. 2, Sect. 3.3, Th. 16. That is $\vdash \forall x$ FORM, from which by Gen, Ax4 and MP 6 may be derived.

7. $\vdash t_1 \neq 0 \supset (t_2 \cdot t_1 = t_3 \cdot t_1 \supset t_2 = t_3)$

Let FORM be $z \neq 0 \supset (x \cdot z = y \cdot z \supset x = y)$.

Let α(x,y) be $x \cdot z = y \cdot z \supset x = y$.

Assume $z \neq 0$ and deduce $\forall x \alpha(x, y)$ by induction on y in α(x,y).

Basis. (1) $x \cdot z = 0 \cdot z$; hyp.

(2) $0 \cdot z = 0$; Th. I.12

(3) $x \cdot z = 0$; (1), (2), Th. I.3

(4) $x = 0$; (3) by Th. II.6

(5) $x \cdot z = 0 \cdot z \supset x = 0$; (1), (4), Ded. Th.

(6) Hence if $z \neq 0$, then $x \cdot z = 0 \cdot z \supset x = 0$, and therefore

(7) $z \neq 0 \vdash \forall x (x \cdot z = 0 \cdot z \supset x = 0)$

Ind. (Assume $\forall x \alpha(x, y)$ and deduce $\forall x \alpha(x, y')$, by induction on x in α(x,y).

(1) $\forall x \alpha(x, y)$; hyp.

(2) α(x,y); (1) Ax4, MP

Basis. (we deduce $\alpha(0, y')$: $(0 \cdot z = y' \cdot z \supset 0 = y')$

(3) $0 \cdot z = y' \cdot z$; hyp.

(4) $z \neq 0$; hyp.
(5) $0 \cdot z = 0$; Th. I.12
(6) $y' \cdot z = 0$; (3), (5), $\text{Conseq}_{\text{PA1}}$
(7) $y' \neq 0 \supset (y' \cdot z = 0 \supset z = 0)$; Th. II.6
(8) $y' \cdot z = 0 \supset (y' \neq 0 \supset z = 0)$; (7), PL
(9) $y' \neq 0 \supset z = 0$; (6), (8), MP
(10) $y' \neq 0$; $\text{Conseq}_{\text{PA3}}$
(11) $z = 0$; (9), (10), MP
But (4) and (11) are contradictory; hence
(12) $z \neq 0,\ 0 \cdot z = y' \cdot z \vdash 0 = y'$
(13) $z \neq 0 \vdash 0 \cdot z = y' \cdot z \supset 0 = y'$; (12), Ded. Th.
I.e., $z \neq 0 \vdash \alpha(0, y')$

Ind. (We show that $\alpha(x,y) \vdash \alpha(x', y')$: $(x' \cdot z = y' \cdot z \supset x' = y')$)
(1) $\alpha(x,y)$: $x \cdot z = y \cdot z \supset x = y$; hyp.
(2) $x' \cdot z = y' \cdot z$; hyp.
(3) $x \cdot z + z = y \cdot z + z$; (2), Th. I.13, $\text{Conseq}_{\text{PA1}}$
(4) $x \cdot z = y \cdot z$; (3), Th. I.19
(5) $x = y$; (1), (4), MP
(6) $x' = y'$; (5), $\text{Conseq}_{\text{PA2}}$
(7) $\alpha(x,y),\ x' \cdot z = y' \cdot z \vdash x' = y'$; (1)-(6)
(8) $\alpha(x,y) \vdash x' \cdot z = y' \cdot z \supset x' = y'$; (7), Ded. Th.
I.e., $\alpha(x,y) \vdash \alpha(x', y')$, i.e.,
$\vdash \alpha(x, y) \supset \alpha(x', y')$;

Therefore, by *Basis* and *Ind* follows $\forall x \alpha(x, y')$.

And, finally, by *Basis*: $\forall x \alpha(x,0)$, and

Ind: $\forall x \alpha(x, y) \supset \forall x \alpha(x, y')$ follows that $\forall y \forall x \alpha(x, y)$. This last formula is got by assuming that $z \neq 0$, i.e.,

$z \neq 0 \vdash \forall y \forall x \alpha(x, y)$

Therefore we have the following derivation:

(a) $\vdash z \neq 0 \supset \forall y \forall x \alpha(x, y)$,

(b) $\vdash \forall y \forall x (z \neq 0 \supset \alpha(x, y))$; since x and y do not appear free in the antecedent of (a).

And then from (b), by Gen, Ax4 and MP, 7 may be derived.

Definitions

Def 1. $t_1 < t_2 =_{df} \exists z(t_1 + z' = t_2)$, where t_1 and t_2 do not contain z

Def 2. $t_1 \leq t_2 =_{df} t_1 < t_2 \vee t_1 = t_2$

Def 3. $t_1 > t_2 =_{df} t_2 < t_1$

Def 4. $t_1 \geq t_2 =_{df} t_2 \leq t_1$

Def 5. $t_1 \not< t_2 =_{df} \neg(t_1 < t_2)$

Def 6. $t_1 \neq t_2 =_{df} \neg(t_1 = t_2)$

Theorem III. *For any terms* $t,\ t_1,\ t_2,\ t_3$ *the following are theorems of* PA.

1. $t \not< t$
2. $t_1 < t_2 \supset (t_2 < t_3 \supset t_1 < t_3)$
3. $t_1 < t_2 \supset t_2 \not< t_1$
4. $t_1 < t_2 \equiv t_1 + t_3 < t_2 + t_3$
5. $t_1 \leq t_2 \supset (t_2 < t_3 \supset t_1 < t_3)$
6. $t_1 = t_2 \supset (t_1 < t_3 \supset t_2 < t_3)$
7. $t_1 = t_2 \supset (t_3 < t_1 \supset t_3 < t_2)$
8. $t_1 \leq t_2 \supset (t_2 \leq t_3 \supset t_1 \leq t_3)$
9. $t_1 \leq t_2 \equiv (t_1 + t_3 \leq t_2 + t_3)$
10. $t \not< 0$
11. $t_1 = t_2 \vee t_1 < t_2 \vee t_2 < t_1$
12. $t \neq 0 \equiv t > 0$
13. $0 \leq t$
14. $0 < t'$
15. $t < t'$
16. $t \leq t$
17. $t_1 + t_2 \geq t_1$
18. $t_2 \neq 0 \supset t_1 + t_2 > t_1$
19. $t_1 \leq t_2 \vee t_2 \leq t_1$
20. $t_1 \leq t_2 \equiv t_1 < t'_2$
21. $t_1 < t_2 \equiv t'_1 \leq t_2$
22. (1) $0 < \overline{1}$, (2) $\overline{1} < \overline{2}, \ldots$
23. $t = 0 \vee \exists w(t = w')$
24. $t_2 > 0 \supset t_1 \cdot t_2 \geq t_1$

Proofs.

1. $t \nless t$ (Irreflexivity of $<$)
   (1) $t < t$; hyp
   (2) $\exists w'(t + w' = t)$; (1) Def. 1
   (3) $u = w'$; (2) C-Rule (comp. Ch. 2, Sect. 3.3, Choice Rule)
   (4) $u + t = t$; (2), (3)
   (5) $t = 0 + t$; Th. I.6
   (6) $u + t = 0 + t$; (4), (5), Th. I.3
   (7) $u = 0$; (6), Th. I.19
   (8) $w' = 0$; (3), (7), Conseq_{PA1} (contradicting Conseq_{PA3})

2. $t_1 < t_2 \supset (t_2 < t_3 \supset t_1 < t_3)$; (Trans. of $<$)
   (1) $t_1 < t_2$; hyp
   (2) $t_2 < t_3$; hyp
   (3) $\exists w(t_1 + w' = t_2)$; (1), Def 1
   (4) $\exists u(t_2 + u' = t_3)$; (2), Def 1
   (5) $t_1 + w' = t_2$; (3) C-Rule
   (6) $t_2 + u' = t_3$; C-Rule
   (7) $(t_1 + w') + u' = t_2 + u'$; (5), Th I.7
   (8) $(t_1 + w') + u' = t_3$; (7), (6), Th I.3
   (9) $t_1 + (w' + u') = t_3$; (8), Th I.10, Conseq_{PA1}
   (10) Let $v = w' + u$, then
   (11) $v' = w' + u'$; (10) by Conseq_{PA2}, Conseq_{PA6}
   (12) $t_1 + v' = t_1 + (w' + u')$; (11), Th I.9
   (13) $t_1 + v' = t_3$; (12), (9), Th I.3
   (14) $\exists v(t_1 + v' = t_3)$; (13), Gen. $\exists$
   (15) $t_1 < t_3$; (14), Def 1
   Hence $t_1 < t_2 ,\ t_2 < t_3 \vdash t_1 < t_3$; (1)-(15), and then
   $\vdash t_1 < t_2 \supset (t_2 < t_3 \supset t_1 < t_3)$; by Ded. Th.

3. $t_1 < t_2 \supset t_2 \nless t_1$
   (1) $t_1 < t_2 \supset (t_2 < t_1 \supset t_1 < t_1)$; 2
   (2) $t_1 < t_2 \supset (t_1 \nless t_1 \supset t_2 \nless t_1)$; (1) PL
   (3) $t_1 \nless t_1 \supset (t_1 < t_2 \supset t_2 \nless t_1)$; (2) PL

(4) $t_1 < t_2 \supset t_2 \not< t_1$; (3), 1, MP

4. $t_1 < t_2 \equiv t_1 + t_3 < t_2 + t_3$; for any terms t_1, t_2, t_3

4.1. $t_1 < t_2 \supset t_1 + t_3 < t_2 + t_3$

(1) $t_1 < t_2$; hyp

(2) $\exists w(t_1 + w' = t_2)$; (1) Def 1

(3) $t_1 + w' = t_2$; (2) C-Rule

(4) $(t_1 + w') + t_3 = t_2 + t_3$; (3), Th I.7

(5) $(t_1 + t_3) + w' = t_2 + t_3$; (4), by Th I.8, Th. I.10

(6) $\exists w((t_1 + t_3) + w' = t_2 + t_3)$; (5), Gen $\exists$

(7) $t_1 + t_3 < t_2 + t_3$; (6) Def 1

Hence

(8) $t_1 < t_2 \supset t_1 + t_3 < t_2 + t_3$; (1)-(7), Ded. Th.

4.2. $t_1 + t_3 < t_2 + t_3 \supset t_1 < t_2$

(1) $t_1 + t_3 < t_2 + t_3$; hyp

(2) $\exists w((t_1 + t_3) + w' = t_2 + t_3)$; (1) Def 1

(3) $(t_1 + t_3) + w' = t_2 + t_3$; (2) C-Rule

(4) $(t_1 + w') + t_3 = t_2 + t_3$; (3), by Th I.8, Th I.10

(5) $t_1 + w' = t_2$; (4), Th. I.19

(6) $\exists w(t_1 + w' = t_2)$; (5), Gen $\exists$

(7) $t_1 < t_2$; (6) Def 1

Hence

(8) $t_1 + t_3 < t_2 + t_3 \supset t_1 < t_2$; (1)-(7), Ded. Th.

4 follows from 4.1 and 4.2 by PL.

5. $t_1 \le t_2 \supset (t_2 < t_3 \supset t_1 < t_3)$

By Def 2 and PL: $((p \supset r) \wedge (q \supset r)) \supset ((p \vee q) \supset r)$, to prove 5 means to prove

5.1. $t_1 < t_2 \supset (t_2 < t_3 \supset t_1 < t_3)$,

already proved by 2 above, and

5.2. $t_1 = t_2 \supset (t_2 < t_3 \supset t_1 < t_3)$

(1) $t_1 = t_2$; hyp

(2) $t_2 < t_3$; hyp

(3) $t_1 = t_2 \supset t_1 + w' = t_2 + w'$; Th I.7

(4) $t_1 + w' = t_2 + w'$; (1), (3), MP

(5) $\exists w(t_2 + w' = t_3)$; (2), Def 1

(6) $t_2 + w' = t_3$; (5), C-Rule

(7) $t_1 + w' = t_3$; (4), (6), Th I.3

(8) $\exists w(t_1 + w' = t_3)$; (7), Gen $\exists$

(9) $t_1 < t_3$; (8), Def 1

5 follows by PL

6. $t_1 = t_2 \supset (t_1 < t_3 \supset t_2 < t_3)$ (exercise)

7. $t_1 = t_2 \supset (t_3 < t_1 \supset t_3 < t_2)$ (exercise)

8. $t_1 \leq t_2 \supset (t_2 \leq t_3 \supset t_1 \leq t_3)$; (exercise)

9. $t_1 \leq t_2 \equiv (t_1 + t_3 \leq t_2 + t_3)$

9.1. $t_1 \leq t_2 \supset (t_1 + t_3 \leq t_2 + t_3)$

(1) $t_1 < t_2 \supset (t_1 + t_3 < t_2 + t_3)$; 4.1

(2) $t_1 = t_2 \supset (t_1 + t_3 = t_2 + t_3)$; Th I.7

(3) $t_1 \leq t_2 \supset (t_1 + t_3 \leq t_2 + t_3)$; (1), (2), PL

$(p \supset q) \supset [(r \supset s) \supset ((p \vee r) \supset (q \vee s))]$, MP, Def 2

9.2. $(t_1 + t_3 \leq t_2 + t_3) \supset t_1 \leq t_2$

(1) $(t_1 + t_3 < t_2 + t_3) \supset t_1 < t_2$; 4.2

(2) $(t_1 + t_3 = t_2 + t_3) \supset t_1 = t_2$; Th I.19

(3) $(t_1 + t_3 \leq t_2 + t_3) \supset t_1 \leq t_2$; (2), (3), PL, MP, Def 2

10. $t \not< 0$

(1) $t < 0$; hyp

(2) $\exists z(t + z' = 0)$; (1) Def 1

(3) $t + z' = 0$; C-Rule

(4) $t + z' = (t + z)'$; Conseq_{PA6}

(5) $0 = (t + z)'$; (3), (4), Conseq_{PA1}

(6) $t < 0 \supset (0 = (t + z)')$; (1)-(5), Ded. Th.

(7) $0 \neq (t + z)'$; Conseq_{PA3}

(8) $t \not< 0$; (6), (7), PL

11. $t_1 = t_2 \vee t_1 < t_2 \vee t_2 < t_1$; (exercise)

12. $t \neq 0 \equiv t > 0$

(1) $t \neq 0$; hyp

(2) $t \not< 0$; 10

(3) $t \neq 0 \wedge t \not< 0$; (2), (3), PL

(4) $(t \neq 0 \wedge t \not< 0) \supset \neg(t = 0 \vee t < 0)$; PL

(5) $\neg(t = 0 \vee t < 0)$; (3), (4), MP

(6) $\neg(t = 0 \vee t < 0) \supset 0 < t$; 11, PL

(7) $0 < t$; (5), (6), MP

(8) $t > 0$; (7), Def 3

Hence

(9) $t \neq 0 \supset t > 0$; (1)-(8), Ded. Th.

Since $t > 0 \supset t \neq 0$, 12 follows by PL.

13. $0 \leq t$ (exercise)

14. $0 < t'$ (exercise)

15. $t < t'$ (exercise)

16. $t \leq t$, (exercise)

17. $t_1 + t_2 \geq t_1$; (exercise)

18. $t_2 \neq 0 \supset t_1 + t_2 > t_1$ (exercise)

19. $t_1 \leq t_2 \vee t_2 \leq t_1$; (exercise)

20. $t_1 \leq t_2 \equiv t_1 < t'_2$; (exercise)

21. $t_1 < t_2 \equiv t'_1 \leq t_2$; (exercise)

22. (1) $0 < \overline{1}$, (2) $\overline{1} < \overline{2}$,... (exercise)

23. $t = 0 \vee \exists w(t = w')$

(1) $t = 0 \vee t \neq 0$; PL

(2) $t \neq 0 \equiv t > 0$; 12

(3) $t = 0 \vee t > 0$; (1), (2), PL

(4) $t > 0$ is $0 < t$; Def 3

(5) $0 < t$ is $\exists w(0 + w' = t)$; Def 1

(6) $0 < t$ is $\exists w(t = w')$; (5), Th. I.8, Th. I.3, Conseq_{PA5}

(7) $t = 0 \vee \exists w(t = w')$; (3)-(6) PL

(A shorter proof can be given using 13. In a similar fashion can be proved: $t = 0 \vee t = \overline{1} \vee \exists w(t = w'')$ etc.; exercise).

24. $t_2 > 0 \supset t_1 \cdot t_2 \geq t_1$ (exercise).

Theorem IV. *For any natural numbers m, n the following holds*:

(1) *If* $m \neq n$, *then* $\vdash \overline{m} \neq \overline{n}$.

(2) (a) $\vdash \overline{m+n} = \overline{m} + \overline{n}$, *and* (b) $\vdash \overline{m \cdot n} = \overline{m} \cdot \overline{n}$.

Proof (1) (*Reductio*)

(a) Suppose that $m \neq n$; i.e., $m > n$ or $n > m$. Let us take $m > n$.

(b) $\overline{m} = \overline{n}$; hyp. (i.e., the terms $\overline{m}$ and $\overline{n}$ are identical)

(c) $\overline{m-n} = 0$; from (b) by *n*-applications of $\text{Conseq}_{\text{PA4}}$

(d) Since $m > n$ (by (a)), it follows that $m - n > 0$, and therefore $m - n - 1 \geq 0$ (by Th. III.20 and Def. 3 and Def. 4).

(e) Let $\mathrm{t} = \overline{m-n-1}$.

(f) $\mathrm{t}' = \overline{m-n} = 0$; (e) and (c)

(g) $\vdash \mathrm{t}' \neq 0$; by $\text{Conseq}_{\text{PA3}}$

(h) $\overline{m} = \overline{n} \vdash \mathrm{t}' = 0 \wedge \mathrm{t}' \neq 0$; (a)-(g)

(i) $\vdash \overline{m} = \overline{n} \supset (\mathrm{t}' = 0 \wedge \mathrm{t}' \neq 0)$; (h), Ded. Th.

(j) $\vdash \overline{m} \neq \overline{n}$; (i) PL

By (a) and (j) it follows (1).

The case $n > m$, similar.

Proof (2(a)) (exercise).

Proof (2(b)) (induction on *n* in the metalanguage)

Basis. (a) Since $m \cdot 0 = 0$ it follows that the terms $\overline{m \cdot 0}$ and 0 are identical, and then $\vdash \overline{m \cdot 0} = 0$ (by Th. I,1).

(b) $\vdash \overline{m} \cdot 0 = 0$; $\text{Conseq}_{\text{PA7}}$

(c) $\vdash \overline{m \cdot 0} = \overline{m} \cdot 0$; (a), (b), Th. I,4.

Ind. (a) $\overline{m \cdot n} = \overline{m} \cdot \overline{n}$; hyp.

(b) Since $m \cdot n' = m \cdot n + m$, it follows that $\overline{m \cdot n'} = \overline{m \cdot n + m}$.

(c) $\overline{m \cdot n + m} = \overline{m \cdot n} + \overline{m}$; by 2(a)

(d) $\overline{m \cdot n} + \overline{m} = \overline{m} \cdot \overline{n} + \overline{m}$; (a), Th. I,7

(e) $\overline{m} \cdot \overline{n} + \overline{m} = \overline{m} \cdot \overline{(n')}$; (b)-(d)

(f) $\overline{m \cdot (n')} = \overline{m} \cdot \overline{(n')}$; (b)-(e)

Theorem V.

1. *For any natural number k*,

$$\vdash (\mathrm{x} = 0 \vee \ldots \vee \mathrm{x} = \overline{k}) \equiv \mathrm{x} \leq \overline{k}$$

1*. *For any natural number k and any formula* $\alpha(\mathrm{x})$,

$\vdash (\alpha(0) \wedge \alpha(\bar{1}) \wedge ... \wedge \alpha(\bar{k})) \equiv \forall x(x \leq \bar{k} \supset \alpha(x))$

2. *For any natural number* $k > 0$,

$\vdash (x = 0 \vee ... \vee x = \overline{k-1}) \equiv x < \bar{k}$

2*. *For any natural number* $k > 0$ *and any formula* $\alpha(x)$,

$\vdash (\alpha(0) \wedge \alpha(\bar{1}) \wedge ... \wedge \alpha(\overline{k-1})) \equiv \forall x(x < \bar{k} \supset \alpha(x))$

Proof. 1 (induction on *k*)

Basis. $k = 0$. $x = 0 \equiv x \leq 0$; by Def. 2, Th. III.10, PL

Induction

(1) $\vdash (x = 0 \vee ... \vee x = \bar{k}) \equiv x \leq \bar{k}$; hyp

(2) $\vdash (x = 0 \vee ... \vee x = \bar{k}) \supset x \leq \bar{k}$; (1) PL

(3) $\vdash x = \overline{k+1} \supset x \leq \overline{k+1}$; Def. 2 and PL

(4) $\vdash x \leq \bar{k} \supset x \leq \overline{k+1}$

(Since $x \leq \bar{k}$ is $x < \bar{k} \vee x = \bar{k}$ (by Def);

but $\vdash x < \bar{k} \supset x < \overline{k+1}$ and $\vdash x = \bar{k} \supset x < \overline{k+1}$.

Hence $\vdash x \leq \bar{k} \supset x \leq \overline{k+1}$; by PL

(5) $\vdash (x = 0 \vee ... \vee x = \bar{k}) \supset x \leq \overline{k+1}$; (2), (4), PL

(6) $\vdash x = \overline{k+1} \supset x \leq \overline{k+1}$; Def. 2, PL

(7) $\vdash (x = 0 \vee ... \vee x = \bar{k} \vee x = \overline{k+1}) \supset x \leq \overline{k+1}$; (5), (6), PL
by $\models (p \supset r) \supset [(q \supset r) \supset ((p \vee q) \supset r)]$
and Rule_p (Ch. 2, 3.2.1)

(8) $\vdash x \leq \bar{k} \supset (x = 0 \vee ... \vee x = \bar{k})$; (1), PL

(9) $\vdash x = \overline{k+1} \supset (x = 0 \vee ... \vee x = \overline{k+1})$; PL

(10) $\vdash (x \leq \bar{k} \vee x = \overline{k+1}) \supset (x = 0 \vee ... \vee x = \bar{k} \vee x = \overline{k+1})$; (8), (9), PL
$\models (p \supset q) \supset [(r \supset s) \supset ((p \vee r) \supset (q \vee s))]$ and Rule_p (Ch. 2, 3.2.1)

(11) $\vdash x \leq \overline{k+1} \supset (x = 0 \vee ... \vee x = \bar{k} \vee x = \overline{k+1})$; (10)

(12) $\vdash (x = 0 \vee ... \vee x = \bar{k} \vee x = \overline{k+1}) \equiv x \leq \overline{k+1}$; (7), (11), PL

1*. For any natural number *k* and any formula $\alpha(x)$,

$\vdash (\alpha(0) \wedge \alpha(\bar{1}) \wedge ... \wedge \alpha(\bar{k})) \equiv \forall x(x \leq \bar{k} \supset \alpha(x))$

Proof. 1*

As we know,[6] for any k the following holds:

$\vdash \alpha(0) \equiv \forall x(x = 0 \supset \alpha(x)), \ldots, \vdash \alpha(\bar{k}) \equiv \forall x(x = \bar{k} \supset \alpha(x))$

And then

$\vdash (\alpha(0) \wedge \ldots \wedge \alpha(\bar{k})) \equiv (\forall x(x = 0 \supset \alpha(x)) \wedge \ldots \wedge \forall x(x = \bar{k} \supset \alpha(x)))$

$\equiv \forall x(x = 0 \supset \alpha(x)) \wedge \ldots \wedge (x = \bar{k} \supset \alpha(x))$; by $(\forall x\alpha \wedge \forall x\beta) \equiv \forall x(\alpha \wedge \beta)$ (cf. Ch. 2, Sect. 3.3, Th.13), PL

$\equiv \forall x((x = 0 \vee \ldots \vee x = \bar{k}) \supset \alpha(x)) \equiv \forall x(x \leq \bar{k} \supset \alpha(x))$; by 1, FOL^{ax}

Proof. 2

(1) $(x = 0 \vee x = \bar{1} \vee \ldots \vee x = \overline{k-1}) \equiv x \leq \overline{k-1}$; by 1

(2) $x \leq \overline{k-1} \equiv x < \bar{k}$; Th. III.20

(3) $(x = 0 \vee x = \bar{1} \vee \ldots \vee x = \overline{k-1}) \equiv x < \bar{k}$; (1), (2), PL

Proof. 2* (By 1*, 2).

Theorem VI.

(a) $\vdash y \neq 0 \supset \exists u \exists v(x = y \cdot u + v \wedge v < y)$
(existence of quotient and remainder)

(b) $\vdash [x = y \cdot u_1 + v_1 \wedge v_1 < y) \wedge (x = y \cdot u_2 + v_2 \wedge v_2 < y)] \supset$
$\supset (u_1 = u_2 \wedge v_1 = v_2)$ (uniqueness of quotient and remainder)

Proof.[7]

2. Number-theoretic functions and relations [8]

Definition 1. *A number-theoretic function is a function whose arguments and values are natural numbers.*

The successor function, for example, is a unary number-theoretic

---

[6] An equivalence of the form $\alpha(\bar{n}) \equiv \forall x(x = \bar{n} \supset \alpha(x))$ is a theorem of FOL^{ax}_{id} comp. Ch. 2, 4.3 Lemma.

[7] For the (long) proof of this theorem comp. *inter alia* S.C. Kleene [1952], Ch. VIII, §38 and E. Mendelson [1964], Ch. 3, Sect. 1.

[8] The content of Sections 2 and 3 are general facts on the number-theoretic functions and relations, and recursivity. For the details of these matters we refer the reader to the textbooks on this topic, e.g., S.C. Kleene [1952], Chs. VIII and IX; Péter, R. [1951]; D. Hilbert, P. Bernays [1934] §7; E. Mendelson [1964], Ch. 3; G.S. Boolos, J.P. Burgess, R.C. Jeffrey [2002], Chs. 6, 7, 15, 16; also used for the present considerations.

function, and addition and multiplication are two binary number-theoretic functions.

Definition 2. *A number-theoretic relation is a relation whose arguments are natural numbers.*

For example, $x = y$ and $x < y$ are two-place number-theoretic relations, and $x + y = z$ is a three-place number-theoretic relation.

Notational distinction. The considerations of this section are essentially *intuitive*. But since some concepts, like the expressibility of a relation in PA^{ax} or the representability of a function in PA^{ax}, do imply the *formal* level, in order to keep the distinction *intuitive-formal* we want to introduce a new symbolism for the intuitive language.[9] Namely, the intuitive symbolism consists of the following symbols: "$\sim$" – negation, "$\wedge$" – conjunction, "$\vee$" – disjunction,[10] "$\rightarrow$" – implication, "$\leftrightarrow$" – equivalence, x_1, x_2, x_3,..., x, y, z,... numerical variables, f_1, f_2, f_3,..., f, g, h,... number-theoretic functions, P_1, P_2, P_3,..., P, Q, R,... intuitive symbols for predicates, "(x)" – universal quantifier, "(Ex)" – existential quantifier, "(E_1x)" – there exists a unique x", "$(x)_{x<y}$" and "$(Ex)_{x<y}$" – bounded quantifiers. The corresponding formal symbolism is that used in the preceding chapters and in the section 1 of this chapter, i.e.,: $\neg$, $\wedge$, $\vee$, $\supset$, $\equiv$, x_1, x_2, x_3,..., f_1, f_2, f_3,..., P_1, P_2, P_3,..., $\forall\mathrm{x}$, $\exists\mathrm{x}$, $\exists_1\mathrm{x}$, $\forall\mathrm{x}(\mathrm{x}<\mathrm{y}...)$, $\exists\mathrm{x}(\mathrm{x}<\mathrm{y}...)$.

Definition 3. *A number-theoretic relation $R(x_1,...,x_n)$ is formally expressible in* PA^{ax} *if and only if there is a formula* $\alpha(\mathrm{x}_1,...,\mathrm{x}_n)$ *of* L_{PA} *with no free variables other than the distinct variables* $\mathrm{x}_1,...,\mathrm{x}_n$ *such that for any numbers $k_1,...,k_n$ holds*:

1. *If $R(k_1,...,k_n)$ is true, then $\vdash \alpha(\bar{k}_1,...,\bar{k}_n)$.*
2. *If $R(k_1,...,k_n)$ is false, then $\vdash \neg\alpha(\bar{k}_1,...,\bar{k}_n)$.*

Example. The relation "less than" is expressed in PA^{ax} by the formula $\mathrm{x}<\mathrm{y}$. Since if $k_1 < k_2$, then there is some n such that $k_2 = k_1 + n'$. And then, $\bar{k}_2$ and $\bar{k}_1 + \bar{n}'$ (*via* Sect. 1.4, Th. IV (2) (a)) are one and the same term. Hence,

---

[9] This distinction was firstly introduced by Hilbert and Ackermann [1928], and then by Gödel [1931], Gentzen [1934], Kleene [1952], Mendelson [1964] and others. With some modifications the present distinction is that of Kleene's [1952], § 45.

[10] "$\wedge$" and "$\vee$" are common to intuitive and to formal symbolism; but the context of their use will show everytime which is the intended meaning.

by Sect. 1.4, Th. I,1, $\vdash \bar{k}_2 = \bar{k}_1 + \bar{n}'$ and therefore $\vdash \exists z(\bar{k}_2 = \bar{k}_1 + \mathrm{z}')$, by Gen∃, i.e., $\vdash \bar{k}_1 < \bar{k}_2$, by Sect. 4.1, Def 1 (preceding Th. III). If, on the other hand, $k_1 \not< k_2$, then $k_1 = k_2$ or $k_2 < k_1$. If $k_1 = k_2$, then $\bar{k}_1$ and $\bar{k}_2$ are one and the same term, and by Sect. 1.4, Th I.1, $\vdash \bar{k}_1 = \bar{k}_2$. And, finally, if $k_2 < k_1$, then, as above, $\vdash \bar{k}_2 < \bar{k}_1$, and by PL, $\vdash \bar{k}_2 \leq \bar{k}_1$. So, for both cases, $\vdash k_2 \leq k_1$. Hence, by Sect. 1.4, Th III.11, it follows $\vdash \bar{k}_1 \not< \bar{k}_2$.

Definition 4. *A number-theoretic function* $f(x_1,...,x_n)$ *is formally representable in* $\mathrm{PA}^{\mathrm{ax}}$ *if and only if there is a formula* $\alpha(\mathrm{x}_1,...,\mathrm{x}_n,\mathrm{x}_{n+1})$ *of* $\mathrm{L_{PA}}$ *with no free variables other than the distinct variables* $\mathrm{x}_1,...,\mathrm{x}_n,\mathrm{x}_{n+1}$ *such that for any numbers* $k_1,...,k_n,k_{n+1}$ *hold*:

1. *If* $f(k_1,...,k_n) = k_{n+1}$, *then* $\vdash \alpha(\bar{k}_1,...,\bar{k}_\mathrm{n},\bar{k}_{\mathrm{n+1}})$.
2. $\exists_1 \mathrm{x}_{\mathrm{n+1}}\alpha(\bar{k}_1,...,\bar{k}_\mathrm{n},\mathrm{x}_{\mathrm{n+1}})$.

The notation $\exists_1\mathrm{x}\alpha(\mathrm{x})$ is an abbreviation for: *there exists a unique* x *such that* $\alpha(\mathrm{x})$. It is equivalent to the following formulas:

$\exists\mathrm{x}(\alpha(\mathrm{x}) \wedge \forall\mathrm{y}(\alpha(\mathrm{y}) \supset \mathrm{x} = \mathrm{y}))$ and

$\exists\mathrm{x}\alpha(\mathrm{x}) \wedge \forall\mathrm{x}\forall\mathrm{y}[(\alpha(\mathrm{x}) \wedge \alpha(\mathrm{y})) \supset \mathrm{x} = \mathrm{y}]$,

respectively, where $\exists\mathrm{x}\alpha(\mathrm{x})$ expresses *existence* and $\forall\mathrm{x}\forall\mathrm{y}[(\alpha(\mathrm{x}) \wedge \alpha(\mathrm{y})) \supset \mathrm{x} = \mathrm{y}]$ expresses *uniqueness*.

The function $f(x_1,...,x_n)$ is *strongly* representable in $\mathrm{PA}^{\mathrm{ax}}$ if 2 is replaced by 2^* $\exists_1\mathrm{x}_{\mathrm{n+1}}\alpha(\mathrm{x}_1,...,\mathrm{x}_\mathrm{n},\mathrm{x}_{\mathrm{n+1}})$. The two notions of representability are provably equivalent.

Example. The successor function, $\mathit{Succ}(\mathrm{x}) = \mathrm{x} + 1$[11] is strongly representable in $\mathrm{PA}^{\mathrm{ax}}$ by the formula $\mathrm{y} = \mathrm{x}'$. Since for any k_1 if $\mathit{Succ}(\mathrm{k}_1) = \mathrm{k}_2$, then $k_2 = k_1 + 1$, i.e., $\bar{k}_2 = (\bar{k}_1)'$, and therefore $\vdash \bar{k}_2 = (\bar{k}_1)'$ and $\vdash \exists_1\mathrm{y}(\mathrm{y} = (\mathrm{x})')$.

Similarly, $x + y$ and $x \cdot y$ will be represented in $\mathrm{PA}^{\mathrm{ax}}$ by the formulas $\mathrm{x} + \mathrm{y} = \mathrm{z}$ and $\mathrm{x} \cdot \mathrm{y} = \mathrm{z}$, respectively (show that!).

Let us remark the fact that Def. 4 is equivalent to the following definition of formal representability of a function in $\mathrm{PA}^{\mathrm{ax}}$. (Let us take in Def.4 n =1).

[11] Or x' = x + 1.

Definition 4*. *A number-theoretic function $f(x)$ is formally representable in* PA^{ax} *if and only if there is a formula* $\alpha(x,y)$ *of* L_{PA} *with no free variables other than the distinct variables* x *and* y, *such that for any numbers* k_1, k_2 *holds*:

If $f(k_1) = k_2$, *then* $PA^{ax} \vdash \forall y(\alpha(\bar{k}_1, y) \equiv y = \bar{k}_2)$.[12]

Now, the formula $\forall y(\alpha(\bar{k}_1, y) \equiv y = \bar{k}_2)$ is logically equivalent to $\forall y[(\alpha(\bar{k}_1, y) \supset y = \bar{k}_2) \wedge (y = \bar{k}_2 \supset \alpha(\bar{k}_1, y))]$ (by PL), equivalently $\forall y(\alpha(\bar{k}_1, y) \supset y = \bar{k}_2) \wedge \forall y(y = \bar{k}_2 \supset \alpha(\bar{k}_1, y))$ (by Ch. 2, Sect. 3.3, Th. 13). But $\forall y(y = \bar{k}_2 \supset \alpha(\bar{k}_1, y)) \equiv \alpha(\bar{k}_1, \bar{k}_2)$ (by Ch. 2, Sect. 4.3, Lemma), and $\forall y(\alpha(\bar{k}_1, y) \supset y = \bar{k}_2) \equiv \forall y(y \neq \bar{k}_2 \supset \neg\alpha(\bar{k}_1, y))$ (by FOL). And therefore the formula $\forall y(\alpha(\bar{k}_1, y) \equiv y = \bar{k}_2)$ (from Def. 4*) is logically equivalent to the conjunction of the following assertions:

(1) $\alpha(\bar{k}_1, \bar{k}_2)$

(2) $\forall y(y \neq \bar{k}_2 \supset \neg\alpha(\bar{k}_1, y))$.

Definition 5. *The characteristic function of a number-theoretic relation* $R(x_1,...,x_n)$ *is the function* $C_R(x_1,...,x_n)$ *defined as follows:*

$$C_R(x_1,...,x_n) = \begin{cases} 0, & \text{if} \quad R(x_1,...,x_n) \quad \text{is true} \\ 1, & \text{if} \quad R(x_1,...,x_n) \quad \text{is false,} \end{cases}$$

or equivalently, $C_R(x_1,...,x_n)$ *is the function taking only* 0 *and* 1 *as values and which satisfies the following equivalence:*

$$R(x_1,...,x_n) \leftrightarrow C_R(x_1,...,x_n) = 0.$$

Theorem. *A number-theoretic relation* $R(x_1,...,x_n)$ *is expressible in* PA^{ax} *if and only if* $C_R(x_1,...,x_n)$ *is representable in* PA^{ax}.

Proof. Suppose that $R(x_1,...,x_n)$ is expressible in PA^{ax} by the formula $\alpha(x_1,...,x_n)$. Let $C_R(x_1,...,x_n)$ be its characteristic function. Let $C_R(x_1,...,x_n) = x_{n+1}$, where x_{n+1} is either 0 or 1. Now, if $x_{n+1} = 0$, then $C_R(x_1,...,x_n) = 0$, and then $R(x_1,...,x_n)$ holds; whence $PA^{ax} \vdash \alpha(\bar{x}_1,...,\bar{x}_n)$ (by Def 3). And since $x_{n+1} = 0$, then $PA^{ax} \vdash \bar{x}_{n+1} = 0$ (by Sect. 1.4, Theorem I,1). And, therefore, by PL

[12] Def 4* is more appropriate to prove some mathematical facts, e.g. *Diagonal Lemma* and its applications in studying of Gödel's Theorem (comp. Sect. 4.2.2).

(1) $\mathrm{PA}^{\mathrm{ax}} \vdash \alpha(\bar{x}_1,\ldots,\bar{x}_n) \wedge \bar{x}_{n+1} = 0$.

On the other hand, if $x_{n+1} = 1$, then $C_R(x_1,\ldots,x_n) = 1$, and then $R(x_1,\ldots,x_n)$ does bot hold; whence $\mathrm{PA}^{\mathrm{ax}} \vdash \neg\alpha(\bar{x}_1,\ldots,\bar{x}_n)$ (by Def 3), and $\mathrm{PA}^{\mathrm{ax}} \vdash \bar{x}_{n+1} = \bar{1}$, i.e.,

(2) $\mathrm{PA}^{\mathrm{ax}} \vdash \neg\alpha(\bar{x}_1,\ldots,\bar{x}_n) \wedge \bar{x}_{n+1} = \bar{1}$.

As (1) and (2) shows, the formula representing $C_R(x_1,\ldots,x_n)$ in $\mathrm{PA}^{\mathrm{ax}}$ is $\beta(\mathrm{x}_1,\ldots,\mathrm{x}_n,\mathrm{x}_{n+1}) : (\alpha(\mathrm{x}_1,\ldots,\mathrm{x}_n) \wedge \mathrm{x}_{n+1} = 0) \vee (\neg\alpha(\mathrm{x}_1,\ldots,\mathrm{x}_n) \wedge \mathrm{x}_{n+1} = \bar{1})$.

Finally, for the converse, suppose that $C_R(x_1,\ldots,x_n)$ is representable in $\mathrm{PA}^{\mathrm{ax}}$ by the formula $\gamma(x_1,\ldots,x_n,x_{n+1})$. As evident, $R(x_1,\ldots,x_n)$ will be expressible in $\mathrm{PA}^{\mathrm{ax}}$ by the formula $\gamma(x_1,\ldots,x_n,0)$.

3. Recursive functions and relations

The *recursive definition* of a function $f(y)$ (or the *definition by induction*) is based on the following ideas: the value $f(0)$ (i.e., the value of f for 0 as the argument) is given, and then for any *y*, the value of f for y', i.e., $f(y')$, is given in terms of *y* and $f(y)$.[13] To illustrate this fact, let us consider the following equations

$$\begin{cases} f(0) = k \\ f(y') = h(y, f(y)), \end{cases}$$

defining a unary function $f(y)$ by induction on *y*, where *k* is a given natural number and $h(y,z)$ is a given binary number-theoretic function.

The two equations enable us to determine the value of $f(y)$ for any *y*. For, by the first equation, $f(0) = k$, with *k* given. Then by the second equation $f(1)$ is the value of $h(0, f(0))$, i.e., $h(0,k)$, a number specifiable by the fact that $h(y,z)$ is a given function. The next value, $f(2)$, by the second equation, will be $h(1, f(1))$, i.e., $h(1, h(0,k))$, again a given number, and so on.

The above equations are functional equations in an unknown function f. Sometimes in the definition occur additional variables $x_1,\ldots,x_n$, called *parameters*, which have fixed values in the process of induction on *y*.

[13] In a similar fashion we speak of a recursive definition of a predicate *P*(*y*).

Such is the case, for example, in the definition of addition, $x+y$, by induction on y with x as parameter and the successor function as a given function:

$$\begin{cases} x+0=x \\ x+y'=(x+y)'. \end{cases}$$

Formally, they are PA5 and PA6, as proper axioms of PA (cf. Sect. 1.2). The same is the case with the equations PA7 and PA8, defining multiplication.

A question arises: which are the number-theoretic functions definable recursively or by induction? The answer depends on what functions we take as known (or given) initially and on what operations (rules) are accepted for obtaining new functions from the initial functions. The class of recursive functions contain the class of primitive recursive functions and the class of those functions obtained by applications of the μ-operator.

3.1. Primitive recursive and recursive functions

The class of *primitive recursive functions* contains all the number-theoretic functions defined by the following equations and systems of equations (I-V):

I. $f(x)=x'$; the successor function

II. $f(x_1,...,x_n)=k$; where *k* is a given natural number; the constant function

III. $f_i^n(x_1,...,x_n)=x_i$; $1\le i\le n$; the projection functions

IV. $f(x_1,...,x_n)=g(h_1(x_1,...,x_n),...,h_m(x_1,...,x_n))$, where $g,h_1,...,h_m$ are given and *n, m* are positive integers; *Substitution*.

V. $\begin{cases} f(x_1,...,x_n,0)=g(x_1,...,x_n) \\ f(x_1,...,x_n,y')=h(x_1,...,x_n,y,f(x_1,...,x_n,y)), \end{cases}$

where *g, h* are given, *n* is a positive integer; *Recursion*

For $n=0$, V will be

$$\begin{cases} f(0)=k, \text{ where } k \text{ is a given natural number} \\ f(y')=h(y,f(y)). \end{cases}$$

I-III are called *initial* functions.

Definition 1. *A function f is primitive recursive if and only if there is a finite sequence* $f_1,...,f_n$ $(n \geq 1)$ *of functions* (*called primitive recursive description of* f), *such that* $f_n = f$ *and for* $1 \leq i \leq n$, *either* f_i *is an initial function or* f_i *comes from preceding functions of the sequence by an application of Substitution or Recursion*.

VI. μ-Operator. Let $g(x_1,...,x_n,y)$ be a function such that for any $x_1,...,x_n$ there is at least one y such that $g(x_1,...,x_n,y)=0$. The least natural number y such that this equation holds will be denoted by $\mu y(g(x_1,...,x_n,y)=0)$. Now, let $f(x_1,...,x_n)=\mu y(g(x_1,...,x_n,y)=0)$. If $g(x_1,...,x_n,y)$ obeys the specified condition, then we say that f is obtained from g by means of the μ-Operator.

Definition 2. *A function* f *is recursive if and only if it can be obtained from the initial functions by a finite number of applications of Substitution, Recursion and* μ-*Operator*.

Theorem 1. *The following functions are primitive recursive:*

1. $x+y$; $\begin{cases} x+0=x \\ x+y'=(x+y)' \end{cases}$ i.e., $\begin{cases} f(x,0)=f_1^1(x) \\ f(x,y')=(f(x,y))' \end{cases}$

2. $x \cdot y$; $\begin{cases} x \cdot 0 = 0 \\ x \cdot y' = x \cdot y + x \end{cases}$ i.e., $\begin{cases} g(x,0)=0 \\ g(x,y')=f(x,g(x,y)) \end{cases}$

   (where f is the addition)

3. x^y; $\begin{cases} x^0 = 1 \\ x^{y'} = x^y \cdot y \end{cases}$

4. $x!$; $\begin{cases} 0! = 1 \\ x'! = x! \cdot x' \end{cases}$

5. $\delta(x)=\begin{cases} x-1 & \text{if} \quad x>0 \\ 0 & \text{if} \quad x=0 \end{cases}$; $\begin{cases} \delta(0)=0 \\ \delta(x')=x \end{cases}$

6. $x \dot{-} y=\begin{cases} x-y & \text{if} \quad x \geq y \\ 0 & \text{if} \quad x<y \end{cases}$; $\begin{cases} x \dot{-} 0 = 0 \\ x \dot{-} y' = \delta(x \dot{-} y) \end{cases}$

7. $|x-y|=\begin{cases} x-y & \text{if} \quad x \geq y \\ y-x & \text{if} \quad x<y \end{cases}$; $|x-y|=(x \dot{-} y)+(y \dot{-} x)$

8. $sg(x)=\begin{cases}0 & \text{if} \quad x=0\\ 1 & \text{if} \quad x>0\end{cases}$; $\qquad \begin{cases}sg(0)=0\\ sg(x')=1\end{cases}$

9. $\overline{sg}(x)=\begin{cases}1 & \text{if} \quad x=0\\ 0 & \text{if} \quad x\neq 0\end{cases}$; $\quad \begin{cases}\overline{sg}(x)=1\dot{-}sg(x)\\ \text{or } \overline{sg}(x)=0^x\end{cases}$ or $\begin{cases}\overline{sg}(0)=1\\ \overline{sg}(x')=0\end{cases}$

10. $min(x,y)$ (minimum of *x* and *y*); $\qquad \min(x,y)=y\dot{-}(y\dot{-}x)$

$\min(x_1,...,x_n)$; If we assume $\min(x_1,...,x_n)$ known as primitive recursive, then

$\min(x_1,...,x_n,x_{n+1})=\min(\min(x_1,...,x_n),x_{n+1})$.

11. $max(x,y)$ (maximum of *x* and *y*); $\qquad \max(x,y)=y+(x\dot{-}y)$

or $\max(x,y)=(x+y)\dot{-}\min(x,y)$

$\max(x_1,...,x_n)$; $\qquad \max(x_1,...,x_n)=\max(\max(x_1,...,x_n),x_{n+1})$

12. $rm(x,y)$ (remainder of division of *x* by *y*);

$$\begin{cases}rm(0,y)=0\\ rm(x',y)=(rm(x,y))'\cdot sg(|y-(rm(x,y))'|)\end{cases}$$

13. $qt(x,y)$ (quotient of division of *x* by *y*)

$$\begin{cases}qt(0,y)=0\\ qt(x',y)=qt(x,y)+\overline{sg}(|\,y-(rm(x,y))'|)\end{cases}$$

Bounded sums and products

Let us now consider some functions, given by the respective definitions.

Definition 1.

$$\sum_{y<z} f(x_1,...,x_n,y)=\begin{cases}0 \text{ if } \; z=0\\ f(x_1,...,x_n,0)+...+f(x_1,...,x_n,z-1) \text{ if } \; z>0\end{cases}$$

(i.e., the *sum* of the numbers $f(x_1,...,x_n,y)$ such that $y<z$, if $z>0$, and 0 if $z=0$).

Definition 2.

$$\prod_{y<z} f(x_1,...,x_n,y)=\begin{cases}1 \text{ if } \; z=0\\ f(x_1,...,x_n,0)\cdot...\cdot f(x_1,...,x_n,z-1) \text{ if } \; z>0\end{cases}$$

(i.e., the *product* of the numbers $f(x_1,...,x_n,y)$ such that $y<z$, if $z>0$, and 1 if $z=0$).

These *bounded* (*finite*) sums and products are functions of $x_1,...,x_n,z$. Some other bounded sums and products can be defined in the way already specified by the definitions above, e.g.,

$$\sum_{y\leq z} f(x_1,...,x_n,y) = \sum_{y<z'} f(x_1,...,x_n,y)$$

$$\sum_{v<y<z} f(x_1,...,x_n,y) = \sum_{y<z\dot{-}v'} f(x_1,...,x_n,y+v')$$

(similarly for $\prod_{y\leq z}$ and $\prod_{v<y<z}$).

Theorem. *If* $f(x_1,...,x_n,y)$ *is (primitive) recursive*[14], *then all these bounded sums and products are (primitive) recursive.*

Proof. As can be seen, the function $\sum_{y<z} f(x_1,...,x_n,y)$ may be given from $f(x_1,...,x_n,y)$ by the following recursion on *z*:

$$\begin{cases} \sum_{y<0} f(x_1,...,x_n,y) = 0 \\ \\ \sum_{y<z'} f(x_1,...,x_n,y) = f(x_1,...,x_n,z) + \sum_{y<z} f(x_1,...,x_n,y). \end{cases}$$

For the function $\sum_{y\leq z} f(x_1,...,x_n,y)$ we can proceed as in the preceding case, or we can argue as follows. Since $\sum_{y<z} f(x_1,...,x_n,y)$ is primitive recursive, let it be $g(x_1,...,x_n,z)$, then $\sum_{y\leq z} f(x_1,...,x_n,y)$ is also primitive recursive, since it is $g(x_1,...,x_n,z')$; that is it can be obtained by *Substitution* from $g(x_1,...,x_n,z)$.

For the other functions, we proceed similarly (exercise).

Example. Let $Div(x)$ be the number of divisors of *x*, if $x>0$, and if $x=0$ let $Div(0)=1$. This function is primitive recursive, since

$$Div(\mathrm{x}) = \sum_{y\leq x} \overline{sg}(rm(x,y)).$$

[14] "(primitive) recursive" means everywhere *primitive recursive or recursive.*

3.2. Primitive recursive and recursive relations

As we saw, a number-theoretic relation is a relation whose arguments are natural numbers. If $R_1(x_1,...,x_n)$ and $R_2(x_1,...,x_n)$ are two such relations, then we can speak of their negation, conjunction, disjunction, implication or equivalence:

$\widetilde{R}_1(x_1,...,x_n)$, $R_1(x_1,...,x_n) \wedge R_2(x_1,...,x_n)$, $R_1(x_1,...,x_n) \vee R_2(x_1,...,x_n)$, $R_1(x_1,...,x_n) \rightarrow R_2(x_1,...,x_n)$, $R_1(x_1,...,x_n) \leftrightarrow R_2(x_1,...,x_n)$, respectively.

The connectives occuring in these new relations have their usual meanings; $R_1(x_1,...,x_n) \wedge R_2(x_1,...,x_n)$, for example, holds for $x_1,...,x_n$ iff $R_1(x_1,...,x_n)$ holds and $R_2(x_1,...,x_n)$ holds. If $R(x_1,...,x_n,y)$ is a number-theoretic relation, then by $(y)_{y<z}R(x_1,...,x_n,y)$ we mean the relation: for all *y*, if *y* is less than *z*, then $R(x_1,...,x_n,y)$ holds. The quantifier $(y)_{y<z}$ is called *bounded quantifier*. In a similar fashion we can use the other bounded quantifiers, $(y)_{y\leq z}$, $(Ey)_{y<z}$, $(Ey)_{y\leq z}$, respectively.

Definition. *A relation* $R(x_1,...,x_n)$ *is* (*primitive*) *recursive if and only if its characteristic function* $C_R(x_1,...,x_n)$ *is* (*primitive*) *recursive*.

Examples

(a) The identity (equality) relation $x_1 = x_2$ is primitive recursive, since its characteristic function, $sg(|x_1 - x_2|)$, is primitive recursive.

(b) The relation $x < y$ is primitive recursive, since its characteristic function $C_<(x,y) = \overline{sg}(y \dot{-} x)$, which is primitive recursive.

(c) The relation $Pr(x)$, *x* is a prime number, is primitive recursive, since its characteristic function is
$C_{Pr}(x) = sg((Div(x) \dot{-} 2) + \overline{sg}(|x-1|) + \overline{sg}(|x-0|))$.

(d) The relation $x \mid y$ (*x* divides *y*, or *y* is divisible by *x*) is primitive recursive, since its characteristic function is $C_{x|y}(x,y) = sg(rm(y,x))$.

Notation. If *R* is a relation, then the expressions of the form $(Ey)_{y<z}R$ and $(y)_{y<z}R$ of the *intuitive* language (containing bounded quantifiers) are simple abbreviations for (Ey) $(y < z \wedge R)$ and (y) $(y < z \supset R)$, respectively. Similarly, for $(Ey)_{y\leq z}R$ and $(y)_{y\leq z}R$. And then, in what follows, they will be used interchangeably.

Theorem 1. (1) *Let* $R_1(x_1,...,x_n)$ *and* $R_2(x_1,...,x_n)$ *be two n-place* (*primitive*) *recursive relations. Then* $\widetilde{R}_1$, $R_1 \circ R_2$ *are* (*primitive*) *recursive, where* "$\circ$" *is any binary connective of propositional logic.*

(2) *Let* $R(x_1,...,x_n,y)$ *be a* $n+1$ *place* (*primitive*) *recursive relation.*

Then (a) $(Ey)_{y<z}R$, $(Ey)_{y\leq z}R$, $(y)_{y<z}R$ *and* $(y)_{y\leq z}R$ *are* (*primitive*) *recursive relations.* (b) $\mu y_{y<z}R$ *and* $\mu y_{y\leq z}R$ *are* (*primitive*) *recursive functions.*

Proofs. (1) Since $R_1(x_1,...,x_n)$ is (primitive) recursive, its characteristic function $C_R(x_1,...,x_n)$ is (primitive) recursive. Then the characteristic function on $\widetilde{R}$ will be $C_{\widetilde{R}}(x_1,...,x_n)=1-C_R(x_1,...,x_n)$, which is (primitive) recursive. The characteristic function of the relation $R_1 \vee R_2$, for example, will be $C_{R_1}\cdot C_{R_2}$, so it is (primitive) recursive.

Now, since all the binary connectives of PL can be defined using ~ and $\vee$ it follows that all relations $R_1 \circ R_2$ are (primitive) recursive.

Examples

$$C_{R_1\wedge R_2}=C_{\sim(\widetilde{R}_1\vee\widetilde{R}_2)}=1-(C_{\widetilde{R}_1}\cdot C_{\widetilde{R}_2})=$$
$$=1-((1-C_{R_1})\cdot(1-C_{R_2}))=C_{R_1}+C_{R_2}-C_{R_1}\cdot C_{R_2}$$
$$C_{R_1\to R_2}=C_{\widetilde{R}_1\vee R_2}=C_{\widetilde{R}_1}\cdot C_{R_2}=(1-C_{R_1})\cdot C_{R_2}=C_{R_2}-C_{R_1}\cdot C_{R_2}$$

(2) (a) Since $R(x_1,...,x_n,y)$ is an $n+1$ relation, the relation $(E_y)_{y<z}R(x_1,...,x_n,y)$ will be the $n+1$ relation $R^*(x_1,...,x_n,z)$, whose characteristic function is $C_{R*}(x_1,...,x_n,z)=\prod_{y<z}C_R(x_1,...,x_n,y)$, and so is (primitive) recursive (by 3.1, Theorem). Similarly we can argue for the other cases.

(b) Let $R(x_1,...,x_n,y)$ be a (primitive) recursive relation and let $w\leq y$. Then $\prod_{w\leq y}C_R(x_1,...,x_n,w)$ takes the value 0 as soon as there is some

$w \le y$ such that $R(x_1,...,x_n,w)$ holds; for all the preceding w it takes the value 1. And then for $y<z$ $\sum_{y<z}\left(\prod_{w\le y} C_R(x_1,...,x_n,w)\right)$ is the number of numbers begining with 0 and ending with the first $y<z$ (but not including it) such that $R(x_1,...,x_n,y)$ holds; and this number will be z if such an y doesn't exist. As can be seen this is the number $\mu y_{y<z}R(x_1,...,x_n,y)$, and therefore the function $\mu y_{y<z}R$ is (primitive) recursive (by 3.1, Theorem).

As a simple example, let us assume that $y=8$, $z=9$ and $w \le y$, and that for $y=0,1,2,3$, $C_R(x_1,...,x_n,w)=1$, for $y=4$, $C_R(x_1,...,x_n,w)=0$ and for $y=5,6,7,8,$ $C_R(x_1,...,x_n,y)=1$. Evidently, the first $w \le y$ such that $R(x_1,...,x_n,w)$ holds is 4, and that for $y<9$, $\sum_{y<9}\left(\prod_{w\le 8} C_R(x_1,...,x_n,w)\right)=4$. Hence the number 4 is $\mu y_{y<z}R(x_1,...,x_n,y)$.

Examples (of some primitive rec. functions used further on in our considerations, in addition to those of the list of Sect. 3.1)

1. Let $p_0, p_1, p_2,...$ be the prime numbers in order of magnitude, with $p_0=2$. The function p_x: *the* x^{th} *prime number in ascending order* is primitive recursive, since we have the following recursion:

$$p_0=2$$

$$p_{x+1}=\mu y_{p_x<y\le p_x!+1}\,Pr(y)$$

The upper bound $(p_x)!+1$ on the first prime after p_x is given by Euclid's argument that to any prime number p there is a prime greater then p and less than or equal $p!+1$ (and therefore an argument for the infinitude of primes).

2. By fundamental theorem of arithmetic, every positive integer x has a unique factorization into a product of prime factors, i.e.,

$$x=p_0^{a_0}\cdot p_1^{a_1}\cdot ...\cdot p_k^{a_k}\ ;\ x\ne 0.$$

Let $(x)_i$ be the exponent a_i; and if $x=0$, then $(x)_i=0$ (by def). Then the function $(x)_i$ is primitive recursive since $(x)_i=\mu y_{y<x}(p_i^y \mid x \wedge \sim(p_i^{y+1}\mid x))$.

3. As can be seen, the finite sequence $a_0, a_1, ..., a_k$ of positive integers can be represented by the number $x = p_0^{a_0} \cdot p_1^{a_1} \cdot ... \cdot p_k^{a_k}$. And then $lh(x)$, the length of x, is $k+1$.

So, let $lh(x)$ be the number of non-vanishing exponents in the factorization of x; and if $x = 0$, $lh(x) = 0$.

The function $lh(x)$ is primitive recursive, since for $x > 0$ this number will be the sum of all prime numbers y such that y divides x. Then if $C_R(x, y)$ is the characteristic function of the relation $R(x, y) = Pr(y) \wedge y \mid x$, then $lh(x) = \sum_{y \le x} \overline{sg}(C_R(x, y))$.

4. If x represents the sequence $Seq_1 : a_0, ..., a_k$, i.e., $x = p_0^{a_0} \cdot ... \cdot p_k^{a_k}$, and y represents the sequence $Seq_2 : b_0, ..., b_m$, then $x * y$ will represent the sequence $Seq_1 Seq_2 : a_0, ..., a_k, b_0, ..., b_m$, and this number is $x \cdot \prod_{j < lh(y)} (p_{lh(x)+j})^{(y)_j}$; so the function $x * y$, called *concatenation* or juxtaposition, is primitive recursive.

Example. If $x = 2^{a_0} \cdot 3^{a_1} \cdot 5^{a_2}$ and $y = 2^{b_0} \cdot 3^{b_1} \cdot 5^{b_2} \cdot 7^{b_3}$, then $x * y$ is the number $x \cdot 7^{b_0} \cdot 11^{b_1} \cdot 13^{b_2} \cdot 17^{b_3}$. Now, since $lh(x) = 3$ and $lh(y) = 4$,

$$x * y = x \cdot 7^{(y)_0} \cdot 11^{(y)_1} \cdot 13^{(y)_2} \cdot 17^{(y)_3} = x \cdot \prod_{j<4} (p_{3+j})^{(y)_j} \quad (j = 0,1,2,3).$$

Theorem 2. *Let $f(x_1, ..., x_n)$ be the function defined in the following way:*

$$f(x_1, ..., x_n) = \begin{cases} g_1(x_1, ..., x_n) & \text{if } R_1(x_1, ..., x_n) \\ g_2(x_1, ..., x_n) & \text{if } R_2(x_1, ..., x_n) \\ \vdots & \\ g_m(x_1, ..., x_n) & \text{if } R_m(x_1, ..., x_n) \end{cases}$$

where $R_1, ..., R_m$ are mutually exclusive and $g_1, ..., g_m$ and $R_1, ..., R_m$ are (primitive) recursive.

Then $f(x_1, ..., x_n)$ is (primitive) recursive.

Proof. We must show that $f(x_1, ..., x_n)$ is strictly one of the functions $g_1, ..., g_m$. And that is the case since

$$f(x_1,...,x_n) = g_1(x_1,...,x_n)\cdot \overline{sg}(C_{R_1}(x_1,...,x_n)) + ... +$$
$$+ g_m(x_1,...,x_n)\cdot \overline{sg}(C_{R_m}(x_1,...,x_n)).$$

Example. If $R_j(1 \le j \le m)$ holds, then $C_{R_j} = 0$ and therefore $\overline{sg}(0) = 1$. Hence $f(x_1,...,x_n) = g_j(x_1,...,x_n)$.

3.3. Gödel's β-function

Gödel's β-function[15] gives us a method of dealing with finite sequences $k_0, k_1,...,k_n$ of natural numbers.

The positive integers $d_0, d_1,...,d_n$ are called *relatively prime* if no two of them have a common positive integral factor except 1. Example of relatively prime: 3, 4, 5.

By Gödel's β-function the numbers $k_0, k_1,...,k_n$ are the values of the β-function $\beta(x_1, x_2, x_3)$.

Let us begin with an example.

Let $c = 0,1,2,...$, $d_0 = 3$, $d_1 = 4$ and the function $rm(c,d)$. As can be seen, if $c = 0,1,...,11$, the pairs $(rm(c,3), rm(c,4))$ are, respectively, (0,0), (1,1), (2,2), (0,3), (1,0), (2,1), (0,2), (1,3), (2,0), (0,1), (1,2) and (2,3); i.e., they exhaust the 12 possible ordered pairs (k_0, k_1) such that $k_0 < 3$ and $k_1 < 4$. And this sequence is repeating as *c* increases from 12 to 23, or from 24 to 35, and so on. This means that if the pair (k_0, k_1) is the pair for $c = j$, then it is also the pair for $c = j + 12$.

Let $k = 12$. Since d_0 and d_1 are relatively prime, the pair (k_0, k_1) occurs again after $d_0 \cdot d_1 = 12$ consecutive values of *c*, and this number is exactly the number of all possible pairs (k_0, k_1) such that $k_0 < d_0$ and $k_1 < d_1$.

Now, in order to show that for the two relative prime numbers d_0, d_1 there is a number *c* such that $rm(c, d_0) = k_0$ and $rm(c, d_1) = k_1$ is sufficient to show that every one of the $d_0 \cdot d_1 = 12$ numbers, with $0 \le c < d_0 d_1$, does

[15] Cf. K. Gödel [1934], in vol. K. Gödel [1986], I, 346-371. Comp. and S.C. Kleene, [1952], §48.

generate one of the 12 pairs (k_0, k_1). And to show that every pair (k_0, k_1) is generated by some number *c*, it suffices to show that no two distinct numbers $0 \leq c < r < d_0 d_1$ do generate the same pair. Let us suppose, by *reductio*, that *c* and *r* produce the same pair. In this case these numbers give the same remainder when divided by d_0, and give the same remainder when divided by d_1. This does imply that their difference $q = r - c$ gives the rest zero when it is divided by d_0 or by d_1, i.e., both numbers d_0 and d_1 divide *q*. Now, since d_0 and d_1 are relatively prime, it follows that the product $d_0 d_1$ also divides *q*. But this is impossible since $0 < q < d_0 d_1$.

By Gödel's β-function the numbers k_0 and k_1 are the values of a function $\beta(x_1, x_2, x_3)$, i.e., $\beta(c, d, i) = k_i$, $i = 0,1$. More exactly, if d_0 and d_1 are relatively prime numbers, a number *c* can be found such that $c < d_0 d_1$ and $rm(c, d_i) = k_i$, $i = 0,1$.

Theorem 1. *For any finite sequence of natural numbers* $k_0, k_1, ..., k_n$, *there exist natural numbers c and d such that* $\beta(c, d, i) = k_i$ $(0 \leq i \leq n)$.

Proof. Gödel's β-function has the following form:

$$\beta(x_1, x_2, x_3) = rm(x_1, 1 + (x_3 + 1)x_2),$$

or equivalently:[16]

$$\beta(c, d, i) = rm(c, 1 + (i + 1)d) \text{, or, finally,}$$

$$\beta(c, d, i) = rm(c, \delta(d, i)) \text{, where } \delta(d, i) = 1 + (i + 1)d.$$ [17]

What must be proved is that

(1) The numbers $d_0, d_1, ..., d_n$ are relatively prime

(2) $k_i < d_i$ $(0 \leq i \leq n)$.

If $k_0, k_1, ..., k_n$ is the given sequence of natural numbers, then let $s = \max(n, k_0, k_1, ..., k_n)$ and $d = s!$.

(1) *Reductio*. Suppose there are two numbers, $1 + (j + 1)s!$ and $1 + (j + m + 1)s!$ which are not relatively prime; i.e., there is a factor in common other than 1, and then there is a prime factor *p* dividing both. In this case *p* will also divide their difference $m \cdot s!$. But

---

[16] "$\beta(c,d,i)$" is Gödel's original notation; comp. K. Gödel [1986], 365.

[17] Kleene's proposal, [1952], 240. As can be seen Gödel's β-function is primitive recursive (by 3.1, Theorem 1,12).

(1)a. The number p cannot divide $s!$, since otherwise it will divide any multiple of $s!$, particularly $(j+1)s!$. And since p divides $1+(j+1)s!$ (by hypothesis) would follow that p divides 1 which is impossible.

(1)b. The number p cannot divide m, since $m \leq n \leq s$ and every number $\leq s$ divides $s!$.

It follows that p does not divide $m \cdot s!$ and therefore the numbers d_i $(0 \leq i \leq n)$ are relatively prime.

(2) $k_i < d_i$ $(0 \leq i \leq n)$, since $k_i \leq s \leq s! < 1+(i+1)s! = \delta(d,i) = d_i$.

Theorem 2. $\beta(x_1,x_2,x_3) = rm(x_1, 1+(x_3+1)\cdot x_2)$ *is representable in* $\mathrm{PA}^{\mathrm{ax}}$.

This means that there is a formula $\beta^*(\mathrm{x}_1,\mathrm{x}_2,\mathrm{x}_3,\mathrm{x}_4)$ of $\mathrm{L_{PA}}$ such that the following holds:

(a) *If* $\beta(k_1,k_2,k_3,k_4)$, *then* $\mathrm{PA}^{\mathrm{ax}} \vdash \beta^*(\bar{k}_1,\bar{k}_2,\bar{k}_3,\bar{k}_4)$.

(b) $\mathrm{PA}^{\mathrm{ax}} \vdash \exists_1 \mathrm{x}_4 \beta^*(\bar{k}_1,\bar{k}_2,\bar{k}_3,\mathrm{x}_4)$.

Proof (a). In order to display the formula representing $\beta(x_1,x_2,x_3,x_n)$ in $\mathrm{PA}^{\mathrm{ax}}$, let us remember that $rm(x,y)$ (i.e., the remainder of the division of x by y) does imply that there is a n such that $x = y\cdot n + r$ and $r < y$. In our case this means that if $\beta(k_1,k_2,k_3) = k_4$, i.e., $rm(k_1, 1+(k_3+1)\cdot k_2) = k_4$, then there is an n such that

(1) $k_1 = (1+(k_3+1)\cdot k_2)\cdot n + k_4$, and

(2) $k_4 < 1+(k_3+1)\cdot k_2$. Hence

(3) $\vdash \bar{k}_1 = (\bar{1} + (\bar{k}_3 + \bar{1})\cdot \bar{k}_2)\cdot \bar{n} + \bar{k}_4$ (by Sect. 1.4 Theorem I.1)

(4) $\vdash \bar{k}_4 < \bar{1} + (\bar{k}_3 + \bar{1})\cdot \bar{k}_2$ (by expressibility of the relation $x < y$ in $\mathrm{PA}^{\mathrm{ax}}$ (cf. Sect. 2, Example to Def.3).

(5) $\vdash \bar{k}_1 = (\bar{1} + (\bar{k}_3 + \bar{1})\cdot \bar{k}_2)\cdot \bar{n} + k_4 \wedge k_4 < \bar{1} + (\bar{k}_3 + \bar{1})\cdot \bar{k}_2$;

(3), (4), PL

(6) $\vdash \exists \mathrm{z}[\bar{k}_1 = (\bar{1} + (\bar{k}_3 + \bar{1})\cdot \bar{k}_2)\cdot \mathrm{z} + k_4 \wedge k_4 < \bar{1} + (\bar{k}_3 + \bar{1})\cdot \bar{k}_2]$;

(5), Gen $\exists$

This last expression suggests the construction of the formula representing $\beta(x_1,x_2,x_3)$ in $\mathrm{PA}^{\mathrm{ax}}$, i.e.,

$\beta^*(\mathrm{x}_1,\mathrm{x}_2,\mathrm{x}_3,\mathrm{x}_4) : \exists \mathrm{z}[\mathrm{x}_1 = (\bar{1} + (\mathrm{x}_3 + \bar{1})\cdot \mathrm{x}_2)\cdot \mathrm{z} + \mathrm{x}_4 \wedge \mathrm{x}_4 < \bar{1} + (\mathrm{x}_3 + \bar{1})\cdot \mathrm{x}_2$.

As (6) shows, part (a) holds.

(b). This part, $\vdash \exists_1 x_4 \beta^*(\bar{k}_1, \bar{k}_2, \bar{k}_3, x_4)$, follows directly from the theorem of the uniqueness of quotient and remainder (cf 1.3, Th. VI).

3.4. Representability of recursive functions and expressibility of recursive relations in PA^{ax}

Theorem. *Every recursive function is formally representable in* PA^{ax}.

Proof.

I. $f(x) = x'$ (the successor function)

This function is representable in PA^{ax} by the formula $\alpha(x_1, x_2)$: $x_2 = (x_1)'$. Since for any number k_1 if $f(k_1) = k_2$, then $k_2 = (k_1)'$ and then the terms $\bar{k}_2$ and $(\bar{k}_1)'$ are identical. Whence (by Sect. 1.4, Th. I.1)

1. If $k_2 = (k_1)'$, then $\vdash \bar{k}_2 = (\bar{k}_1)'$ (i.e., $\vdash \alpha(\bar{k}_1, \bar{k}_2)$).
2. $\vdash \exists_1 x_2 (x_2 = (\bar{k}_1)')$ (i.e., $\vdash \exists_1 x_2 \alpha(\bar{k}_1, x_2)$).

II. $f(x_1, ..., x_n) = q$ (the constant function)

This function is representable in PA^{ax} by the formula

$\alpha(x_1, ..., x_n, x_{n+1})$: $(x_1 = x_1 \wedge ... \wedge x_n = x_n \wedge x_{n+1} = \bar{q})$.

If $f(k_1, ..., k_n) = k_{n+1}$, then $k_{n+1} = q$ and therefore the terms $\bar{k}_{n+1}$ and $\bar{q}$ are identical. Whence

1. If $f(k_1, ..., k_n) = q$, then $\vdash \bar{k}_1 = \bar{k}_1 \wedge ... \wedge \bar{k}_n = \bar{k}_n \wedge \bar{q} = \bar{q}$.
2. $\vdash \exists_1 x_{n+1} (\bar{k}_1 = \bar{k}_1 \wedge ... \wedge \bar{k}_n = \bar{k}_n \wedge x_{n+1} = \bar{q})$.

III. $f_i^n(x_1, ..., x_n) = x_i$ (the projection functions)

This function is representable in PA^{ax} by the formula

$$\alpha(x_1, ..., x_n, x_{n+1}): x_1 = x_1 \wedge ... \wedge x_n = x_n \wedge x_{n+1} = x_i.$$

Since if $f(k_1, ..., k_n) = k_{n+1}$, then the terms $\bar{k}_{n+1}$ and $\bar{k}_i$ are identical, and then

1. If $f(k_1, ..., k_n) = k_{n+1}$, then $\vdash \bar{k}_1 = \bar{k}_1 \wedge ... \wedge \bar{k}_n = \bar{k}_n \wedge \bar{k}_{n+1} = \bar{k}_i$.
2. $\vdash \exists_1 x_{n+1} (\bar{k}_1 = \bar{k}_1 \wedge ... \wedge \bar{k}_n = \bar{k}_n \wedge x_{n+1} = \bar{k}_i)$.

(IV) Substitution

$$f(x_1, ..., x_n) = g(h_1(x_1, ..., x_n), ..., h_m(x_1, ..., x_n))$$

(where f is obtained by substitution from $g(y_1, .., y_m)$ and $h_1(x_1, ..., x_n)$,..., $h_m(x_1, ..., x_n)$.

Proof. Firstly, we have to prove 1. if $f(k_1,\ldots,k_n)=k$ then $\vdash \alpha(\bar{k}_1,\ldots,\bar{k}_n,\bar{k})$, where $\alpha(x_1,\ldots,x_n,x_{n+1})$ is the formula which represents f in PA^{ax}.

Let $f(x_1,\ldots,x_n)=g(h_1(x_1,\ldots,x_n),\ldots h_m(x_1,\ldots,x_n))$. To find the formula representing f in PA^{ax} we proceed as follows:

Suppose that the functions h_i^n and g^m are representable in PA^{ax} by $\alpha_i(x_1,\ldots,x_n,x_{n+1})$ and $\beta(x_1,\ldots,x_m,x_{m+1})$, respectively.

Suppose now that $f(k_1,\ldots,k_n)=k$, $h_i(k_1,\ldots,k_n)=a_i$. Then $g(a_1,\ldots,a_m)=k$ (by def. of f). Then we have:

(1) $\vdash \alpha_i(\bar{k}_1,\ldots,\bar{k}_n,\bar{a}_i)$, i.e., $\vdash \alpha_1(\bar{k}_1,\ldots,\bar{k}_n,\bar{a}_1)$
$$\vdash \alpha_2(\bar{k}_1,\ldots,\bar{k}_n,\bar{a}_2)$$
$$\vdots$$
$$\vdash \alpha_m(\bar{k}_1,\ldots,\bar{k}_n,\bar{a}_m)$$

(2) $\vdash \beta(\bar{a}_1,\ldots,\bar{a}_m,\bar{k})$ (since β represents g)

(3) $\vdash \alpha_1(\bar{k}_1,\ldots,\bar{k}_n,\bar{a}_1)\wedge\ldots\wedge\alpha_m(\bar{k}_1,\ldots,\bar{k}_n,\bar{a}_m)\wedge\beta(\bar{a}_1,\ldots,\bar{a}_m,\bar{k})$;
(1), (2), PL

(4) $\vdash \exists y_1\ldots\exists y_m[\alpha_1(\bar{k}_1,\ldots,\bar{k}_n,y_1)\wedge\ldots\wedge\alpha_m(\bar{k}_1,\ldots,\bar{k}_n,y_m)\wedge$
$\wedge\beta(y_1,\ldots,y_m,\bar{k})]$; (3), Gen $\exists$ *m*-times

Therefore, if $f(k_1,\ldots,k_n)=k$, then the formula in (4) is provable. And then the formula representing $f(x_1,\ldots,x_n)$ in PA^{ax} is:

FORM. $\exists y_1\ldots\exists y_m[\alpha_1(x_1,\ldots,x_n,y_1)\wedge\ldots\wedge\alpha_m(x_1,\ldots,x_n,y_m)\wedge$
$\wedge\beta(y_1,\ldots,y_m,x_{n+1})]$.

2. $\vdash \exists_1 x_{n+1}\alpha(\bar{k}_1,\ldots,\bar{k}_n,x_{n+1})$.

The *existence* part follows directly from (a) by Gen $\exists$. For the proof of *uniqueness* we proceed as follows:

(1) Take the formula $\alpha(\bar{k}_1,\ldots,\bar{k}_n,x_{n+1})$, i.e.,
$\exists y_1\ldots\exists y_m(\alpha(\bar{k}_1,\ldots,\bar{k}_n,y_1)\wedge\ldots\wedge\alpha_m(\bar{k}_1,\ldots,\bar{k}_n,y_m)\wedge\beta(y_1,\ldots,y_m,x_{n+1}))$.

(2) From (1) construct the following two formulas by *m* applications of C-Rule (replacing $y_1,\ldots,y_m$ alternatively with $n_1,\ldots,n_m$ and $n'_1,\ldots,n'_m$, respectively), and by setting $x_{n+1}=z$ and $x_{n+1}=w$, respectively, i.e.,
(a) $\alpha_1(\bar{k}_1,\ldots,\bar{k}_n,n_1)\wedge\ldots\wedge\alpha_m(\bar{k}_1,\ldots,\bar{k}_n,n_m)\wedge\beta(n_1,\ldots,n_m,z)$

(b) $\alpha_1(\bar{k}_1,...,\bar{k}_n,n'_1)\wedge...\wedge\alpha_m(\bar{k}_1,...,\bar{k}_n,n'_m)\wedge\beta(n'_1,...,n'_m,w)$

(3) Now, we reason as follows: since $\alpha_i(x_1,...,x_n,x_{n+1})$ represent $h_i(x_1,...,x_n)$ in PA^{ax} it follows that $PA^{ax}\vdash\exists_1 x_{n+1}\alpha_i(\bar{k}_1,...,\bar{k}_n,x_{n+1})$, which with $\alpha_i(\bar{k}_1,...,\bar{k}_n,n_i)$ and $\alpha_i(\bar{k}_1,...,\bar{k}_n,n'_i)$ give: $n_i=n'_i$. And from $\beta(n_1,...,n_m,z)$ (the last conjunct of (a)) and $n_i=n'_i$ it follows $\beta(n'_1,...,n'_m,z)$.[18] And since $\beta(x_1,...,x_{m+1})$ represents $g(x_1,...,x_m)$ in PA^{ax}, it follows that $PA^{ax}\vdash\exists_1 x_{n+1}\beta(n'_1,...,n'_m,x_{n+1})$, which together with $\beta(n'_1,...,n'_m,w)$ (the last conjunct of (b)) give $z=w$.

(4) Re-introduce, by Gen $\exists$, the existential quantifiers in (a) and (b) and obtain:

(a*) $\exists y_1...\exists y_m(\alpha_1(\bar{k}_1,...,\bar{k}_n,y_1)\wedge...\wedge\alpha_m(\bar{k}_1,...,\bar{k}_n,y_m)\wedge$
$\wedge\beta(y_1,...,y_m,z)$; i.e., $\alpha(\bar{k}_1,...,\bar{k}_n,z)$

(b*) $\exists y_1...\exists y_m(\alpha_1(\bar{k}_1,...,\bar{k}_n,y_1)\wedge...\wedge\alpha_m(\bar{k}_1,...,\bar{k}_n,y_m)\wedge$
$\wedge\beta(y_1,...,y_m,w)$; i.e., $\alpha(\bar{k}_1,...,\bar{k}_n,w)$

From (1)-(4) it follows therefore that

$$\alpha(\bar{k}_1,...,\bar{k}_n,z)\wedge\alpha(\bar{k}_1,...,\bar{k}_n,w)\vdash z=w,$$

whence, by Ded. Th. $\vdash\alpha(\bar{k}_1,...,\bar{k}_n,z)\wedge\alpha(\bar{k}_1,...,\bar{k}_n,w)\supset z=w$, whence $PA^{ax}\vdash\exists_1 x_{n+1}\alpha(\bar{k}_1,...,\bar{k}_n,x_{n+1})$, i.e., (2).

(V) Recursion

$$\begin{cases} f(x_1,...,x_n,0)=g(x_1,...,x_n) \\ f(x_1,...,x_n,y+1)=h(x_1,...,x_n,y,f(x_1,...,x_n,y)) \end{cases}$$

Suppose that the functions $g(x_1,...,x_n)$ and $h(x_1,...,x_n,y,z)$ are representable in PA^{ax} by the formulas $\alpha(x_1,...,x_n,x_{n+1})$ and $\beta(x_1,...,x_n,x_{n+1},x_{n+2},x_{n+3})$, respectively. We must show that the recursive function $f(x_1,...,x_n,y)$, defined by (V) is representable in PA^{ax} i.e., there exists a formula $\gamma(x_1,...,x_n,x_{n+1},x_{n+2})$ such that for any $n+2$-tupel of natural numbers $k_1,...,k_n,k_{n+1},k_{n+2}$ the following holds:

(a) If $f(k_1,...,k_n,k_{n+1})=k_{n+2}$, then $\vdash\gamma(\bar{k}_1,...,\overline{k_{n+1}},\overline{k_{n+2}})$

[18] By x = y, $\alpha(x)\vdash\alpha(y)$.

(b) $\vdash \exists_1 \mathrm{x}_{n+2}\gamma(\bar{k}_1,...,\bar{k}_{n+1},\mathrm{x}_{n+2})$.

Proof.[19] (a). If $f(k_1,...,k_n,k_{n+1}) = k_{n+2}$, then $\vdash \gamma(\bar{k}_1,...,\bar{k}_n,\bar{k}_{n+1},\bar{k}_{n+2})$. The following considerations alows us to construct effectively the formula $\gamma(\mathrm{x}_1,...,\mathrm{x}_n,\mathrm{x}_{n+1},\mathrm{x}_{n+2})$ which represents $f(x_1,...,x_n,x_{n+1})$ in PA$^{\mathrm{ax}}$, and therefore to prove (a).

First of all, assume that $f(k_1,...,k_n,k) = m$ and consider the two cases: $k = 0$ and $k > 0$. In the first case, if we take the sequence whose only member is m, then by Theorem 1 (of Sect. 3.3) there are (and can be found) the numbers c, d such that $\beta(c,d,0) = m$. And then $\beta(c,d,0) = m = f(k_1,...,k_n,0) = g(k_1,...,k_n)$, where $g(x_1,...,x_n)$ is represented in PA$^{\mathrm{ax}}$ by $\alpha(\mathrm{x}_1,...,\mathrm{x}_n,\mathrm{x}_{n+1})$. And then

(a) $\vdash \beta^*(\bar{c},\bar{d},0,\bar{m})$ and

(b) $\vdash \alpha(\bar{k}_1,...,\bar{k}_n,\bar{m})$. So

(c) $\vdash \beta^*(\bar{c},\bar{d},0,\bar{m}) \wedge \alpha(\bar{k}_1,...,\bar{k}_n,\bar{m})$; (a), (b), PL; whence, by Gen $\exists$, we get

(d) $\vdash \exists \mathrm{w}[\beta^*(\bar{c},\bar{d},0,\mathrm{w}) \wedge \alpha(\bar{k}_1,...,\bar{k}_n,\mathrm{w})]$. And then it follows that

(e) $\vdash \exists \mathrm{w}[\beta^*(\bar{c},\bar{d},0,\mathrm{w}) \wedge \alpha(\bar{k}_1,...,\bar{k}_n,\mathrm{w})] \wedge \beta^*(\bar{c},\bar{d},0,\bar{m})$; (d), (a), PL.

For the second case, $k > 0$, suppose that $f(k_1,...,k_n,i) = b_i$ and that $b_0,b_1,...,b_k$ is the corresponding sequence of values ($0 \leq i \leq k$). By Sect. 3.3, Theorem 1, there exist the numbers c and d such that $\beta(c,d,i) = b_i$. And then we have:

$\vdash \beta^*(\bar{c},\bar{d},0,\bar{b}_0)$, $\vdash \beta^*(\bar{c},\bar{d},\bar{1},\bar{b}_1)$,..., $\vdash \beta^*(\bar{c},\bar{d},\bar{k},\bar{b}_k)$,

i.e., for $0 \leq i \leq k$, $\vdash \beta^*(\bar{c},\bar{d},\bar{i},\bar{b}_i)$.

For $i = 0$, for example, we have $\beta(c,d,0) = b_0 = f(k_1,...,k_n,0) = g(k_1,...,k_n)$. And since $\alpha(\mathrm{x}_1,...,\mathrm{x}_n,\mathrm{x}_{n+1})$ represents $g(x_1,...,x_n)$ in PA$^{\mathrm{ax}}$, it follows that

(a) $\vdash \beta^*(\bar{c},\bar{d},0,\bar{b}_0)$ and

(b) $\vdash \alpha(\bar{k}_1,...,\bar{k}_n,\bar{b}_0)$, and then

---

[19] This proof is based on Kleene's proof; comp. S.C. Kleene [1952], §49, Theorem I, Theorem 27 and Mendelson's proof; cf. E. Mendelson [1964], Ch. 3, 131-132, Proposition 3.23.

(c) $\vdash \beta^*(\bar{c},\bar{d},0,\bar{b_0}) \wedge \alpha(\bar{k_1},...,\bar{k_n},b_0)$; (a), (b), PL, whence

(d) $\vdash \exists w(\beta^*(\bar{c},\bar{d},0,w) \wedge \alpha(\bar{k_1},...,\bar{k_n},w))$; (c), Gen $\exists$

(e) $\vdash \beta^*(\bar{c},\bar{d},\bar{k},\bar{m})$; since $b_k = f(k_1,...,k_n,k) = m$ and $\beta(c,d,k) = m$.

As we saw, for $0 \le i \le k-1$, $\beta(c,d,i) = b_i$ and the following holds:

(EQ) $\beta(c,d,i+1) = b_{i+1} = f(k_1,...,k_n,i+1) = h(k_1,...,k_n,i,b_i)$.

And then

(f) $\vdash \beta^*(\bar{c},\bar{d},\bar{i},\bar{b_i}) \wedge \beta^*(\bar{c},\bar{d},\overline{i+1},\overline{b_i+1}) \wedge \beta(\bar{k_1},...,\bar{k_n},\bar{i},\bar{b_i},\overline{b_{i+1}})$;

by (EQ) and the fact that $\beta(x_1,...,x_n,x_{n+1},x_{n+2},x_{n+3})$ represents $h(x_1,...,x_n,y,z)$ in PA^{ax}.

(g) $\vdash \exists y \exists z(\beta^*(\bar{c},\bar{d},\bar{i},y) \wedge \beta^*(\bar{c},\bar{d},\overline{i+1},z) \wedge \beta(\bar{k_1},...,\bar{k_n},\bar{i},y,z)$;

(f) by Gen $\exists$ (two times).

(h) $\vdash \forall w(w < \bar{k} \supset \exists y \exists z(\beta^*(\bar{c},\bar{d},w,y) \wedge \beta^*(\bar{c},\bar{d},w',z)$
$\wedge \beta(\bar{k_1},...,\bar{k_n},w,y,z)))$; (g) by Sect. 1.4, Theorem V.2*.

(i) $\vdash \exists w(\beta^*(\bar{c},\bar{d},0,w) \wedge \alpha(\bar{k_1},...,\bar{k_n},w) \wedge \beta^*(\bar{c},\bar{d},\bar{k},\bar{m})$
$\wedge \forall w(w < \bar{k} \supset \exists y \exists z(\beta^*(\bar{c},\bar{d},w,y) \wedge \beta^*(\bar{c},\bar{d},w',z) \wedge \beta(\bar{k_1},...,\bar{k_n},w,y,z)))$;

(d), (e), (h), PL.

(j) $\vdash \exists u \exists v\{\exists w(\beta^*(u,v,0,w) \wedge \alpha(\bar{k_1},...,\bar{k_n},w) \wedge \beta^*(u,v,\bar{k},\bar{m})$
$\wedge \forall w(w < \bar{k} \supset \exists y \exists z(\beta^*(u,v,w,y) \wedge \beta^*(u,v,w',z) \wedge \beta(\bar{k_1},...,\bar{k_n},w,y,z))\}$;

(i) by Gen $\exists$ (two times).

But this is just the formula $\gamma(\bar{k_1},...,\bar{k_n},\overline{k_{n+1}},\overline{k_{n+2}})$, and the way it is obtained suggests the construction of the formula representing $f(x_1,...,x_n,x_{n+1})$ in PA^{ax}, i.e.,

$\gamma(x_1,...,x_n,x_{n+1},x_{n+2}) : \exists u \exists v\{\exists w(\beta^*(u,v,0,w) \wedge \alpha(x_1,...,x_n,w))$
$\wedge \beta^*(u,v,x_{n+1},x_{n+2}) \wedge \forall w(w < x_{n+1} \supset \exists y \exists z(\beta^*(u,v,w,y)$
$\wedge \beta^*(u,v,w',z) \wedge \beta(x_1,...,x_n,w,y,z))\}$.

And since $\vdash \gamma(\bar{k_1},...,\bar{k_n},\overline{k_{n+1}},\overline{k_{n+2}})$ this is the proof of (a).

Remark. For the first case, $k = 0$, in the proof above, let us observe that the implication in the scope of "$\forall w$" in (j) is provable, since its antecedent is $w < \bar{k}$ and since PA$^{ax} \vdash \neg(w < 0)$, and then the implication follows by PL: $\vdash \neg p \supset (p \supset q)$ and Rule$_P$.

It remains to show that

(b) $\vdash \exists_1 x_{n+2}\gamma(\bar{k}_1,...,\bar{k}_n,\bar{k},x_{n+2})$

But the *existence* of such an x_{n+2} follows by Gen $\exists$ from the part (a), i.e., $\vdash \gamma(\bar{k}_1,...,\bar{k}_n,\bar{k},\bar{m})$. So only *uniqueness* is still to be proved.

Proof. (b) (intuitive induction[20] on *k*)

Basis. For $k = 0$ the result follows directly from what we said above: by (1) and (2) of the part (a) and *Remark.*

Induction. Let $n_1 = g(k_1,...,k_n)$, $n_2 = f(k_1,...,k_n,k)$ and

$n_3 = f(k_1,...,k_n,k+1) = h(k_1,...,k_n,k,f(k_1,...,k_n,k)) = h(k_1,...,k_n,k,n_2)$.

As above, the functions $g(x_1,...,x_n)$ and $h(x_1,...,x_n,y,z)$ are representable in PA^{ax} by $\alpha(x_1,...,x_n,x_{n+1})$ and $\beta(x_1,...,x_n,x_{n+1},x_{n+2},x_{n+3})$, respectively. And therefore we have

1. $\vdash \alpha(\bar{k}_1,...,\bar{k}_n,\bar{n}_1)$; by Part (a) (of the theorem)
2. $\vdash \beta(\bar{k}_1,...,\bar{k}_n,\bar{k},\bar{n}_2,\bar{n}_3)$; by Part (a)
3. $\vdash \gamma(\bar{k}_1,...,\bar{k}_n,\bar{k},\bar{n}_2)$; by Part (a)
4. $\vdash \gamma(\bar{k}_1,...,\bar{k}_n,\overline{k+1},\bar{n}_3)$; by Part (a)
5. $\vdash \exists_1 x_{n+2}\gamma(\bar{k}_1,...,\bar{k}_n,\bar{k},x_{n+2})$; by hyp. of induction.

We must show that $\vdash \exists_1 x_{n+2}\gamma(\bar{k}_1,...,\bar{k}_n,\bar{k}_{n+1},x_{n+2})$.

6. $\gamma(\bar{k}_1,...,\bar{k}_n,\bar{k}_{n+1},x_{n+2})$; hyp; and prove that $x_{n+2} = \bar{n}_3$

In extenso, the formula $\gamma(\bar{k}_1,...\bar{k}_n,\overline{k_{n+1}},x_{n+2})$ is

FORM. $\exists u\exists v\{\exists w(\beta^*(u,v,0,w) \wedge \alpha(\bar{k}_1,...,\bar{k}_n,w) \wedge \beta^*(u,v,\overline{k+1},x_{n+2})$
$\wedge \forall w(w < \overline{k+1} \supset \exists y\exists z(\beta^*(u,v,w,y) \wedge \beta^*(u,v,w',z) \wedge \beta(\bar{k}_1,...,\bar{k}_n,w,y,z))\}$.

(1) $\exists w(\beta^*(c,d,0,w) \wedge \alpha(\bar{k}_1,...,\bar{k}_n,w)$; from 6 (comp. FORM), using *C-Rule* twice for the variables u and v

(2) $\forall w(w < \overline{k+1} \supset \exists y\exists z[\beta^*(c,d,w,y) \wedge$
$\wedge \beta^*(c,d,w',z) \wedge \beta(\bar{k}_1,...,\bar{k}_n,w,y,z)]$;

(3) $\beta^*(c,d,\overline{k+1},x_{n+2})$

(4) $\forall w(w < \bar{k} \supset \exists y\exists z[\beta^*(c,d,w,y) \wedge$
$\wedge \beta^*(c,d,w',z) \wedge \beta(\bar{k}_1,...,\bar{k}_n,w,y,z)]$, from (2)

---

[20] I.e. the application of induction in the metalanguage.

(5) $\beta^*(c,d,\bar{k},r) \wedge \beta^*(c,d,\overline{k+1},s) \wedge \beta(\bar{k}_1,...,\bar{k}_n,\bar{k},r,s)$; from (2), using *C-Rule* twice for y and z.

(5) (a) $\beta^*(c,d,\bar{k},r)$; from (5), by PL

(b) $\beta^*(c,d,\overline{k+1},s)$; from (5), by PL

(c) $\beta(\bar{k}_1,...,\bar{k}_n,\bar{k},r,s)$; from (5), by PL

(6) $\gamma(\bar{k}_1,...,\bar{k}_n,\bar{k},r)$; (1), (4), (5) (a)

(7) $r=\bar{n}_2$; by (6) and 5

(8) $\beta(\bar{k}_1,...,\bar{k}_n,\bar{k},\bar{n}_2,s)$; from (5) (c), (7)

(9) $\bar{n}_3 = s$; by hyp. of ind., 2 and (8)

(10) $\beta^*(c,d,\overline{k+1},\bar{n}_3)$; from (5) (b) and (9)

(11) $x_{n+2} = \bar{n}_3$; from (3) and (10), by Sect. 3.3 Th. 2 (Gödel's β-function is representable in PA^{ax} by $\beta^*(x_1,x_2,x_3,x_4)$ and so we have $\exists_1 x_4 \beta(c,d,\overline{k+1},x_4)$).

(VI) μ-*Operator*

$$f(x_1,...,x_n)=\mu y(g(x_1,...,x_n,y)=0)$$

(where $g(x_1,...,x_n,y)$ is a function such that for any $x_1,...,x_n$ there is at least one y such that $g(x_1,...,x_n,y)=0$).

The function $f(x_1,...,x_n)$ so defined is representable in PA^{ax}. I.e., there exists a formula $\alpha(x_1,...,x_n,x_{n+1})$ such that for any numbers $k_1,...,k_{n+1}$ holds:

1. If $f(k_1,...,k_n)=k_{n+1}$, then $\vdash \alpha(\bar{k}_1,...,\bar{k}_n,\bar{k}_{n+1})$
2. $\vdash \exists_1 x_{n+1} \alpha(\bar{k}_1,...,\bar{k}_n,x_{n+1})$.

Proof. Suppose that $g(x_1,...,x_n,y)$ is representable in PA^{ax} by $\beta(x_1,...,x_n,x_{n+1},x_{n+2})$ and this means that

(a) If $g(k_1,...,k_n,k_{n+1})=k_{n+2}$, then $\vdash \beta(\bar{k}_1,...,\bar{k}_n,\bar{k}_{n+1},\bar{k}_{n+2})$.

(b) $\vdash \exists_1 x_{n+2}(\bar{k}_1,...,\bar{k}_{n+1},x_{n+2})$.

Let $f(x_1,...,x_n)=\mu y(g(x_1,...,x_n,y)=0)$.

To find the formula α which represents *f* in PA^{ax} we proceed as follows:

(1) Suppose that $f(k_1,...,k_n) = k$. From this follows

(a) $g(k_1,...,k_n,k) = 0$.

(b) For any $m < k$: $g(k_1,...,k_n,m) \neq 0$.

(2) From (a) it follows that $\vdash \beta(\bar{k}_1,...,\bar{k}_n,\bar{k},0)$ and from (b) it follows that $\vdash \forall y(y < \bar{k} \supset \neg\beta(\bar{k}_1,...,\bar{k}_n,y,0)$, by Sect. 1.4, Th. V.2. Whence, by PL

$\vdash \beta(\bar{k}_1,...,\bar{k}_n,\bar{k},0) \wedge \forall y(y < \bar{k} \supset \neg\beta(\bar{k}_1,...,\bar{k}_n,y,0))$.

From (1) and (2) it follows that the formula which represents $f(x_1,...,x_n)$ so defined in PA^{ax} is $\alpha(x_1,...,x_n,x_{n+1})$:

$\beta(x_1,...,x_n,x_{n+1},0) \wedge \forall y(y < x_{n+1} \supset \neg\beta(x_1,...,x_n,y,0)$.

By (1) and (2) follows the part (a) of the proof.

The *existence* part of the proof of (b) follows from (a) and Gen ∃. The *uniqueness* part is provable by the following argument.

(1) Take the formula $\alpha(\bar{k}_1,...,\bar{k}_n,x_{n+1},0)$, i.e.,

$\beta(\bar{k}_1,...,\bar{k}_n,x_{n+1},0) \wedge \forall y(y < x_{n+1} \supset \neg\beta(\bar{k}_1,...,\bar{k}_n,y,0)$.

(2) From (1) we construct the following two formulas, taking alternatively for x_{n+1} z and w, respectively.

(a) $\beta(\bar{k}_1,...,\bar{k}_n,z,0) \wedge \forall y(y < z \supset \neg\beta(\bar{k}_1,...,\bar{k}_n,y,0)$

(b) $\beta(\bar{k}_1,...,\bar{k}_n,w,0) \wedge \forall y(y < w \supset \neg\beta(\bar{k}_1,...,\bar{k}_n,y,0)$.

(3) Now, we reason as follows: We consider the following two cases: $z < w$ and $w < z$. In the first case, we derive $\neg\beta(\bar{k}_1,...,\bar{k}_n,z,0)$ (from the second conjunct of (b), Ax 4 and MP), a result which contradicts $\beta(\bar{k}_1,...,\bar{k}_n,z,0)$, a formula derivable from (a) by PL. Similarly, if $w < z$ we derive $\neg\beta(\bar{k}_1,...,\bar{k}_n,w,0)$ and $\beta(\bar{k}_1,...,\bar{k}_n,w,0)$. But, as we know (Sect. 1.4, Th. III.11), $PA^{ax} \vdash w = z \vee w < z \vee z < w$. Whence, by this argument it follows $z = w$ (i.e., uniqueness required by part (b) of representability).

This is the proof of representability in PA^{ax} of every recursive function.

Remark. From the above theorem easily follows that every recursive *relation* is formally expressible in PA^{ax}, since if R^n is recursive, then its characteristic function C_{R^n} is recursive (by Sect. 3.2, Definition), and then by the above theorem C_{R^n} is representable in PA^{ax}. Whence, by Sect. 2 Theorem, R^n is expressible in PA^{ax}.

4. Gödel's Theorems for PA^{ax}

4.1. Arithmetization of the syntax of PA^{ax}

A fundamental distinction regarding an axiomatic theory is that between its object language (OL) and its metalanguage (ML). As we know, $x\cdot(y+z)=x\cdot y+x\cdot z$ is a formula of L_{PA}; so are $x=x$ or $\exists z(x+z'=y)$. But expressions like "*x* is a formula", "*x* is a proof", "PA^{ax} is consistent" do not belong to OL of PA, but to its ML. I.e., they are metamathematical expressions[21], and so they belong to an informal mathematical theory, loosley defined as containing a part of English language, the primitive symbols of L_{PA} or finite sequences of them, concepts like "term", "formula", "axiom", "substitution", "proof", "provable" and "sentence".

But a simple question arises: can the formal system of PA^{ax} "talk about" its own syntax? The answer is "yes!" and, as can be shown, PA^{ax} can define and prove a lot of concepts and sentences of its metatheory. In order to do that two tasks must be carried out:

(1) The "translations" of metamathematical expressions in (primitive) recursive functions and relations.

(2) The representability and expressibility of these functions and relations, respectively, in PA^{ax} (*via* Sect. 3.4. Theorem and final Remark).

The item (1) can be achieved by *arithmetization* of the syntax of PA^{ax}.

The arithmetization of PA^{ax} is a 1-1 function *g* from the set of symbols, expressions and finite sequences of expressions of L_{PA} into the set of positive integers. The number associated to an expression *e* (consisting in a single symbol or a finite sequence of symbols) will be called its *Gödel number* (code). In what follows by $g(e)=n$ we mean: *the Gödel number of e is n*.

There are many ways to *arithmetize* the syntax of PA^{ax}.[22] Let us give some examples, beginning with the original one, i.e., Gödel's proposal.

---

[21] In what follows we take "the metatheory of PA", "the metamathematics" and "the syntax of PA" as synonyms.

[22] Comp. *inter alia*, K. Gödel [1931] Sect 2, D. Hilbert and P. Bernays [1939] §4, W.V.O. Quine [1940], S.C. Kleene [1952] Ch. X, H. Scholz and G. Hasenjaeger [1961], §§230-238, E. Mendelson [1964] Ch. 3, Sect. 4, R.M. Smullyan [1961] Ch. I, §6, [1992] Ch. I, II, G. Boolos [1993] Ch. 2, S. Kripke [1996] Lect. IV, G. Boolos, J. Burgess and R. Jeffrey [2002] Ch. 15.

Gödel's strategy[23] of coding expressions is the following. With each primitive symbol, 0, f, ¬, ∨, ∀, (,), is associated an odd number: 1, 3, 5, 7, 9, 11, 13, respectively. To every variable of the type *n* is associated a number >13 of the form p^n. In this way to every sequence of symbols (whether primitive symbols or formulas) is associated a sequence of numbers, $n_1,...,n_k$. Then to such a sequence of numbers is, finally, associated a number $n = 2^{n_1} \cdot 3^{n_2} \cdot ... \cdot p_k^{n_k}$, where $2,3,...,p_k$ are the first *k* prime numbers in order of magnitude. The number *n* is the Gödel number of the respective sequence of symbols. Briefly, the "translation" of syntactic (metamathematical) concepts in arithmetical expressions will be given in the following way. With $\mathrm{R}(\mathrm{a}_1,...,\mathrm{a}_\mathrm{n})$, a relation between the syntactic objects $\mathrm{a}_1,...,\mathrm{a}_\mathrm{n}$ (primitive symbols or sequences of primitive symbols), is associated a number theoretic relation $R^*(x_1,...,x_n)$ that holds between the numbers $x_1,...,x_n$ if and only if there exist $\mathrm{a}_1,...,\mathrm{a}_\mathrm{n}$ such that $g(\mathrm{a}_1) = x_1$,..., $g(\mathrm{a}_\mathrm{n}) = x_n$ and $\mathrm{R}(\mathrm{a}_1,...,\mathrm{a}_\mathrm{n})$ holds; where $g(\mathrm{a}_\mathrm{i})$, $1 \le i \le n$, is the Gödel number of the syntactic object a_i. By this procedure, the syntactic entities ("formula", "axiom", "proof", "provable formula" etc.) turn into arithmetical expressions.

Some other way to arithmetize the syntax of $\mathrm{PA}^{\mathrm{ax}}$ is based on establishing a 1-1 correspondence between the ordered *pairs* of natural numbers and the set of natural numbers, using recursive functions. A simple way to do this is to use a recursive pairing function $J(x, y)$ with the following property: if $J(x, y) = n$, then (x, y) is the *n*th pair in the enumeration. Let us detail.

Consider firstly that we arrange the ordered pairs in the following way (Cantor's enumeration): we take firstly all the pairs (x, y) whose sum is 0. There is only one such pair, (0,0). Then follow all the pairs whose sum is 1 and we take them in the order (0,1) and (1,0). We continue with all ordered pairs whose sum is 3, i.e., (0,3), (1,2), (2,1), (3,0) and so on. Now, if we take the function

$$J(x, y) = \frac{1}{2}[(x + y)^2 + 3x + y],$$

[23] Cf. K. Gödel, [1931] (cited apud K. Gödel, [1986], 157).

we observe that this function do the job. For each ordered pair of natural numbers (x, y), $J(x, y)$ is a single natural number *z*; and, conversely, for every natural number *z* there exists strictly one ordered pair of numbers (x, y) such that $z = J(x, y)$.[24]

For this function $J(x, y) = z$ there exist the inverse recursive functions $K(z)$ and $L(z)$ such that: $K(z) = x$, $L(z) = y$; i.e., $x = K(J(x, y))$ and $y = L(J(x, y))$. And therefore $J(K(z), L(z)) = z$. So, $J(x, y)$ and its inverse functions gives us a method to generate the ordered pairs of natural numbers. This result can be extended to all finite sequences of natural numbers, by defining the function $J_n(x_1, ..., x_n)$ $(n \geq 2)$ in the following way:

$$J_2(x_1, x_2) = J(x_1, x_2)$$
$$J_3(x_1, x_2, x_3) = J(x_1, J(x_2, x_3))$$
$$J_4(x_1, x_2, x_3, x_4) = J(x_1, J_3(x_2, x_3, x_4))$$
$$\vdots$$
$$J_{n+1}(x_1, ..., x_n, x_{n+1}) = J(x_1, J_n(x_1, ..., x_n))$$

By induction we may prove that for each $n \geq 2$, $J_n(x_1, ..., x_n)$ is a 1-1 recursive function with its range the set of natural numbers.

Now, if we apply this method to a finite sequence of numbers $x_1, ..., x_m$, then a number *z* is obtained, from which the sequence $x_1, ..., x_m$ can be regained in the following way:

(1) We set up $L^1(z) = z$, $L^{n+1}(z) = L(L^n(z))$. Then for $i < m$, $L^i(z) = J(x_i, ...)$, but $L^m(z) = x_m$.

(2) We define $f(z, i) = K(L^i(z))$, and in this way we have for $i < m$, $x_i = f(z, i)$, $x_m = L^m(z)$.

This method sets up the coding of expressions of the language L_{PA}.[25]

A similar method of coding is used by Boolos[26], by assigning to the eight primitive symbols, ⊥, ⊃, ∀, =, 0, s, + and × the *odd* numbers 1, 3, 5, 7, 9, 11, 13 and 15, respectively, to the variable v_i the number $2i + 17$, and to the ordered pair (x, y), where *x* and *y* are arbitrary objects (symbols or

---

[24] Details of proof, cf. M. Davis, [1958], 44.

[25] For details, comp. Scholz/Hasenjaeger, [1961], §233.

[26] G. Boolos, [1993], Ch. 2; comp. also C. Smorynski [1985], Ch. 0, §5.

ordered pairs), the *even* number $\pi(x,y)=z[(i+j)^2+i+1]$. If an expression is an ordered triple (x,y,z) then, as above, it can be read as $(x,(y,z))$, i.e., again as an ordered pair.

The Gödel number of a formula excepting $\perp$ is a number of the form $\pi(i,\pi(a,b))$[27], where *i* is odd and $\pi(a,b)$ is even. In order to distinguish the Gödel number of a proof from the other Gödel numbers, to such a proof is assigned a number of the form $\pi(\pi(a,b),k)$.

This method allows us to define all the main concepts of the syntax. Let us take two key concepts.

$Pf(y,x)$: *y* is (the Gödel number of) a proof for (the formula with Gödel number) *x*.

Usually, in metamathematical terms, $Pf(y,x)$ is: *y* is a finite sequence of formulas, whose last formula is *x*, and such that every formula of the sequence is either an axiom or follows from the two preceding formulas by *modus ponens*, or it follows from a preceding formula by application of Generalization Rule.

Now, arithmetically expressed, this is the primitive recursive relation

$$\begin{aligned} Pf(y,x):\quad & \mathrm{FinSeq}(y) \wedge s_{lh(y)-1} = x \wedge \forall i < lh(y)-1[Ax(y_i) \\ & \vee \exists j < i\ \exists k < i\ \ \mathrm{ConseqByModPon}(y_i, y_j, y_k) \vee \\ & \vee \exists j < i\ \ \mathrm{ConseqByGen}(y_i, y_j)], \end{aligned}$$

where "FinSeq(*y*)" means *y* is a finite sequence, "$s_{lh(y)-1}=x$" means the last term of the sequence is *x*, "$\mathrm{Ax}(y_i)$" means y_i is an axiom, "$\mathrm{ConseqByModPon}(y_i,y_j,y_k)$" means y_i follows by *modus ponens* from y_j and y_k, and "$\mathrm{ConseqByGen}(y_i,y_j)$" means the formula y_j follows by Generalization Rule from the formula y_i.

Bew(*x*) is the semi-recursive[28] relation (predicate) of provability in PA^{ax}. Its meaning is: "*x* is provable in PA^{ax}". Its arithmetical expression is, simply, $\exists y Pf(y,x)$.

A strategy of coding, similar to that of Gödel's, is Mendelson's.[29] Given its elegant form we take it, with minor changes, as basis of our

---

[27] Since the formulas are represented, for example, as $=(t_1,t_2)$, $\supset(F_1,F_2)$ etc., where the first symbol is a primitive one.

[28] Or recursively enumerable; comp. Sect. 4.2.5.

[29] Cf. E. Mendelson, [1964], Ch. 3, Sect.4.

considerations.

The symbols of PA^{ax} (first row) are assigned the following Gödel numbers (second row):

(	)	,	$\neg$	$\supset$	$\forall$	x_i	a_i	f_i^n	P_i^n
3	5	7	9	11	13	$7+8i$	$9+8i$	$11+8(2^n \cdot 3^i)$	$13+8(2^n \cdot 3^i)$

for $i, n \geq 1$.

As can be seen, the Gödel numbers of the symbols are odd numbers ≥ 3, and for different symbols their Gödel numbers are different.

Example. $g(x_3) = 7 + 8 \cdot 3 = 31$, $g(a_3) = 9 + 8 \cdot 3 = 33$,

$$g(f_2^1) = 11 + 8(2^1 \cdot 3^2) = 155, \ g(P_1^2) = 13 + 8(2^2 \cdot 3^1) = 109.$$

Once the symbols are given the respective Gödel numbers, the Gödel number of an expression (as a finite sequence of symbols) and of a finite sequence of expressions can be determined in the following way.

If $s_1 s_2 ... s_k$ is an expression consisting of the respective symbols $s_1, ..., s_k$, then its Gödel number will be $2^{g(s_1)} \cdot 3^{g(s_2)} \cdot ... \cdot p_{k-1}^{g(s_k)}$, where $2, 3, ..., p_{k-1}$ are the first *k* prime numbers with $p_0 = 2$.

Example. $g(\forall x_1 P_1^1(x_1)) = 2^{13} \cdot 3^{15} \cdot 5^{61} \cdot 7^3 \cdot 11^{15} \cdot 13^5$

As we know, the factorization of an integer into primes is unique, and then if e_1 and e_2 are different expressions, their Gödel numbers will be different. Moreover, the symbols and the expressions have different Gödel numbers, due to the fact that Gödel numbers of symbols are odd, but that of an expression is even. As can be observed, the Gödel number of a symbol, e.g. $g(\supset) = 11$, is different from the Gödel number of the expression consisting only of the respective symbol, i.e., 2^{11}.

Finally, if $e_1, e_2, ..., e_k$ is a sequence of expressions, then this sequence will have the following Gödel number: $2^{g(e_1)} \cdot 3^{g(e_2)} \cdot ... \cdot p_{k-1}^{g(e_k)}$. If $e_i \neq e_j$ then, as can be argued, $g(e_i) \neq g(e_j)$. And since the Gödel number of a sequence of expressions is even, it is different from the Gödel number of a symbol, and since in its factorization the first prime 2 has an even exponent, it is different from the Gödel number of an expression.

Using this method of coding essential parts of the syntax of PA^{ax} can be arithmetized; i.e., the metamathematical assertions can be replaced with number-theoretic statements.[30]

Let us take some examples of such a "translation" of the syntax in arithmetical expressions.[31] As can be argued, all these functions and relations are (primitive) recursive.[32]

1. *EVbl*(x); x is the Gödel number of an expression consisting of a variable.

$(Ez)_{z<x}(1 \leq z \wedge x = 2^{7+8z})$

By 3.1 and 3.2 this relation is primitive recursive.

2. $Arg_P(x) = (qt(x \dot{-} 13,8))_0$: If x is the Gödel number of a predicate symbol P_i^n, then $Arg_P(x) = n$.

For example, if the predicate symbol is P_1^2, then

$g(\mathrm{P}_1^2) = 13 + 8(2^2 \cdot 3^1) = 109$.

Then $109 \dot{-} 13 = 96$,

$(qt(96,8))_0 = (12)_0 = (2^2 \cdot 3^1)_0 = 2$.

3. *Gen*(*x,y*): The expression with Gödel number y comes from the expression with Gödel number x by *Gen*.

$(Ev)_{v<y}(EVbl(v) \wedge y = 2^3 * 2^{13} * v * x * 2^5)$,

where "$*$" is the concatenation operation (cf. Sect. 3.2, Example 4).

4. $Ax_1(x)$: x is the Gödel number of an instance of Ax1.

Assuming that *Fml*(x) is the number-theoretic expression of "x is a formula of $\mathrm{L_{PA}}$" and that this relation is (primitive) recursive,[33] the number-theoretic "translation" of $\mathrm{Ax}_1(x)$ is

$(Eu)_{u<x}(Ev)_{v<x}(Fml(u) \wedge Fml(v) \wedge x = 2^3 * u * 2^{11} * 2^3 * v * 2^{11} * u * 2^5 * 2^5$,

i.e., the number-theoretic expression of an instance of Ax1: $\alpha \supset (\beta \supset \alpha)$.

---

[30] An approach by which the metatheory of PA^{ax} becomes a branch of the arithmetic of natural numbers.

[31] For the full list, comp. Mendelson, [1964], Ch. 3, Sect. 4.

[32] They remain (primitive) recursive for any first order theory for which the sets of Gödel numbers of its symbols (symbols for predicates/ functions/ constants) and the set of Gödel numbers of its *proper axioms* are (primitive) recursive.

[33] Since the (primitive) recursiveness of *Fml*(x) requires the (primitive) recursiveness of the relations: "x is the Gödel number of a predicate symbol", "x is the Gödel number of a function symbol" and "x is the Gödel number of a constant symbol", in the definition 4 the (primitive) recursiveness of these three relations is assumed. And their (primitive) recursiveness requires that the respective sets of symbols be *finite* (as is the case with PA^{ax}).

5. $Num(y)$ = the Gödel number of $\bar{y}$.
(I.e., the Gödel number of the numeral corresponding to the number y).

$$Num(0) = 2^{17}$$

$$Num(y+1) = 2^{59} * 2^{3} * Num(y) * 2^{5}$$

6. $Sub(y,u,v)$ = the Gödel number of the expression resulting from the expression with the Gödel number y by substituting the term with Gödel number u for all free occurrences of the variable with Gödel number v.

$$Sub(y,u,v) = \mu x_{x<uy^{lh(y)}} Subst(x,y,u,v),$$

where $Subst(x,y,u,v)$ is the primitive recursive relation: x is the Gödel number of the expression obtained from the expression with Gödel number y by substituting the term with Gödel number u for all free occurrences of the variable with Gödel number v.

7. $Pf(y,x)$: y is the Gödel number of a proof of the formula with Gödel number x. $Pf(y,x) = Prf(y) \wedge x = (y)_{lh(y) \dot{-} 1}$,
where $Prf(y)$ is the relation "y is the Gödel number of a proof in PA^{ax}", assumed to be (primitive) recursive.

8. Let $\alpha(x_1)$ be a formula of L_{PA}, containing only x_1 free. Let m be its Gödel number. Let $Bw_\alpha(x,y)$ be defined as follows: "y is the Gödel number of a proof in PA^{ax} of the formula $\alpha(\bar{x})$".[34] The arithmetical expression of this metamathematical predicate is:

$$Pf(y, Sub(m, Num(x), 2^{15})).$$

If $\alpha(x_1,x_2)$ is a given formula of L_{PA}, with x_1 and x_2 its only free variables and m is its Gödel number, then $Bw_\alpha(x_1,x_2,y)$ is the metamathematical predicate: "y is a proof in PA^{ax} of the formula $\alpha(\bar{x}_1,\bar{x}_2)$. Its arithmetical counterpart will be:

$$Pf(y, Sub(Sub(m, Num(x_1), 2^{15}), Num(x_2), 2^{23})$$

(similarly for $\alpha(x_1,...,x_n)$).

In a similar fashion, many other metamathematical assertions can be "translated" into the number-theoretic expressions.

9. *Diagonal function*: If n is the Gödel number of a formula $\alpha(x_1)$, with x_1 free, then $\delta(n)$ is the Gödel number of the formula $\alpha(\bar{n})$ ($\alpha(\bar{n})$ is called the *diagonal*(*ization*) of $\alpha(x_1)$).[35]

[34] More on this predicate, cf. S.C. Kleene [1952], §60.
[35] More on the diagonalization: next section (4.2.1.1).

$\delta(n) = Sub(n, Num(n), 2^{15})$

10. $R_1(n, y)$: n is the Gödel number of the formula $\alpha(\mathrm{x}_1)$, with x_1 free, and y is the Gödel number of a proof of its diagonalization, i.e., $\alpha(\overline{n})$.

Assuming that $Fr(n,x)$ (n is the Gödel number of a formula containing the variable with Gödel number x free), $Fml(n)$ (n is the Gödel number of a formula of L_{PA}) are primitive recursive, using $Sub(y,u,v)$ $R_1(n, y)$ can be rendered by the following expression:

$Fml(n) \wedge Fr(n, 2^{15}) \wedge Pf(y, Sub(n, Num(n), 2^{15}))$.

11. $R_2(n, y)$: n is the Gödel number of the formula $\alpha(\mathrm{x}_1)$, with x_1 free, and y is the Gödel number of a proof of $\neg\alpha(\overline{n})$.

By arithmetization, this metamathematical expression becomes:

$Fml(n) \wedge Fr(n, 2^{15}) \wedge Pf(y, Sub(2^3 * 2^9 * n * 2^5, Num(n), 2^{15}))$.

Two equivalences

Theorem. *Every function* $f(x_1,...,x_n)$, *representable in* $\mathrm{PA}^{\mathrm{ax}}$, *is recursive*.

Proof. By Def. 4 of Sect. 2, if $f(x_1,...,x_n)$ is representable in $\mathrm{PA}^{\mathrm{ax}}$, then there is a formula $\alpha(\mathrm{x}_1,...,\mathrm{x}_n,\mathrm{y})$ of L_{PA} such that for any numbers $x_1,...,x_n$

If $f(x_1,...,x_n) = y$, then $\mathrm{PA}^{\mathrm{ax}} \vdash \alpha(\overline{x}_1,...,\overline{x}_n, \overline{y})$.

The gist of the proof is to consider a number z whose first factors (in its factorization) are 2^{z_0} and 3^{z_1}, where $z_0 = (z)_0$ is the number y (i.e. the value of f for arguments $x_1,...,x_n$) and $z_1 = (z)_1$ is the Gödel number of a proof of the formula $\alpha(\overline{x}_1,...,\overline{x}_n, \overline{y})$. Then $Bw_\alpha(x_1,...,x_n,(z)_0,(z)_1)$ holds, and therefore $f(x_1,..,x_n) = (\mu z(Bw_\alpha(x_1,...,x_n,(z)_0,(z)_1))_0$. Now, since $(x)_i$ is primitive recursive (cf. Sect. 3.2, Example 2) and Bw_α is primitive recursive (cf. Sect. 4.1, Example 8), it follows (*via* Sect. 3.1, μ-Operator) that $f(x_1,...,x_n)$ is a recursive function.

By Theorem of Sect. 3.4 every recursive function is representable in $\mathrm{PA}^{\mathrm{ax}}$, and by the theorem above the converse is also the case. Hence, the first metalinguistic equivalence holds:

Eq1. *A function* $f(x_1,...,x_n)$ *is recursive if and only if it is representable in* $\mathrm{PA}^{\mathrm{ax}}$.

On the other hand, by Def. of Sect. 3.2, a relation $R(x_1,...,x_n)$ is recursive if and only if its characteristic function $C_R(x_1,...,x_n)$ is recursive. By Eq1, C_R is recursive if and only if C_R is representable in PA^{ax}. And, finally, by Sect. 2 Theorem, $R(x_1,...,x_n)$ is expressible in PA^{ax} if and only if $C_R(x_1,...,x_n)$ is representable in PA^{ax}. Hence, the second metalinguistic equivalence also holds:

Eq2. *A relation* $R(x_1,...,x_n)$ *is recursive if and only if it is expressible in* PA^{ax}.

4.2. Gödel's Theorems for PA^{ax}

The construction of the celebrated Gödel's undecidable sentence *G* and the proof of his theorems can be carried out in a variety of ways. In what follows we analyze this topic as a result of applying self-reference in the following forms:

1. Diagonalization (without Diagonal Lemma) (4.2.1)
2. Diagonalization (*via* Diagonal Lemma) (4.2.2)
3. Semi-recursivity (*via* Paradoxes) (4.2.3).

In all these cases the use of *diagonal arguments* plays a key role.

4.2.1. Gödel's Theorems for PA^{ax} (*via* diagonalization)

4.2.1.1. Diagonalization

Let us take the sentence "This sentence has five words". As can be seen, it contains the indexical word "this", indicating the fact that the sentence refers to itself. Hence the sentence is self-referential, since it ascribes to itself the property of having five words, and then it is true.

Now, if we take the sentence "This sentence is false"[36], again, the indexical "this" shows that it is self-referential. But, in contrast to the previous example, this sentence cannot be true or false (argue!).

But if we do not want to use the indexicals in constructing self-reference, we can proceed by *diagonalization*. Intuitively, the diagonalization (or diagonal) of an expression containing *x* free is the expression resulting by substituting the quotation (i.e., the name) of that expression for every free occurrence of *x* in the expression itself. As an example, let us take the following expression:

[36] It is a version of the well-known Liar-paradox.

Exp. *It is false the expression asserting the diagonalization of x*.

Containing *x* free, *Exp* is not a sentence, hence cannot be true or false. Let us construct its diagonalizaton, i.e.,

Diag. *It is false the expression asserting the diagonalization of* "*It is false the expression asserting the diagonalization of x*".

As can be seen, this is a sentence asserting its own falsity, hence it is self-referential.

Let us now couch this idea of diagonalization in formal terms, since the theory we are dealing with in what follows is the formal system PA^{ax}.

Definition. *Let* $\alpha(x)$ *be a formula of* L_{PA} *containing* x *free, let n be its Gödel number and* $\overline{n}$ *the corresponding numeral. Then the diagonalization* (*or diagonal*) *of* $\alpha(x)$ *is the formula obtained from* $\alpha(x)$ *by substituting the numeral* $\overline{n}$ *for all free occurrences of* x *in* $\alpha(x)$; *i.e., it is the formula* $\alpha(\overline{n})$.

Of course, diagonalization involves the substitution, an operation difficult to arithmetize. This is why the diagonal of the formula $\alpha(x)$ is sometimes defined as being the formula $\exists x(x=\overline{n}\wedge\alpha(x))$ or the formula $\forall x(x=\overline{n}\supset\alpha(x))$. And this fact is perfectly licit, since the equivalences $\alpha(\overline{n})\equiv\exists x(x=\overline{n}\wedge\alpha(x))$ and $\alpha(\overline{n})\equiv\forall x(x=\overline{n}\supset\alpha(x))$ are theorems of FOL^{ax}_{id}, and therefore theorems of PA^{ax} (comp. Ch. 2, 4.3, Lemma).[37]

4.2.1.2. Gödel's Theorems for PA^{ax}

The formal system *P* for which Gödel states and proves his theorems "is essentially the system obtained when the logic of PM [Principia Mathematica] is superposed upon the Peano axioms [...]".[38] In what follows we only consider the system PA^{ax}, for which the existence of undecidable sentences can also be proved.[39] To begin with, let us display the Gödel's view of the existence of an undecidable sentence in its intuitive form, and then in the formalized one.

(1) **Gödel's intuitive argument**

As we saw in Sect. 4.1, using the Gödel numbering a lot of metamathematical expressions applied to syntactical objects (such as

---

[37] Even though $\alpha(\overline{n})$, $\exists x(x=\overline{n}\wedge\alpha(x))$ and $\forall x(x=\overline{n}\supset\alpha(x))$ are different formulas of L_{PA}, and then they have different Gödel numbers.

[38] K. Gödel, [1986], 151.

[39] Cf. K. Gödel [1931], Theorem VI, Theorem VIII.

"formula", "proof", "free variable", "relation sign") become *arithmetical* expressions applied to their Gödel numbers. In a second step these last expressions are rendered (expressed/ defined)[40] in the formal language of PA. In this way we can, for example, construct a formula R(x), with exactly one free variable x (in Gödel's terms a *class sign*), whose meaning is "*x* is a provable formula".

Let $R_0(x), R_1(x), R_2(x), ...$ be an enumeration of all class signs. $R_i(x)$ (in Gödel's notation $R(i)$) is the *i*th term in this enumeration. By $R_i(\overline{n})$ (in Gödel's notation [R;n]) we understand the formula resulting from the formula $R_i(x)$ by substitution of the numeral of *n* for the free variable x; i.e., $R_i(\overline{n})$. The relation "*x* is the Gödel number for such formula" (in Gödel's terms $x = [y; z]$) is also definable in PA.

The notions grounding the construction of the Gödel's argument are the following:

1. The formal system PA^{ax} is *correct*;[41] i.e., for any formula α the following holds:

If PA $\vdash \alpha$, then α is true (in *M*[42]).

2. *Definability* in L_{PA} of a numerical set *K*.

Definition. *A number-theoretic set K*[43] *is definable in* L_{PA}, *iff there is a formula of* L_{PA} α(x) (*in Gödel's terms a class sign*) *such that for any number x holds*:

$x \in K$ *iff* $\alpha(\overline{x})$ *is true* (*in M*),

where $\alpha(\overline{x})$ *is obtained from* α(x) *by substitution of the numeral of x*, *i.e.*, $\overline{x}$, *for* x.

[40] In this informal argument Gödel uses "ausdrückbar" (expressible) and "definierbar" (definable) as synonyms (in the sense given in 2 below).

[41] The terms "correct" and "correctness" are synonyms with "sound" and "soundness", respectively (as we used them in Chs. 1 and 2). In the celebrated paper [1931] Gödel used this meaning of soundness in his intuitive argument (without any mention of the terms "korrekt" or "Korrektheit"): "...suppose that the proposition [R(q);q] were provable; then it would also be true". [K. Gödel [1986], 148, 150. Some other authors use the term "correct" (for sound), e.g. S.C. Kleene [1952], 302; R.M. Smullyan [1992], 6; H. Hermes [1963], 141 ("Korrektheit").

[42] Where *M* is the standard model of L_{PA}.

[43] "$x \in K$" can be expressed equivalently in the form of the 1-place numerical relation *K*(*x*).

3. *Definition.* $x \in K =_{\mathrm{df}} \sim \mathrm{Bew}R_x(\bar{x})$,[44]
whose meaning is the following: "x is a number in K iff the formula obtained from the xth class sign in the above enumeration, i.e., $R_x(\mathrm{x})$, is not provable for the argument x".

Now, since all the notions in the expression "$\sim \mathrm{Bew}R_x(\bar{x})$" are definable in L$_{\mathrm{PA}}$, it follows that the relation $x \in K$ is definable in L$_{\mathrm{PA}}$ by a class sign. Let it be qth class sign in the above enumeration, i.e., $\mathrm{R}_q(\mathrm{x})$ defines $x \in K$ in L$_{\mathrm{PA}}$. And then we have the following equivalence:

(***Eq***) For *any x*: $x \in K$ iff $\sim \mathrm{BewR}_x(\bar{x})$ (by 3)

iff $\mathrm{R}_q(\bar{x})$ is true in *M* (by 2).

Let $x = q$ (diagonalization!). By (***Eq***) we deduce:

(***Eq****) $q \in K$ iff $\sim \mathrm{BewR}_q(\bar{q})$ iff $\mathrm{R}_q(\bar{q})$ is true in *M*.

As can be seen, since $\mathrm{R}_q(\mathrm{x})$ defines *K* in L$_{\mathrm{PA}}$, the sentence $\mathrm{R}_q(\bar{q})$ says that $q \in K$ (by 2), and therefore that it is not provable (by 3). Hence $\mathrm{R}_q(\bar{q})$ is self-referential, asserting about itself that it is not provable in $\mathrm{PA}^{\mathrm{ax}}$.

Theorem. $\mathrm{R}_q(\bar{q})$ *is undecidable in* $\mathrm{PA}^{\mathrm{ax}}$.

(a) $\mathrm{R}_q(\bar{q})$ *is not provable in* $\mathrm{PA}^{\mathrm{ax}}$.

(*Reductio*). Suppose that $\mathrm{R}_q(\bar{q})$ is provable in $\mathrm{PA}^{\mathrm{ax}}$. It follows that $\mathrm{R}_q(\bar{q})$ is true (in *M*) (by 1). And then $q \in K$ (by 2). Hence $\sim\mathrm{Bew}\,\mathrm{R}_q(\bar{q})$ holds (by 3); i.e., $\mathrm{R}_q(\bar{q})$ is not provable in $\mathrm{PA}^{\mathrm{ax}}$. Therefore, $\mathrm{R}_q(\bar{q})$ is not provable in $\mathrm{PA}^{\mathrm{ax}}$.[45]

(b) $\neg\mathrm{R}_q(\bar{q})$ is not provable in $\mathrm{PA}^{\mathrm{ax}}$.

(*Reductio*). Suppose that $\neg\mathrm{R}_q(\bar{q})$ is provable in $\mathrm{PA}^{\mathrm{ax}}$. It follows that $\neg\mathrm{R}_q(\bar{q})$ is true in *M* (by 1). And then $\mathrm{R}_q(\bar{q})$ is false in *M*; and this implies that $q \notin K$ (by 2); equivalently $\mathrm{Bew}\,\mathrm{R}_q(\bar{q})$ (by 3), contradicting 1.

By (a) and (b) it follows that $\mathrm{R}_q(\bar{q})$ is undecidable in $\mathrm{PA}^{\mathrm{ax}}$. Finally, since by (a) $\mathrm{R}_q(\bar{q})$ is not provable in $\mathrm{PA}^{\mathrm{ax}}$, and since the meaning of $\mathrm{R}_q(\bar{q})$

---

44 In Gödel's notation: $x \in K \equiv \overline{\mathrm{Bew}}[\mathrm{R}(x);x]$, cf. 149; "Bew" is for the German "beweisbar" (provable).

45 A shorter form of this argument can be given *via* (***Eq****) plus 1 (the correctness of $\mathrm{PA}^{\mathrm{ax}}$).

is "$\mathrm{R}_q(\overline{q})$ is not provable in PA^{ax}" it follows (metamathematically) that $\mathrm{R}_q(\overline{q})$ is *true*. Hence $\mathrm{R}_q(\overline{q})$ is an example of a true sentence of L_{PA} but not provable in PA^{ax}.

Remark. A similar strategy of constructing intuitively an undecidable sentence is given by Smullyan. [46] In his *abstract* form of Gödel's Theorem the index *n* of a formula R(x), i.e., $\mathrm{R}_\mathrm{n}(\mathrm{x})$, in an enumeration is taken to be its Gödel number. Then for $\mathrm{R}_\mathrm{n}(\mathrm{x})$ its diagonalization will be $\mathrm{R}_n(\overline{n})$ and will have the Gödel number $\delta(n)$ (where $\delta(x)$ is the diagonal function; cf. Sect. 4.1, Example 9). The symbol *P* denotes the set of Gödel numbers of *provable* formulas,[47] and $\widetilde{P}$ will be its complement.

The three ingredients of Smullyan's account are the same as in Gödel's intuitive argument, i.e.,

1^*. *L* is correct.

2^*. Definability[48] of a set *K* in *L* by a formula R(x):

$x \in K$ iff $\mathrm{R}(\overline{x})$ is true in *L*.

3^*. *Definition.* $x \in K$ *iff* $\delta(x) \in \widetilde{P}$; *i.e.,* $K = \{x \mid \delta(x) \in \widetilde{P}\}$ [49].

The Smullyan's argument proceeds as follows. Let *q* be the Gödel number of R(x) defining *K* in *L*, i.e., $\mathrm{R(x)} = \mathrm{R}_q(\mathrm{x})$. And then for *all x*:

(***Equiv***) $x \in K$ iff $\delta(x) \in \widetilde{P}$ (by 3^*) iff $\mathrm{R}_q(\overline{x})$ is true (by 2^*).

As can be observed, $\delta(x) \in \widetilde{P}$ means "the formula with Gödel number $\delta(x)$, i.e., $\mathrm{R}_x(\overline{x})$, is not provable".

Let $x = q$ (diagonalization!) From (***Equiv***) we derive

(***Equiv****) $q \in K$ iff $\delta(q) \in \widetilde{P}$ iff $\mathrm{R}_q(\overline{q})$ is true in *L*.

Since $\delta(q)$ is the Gödel number of $\mathrm{R}_q(\overline{q})$, it follows that, by (***Equiv****), we derive:

$\mathrm{R}_q(\overline{q})$ is not provable iff $\mathrm{R}_q(\overline{q})$ is true in *L*.

And then we have the following alternatives:

---

[46] R. Smullyan, [1992], 7 Theorem (GT).

[47] In a language *L* (taken as abstract system).

[48] Smullyan uses "expresses" for "definable"; cf. [1992], 6. In what follows these notions will be taken as synonyms and therefore interchangeable.

[49] $K = \delta^{-1}(\widetilde{P})$; i.e., *K* is the inverse image of $\widetilde{P}$ under diagonal function $\delta(x)$.

(1) $\mathrm{R}_q(\overline{q})$ is not provable and true.

(2) $\mathrm{R}_q(\overline{q})$ is provable and not true.

As can be observed, (2) is excluded by 1^* (*L* is correct). Therefore $\mathrm{R}_q(\overline{q})$ is a true but undecidable sentence of *L*.

(2) **Gödel's Theorems for PA$^{\text{ax}}$**

As we saw in (1), Gödel's *intuitive* argument for the existence of an undecidable sentence of PA$^{\text{ax}}$ is based on the strong assumption of the *correctness* of PA. In the *formal* construction of it Gödel replaces this assumption "by a purely formal and much weaker one"[50], that of ω-*consistency*.

In order to display this argument we need some definitions.

Definition 1. *A formal system S is* ***inconsistent*** *iff there is a formula* α *of* L_S *such that* $S \vdash \alpha$ *and* $S \vdash \neg\alpha$; *otherwise it is* ***consistent***.

Definition 1*. *A formal system S is* ***inconsistent*** *iff S proves any formula; otherwise it is* ***consistent***.

The definitions 1 and 1^* define one and the same concept (argue!).

Definition 2. *A formal system S is* ω-***inconsistent*** *iff there is a formula* α(x) *of* L_S, *with* x *free variable, such that the following hold:*

(a) *For any n*: $S \vdash \neg\alpha(\overline{n})$; (*i.e.,* $\vdash \neg\alpha(0)$, $\vdash \neg\alpha(\overline{1})$, $\vdash \neg\alpha(\overline{2})$,...).

(b) $S \vdash \exists x \alpha(x)$;

otherwise S is ω-***consistent***.

Of course, it is the same thing if ω-inconsistency is defined by (a^*) for any *n*: $S \vdash \alpha(\overline{n})$, and ($b^*$) $S \vdash \neg\forall x\alpha(x)$.[51]

Exercises

1. Argue that the following holds: If PA$^{\text{ax}}$ is ω-consistent, then PA$^{\text{ax}}$ is consistent.

2. Let $\mathrm{PA}^* = \mathrm{PA}^{\text{ax}} \cup \{\neg G\}$. Is PA^* consistent, ω-consistent? (Argue!)

---

[50] Cf. K. Gödel, [1986], 151.

[51] I.e. all instances of α(x) are provable and the *negation* of its universal quantification is also provable.

The Gödelian sentence G

In his paper [1931], Gödel refers to the syntactic entities using their Gödel numbers. Instead of "the formula $\mathrm{R}(\mathrm{x}_1,...,\mathrm{x}_\mathrm{n})$ with just the free variables $(\mathrm{x}_1,...,\mathrm{x}_\mathrm{n})$" we find its equivalent in the following terms: "an *n*-place RELATION SIGN r (with the FREE VARIABLES $u_1,u_2,...,u_n$)".[52] Or, let $Sb(r^{u_1,...,u_n}_{Z(x_1),...,Z(x_n)})$ be the Gödel number of the formula resulting from the formula with Gödel number *r* by substitution of the numerals for the natural numbers $x_1,...,x_n$ for its free variables whose Gödel numbers are $u_1,...,u_n$, respectively; i.e., $Sb(r^{u_1,...,u_n}_{Z(x_1),...,Z(x_n)})$ is the Gödel number of $\mathrm{R}(\bar{x}_1,...,\bar{x}_n)$. An assertion of the form "$\mathrm{R}(\bar{x}_1,...,\bar{x}_n)$ is provable", in Gödel's terms, becomes $\mathrm{Bew}[Sb(r^{u_1,...,u_n}_{Z(x_1),...,Z(x_n)})]$, i.e., $Sb(r^{u_1,...,u_n}_{Z(x_1),...,Z(x_n)})$ is a PROVABLE FORMULA.

Gödel's construction begins with the definition of the following relation:

$Q(x,y) =_{df} \overline{xB_k[Sb(y^{19}_{Z(y)})]}$, where $xB_k y$ is the relation: "*x* is the Gödel number of a proof[53] of a formula with Gödel number *y*". Now, since all the notions in *definiens* are primitive recursive,[54] the relation $Q(x,y)$ is primitive recursive. And then, by his Theorem V, there is a RELATION SIGN *q* (with the FREE VARIABLES 17 and 19) such that

(a) $Q(x,y) \rightarrow \mathrm{Bew}_k[Sb(q^{17}_{Z(x)}\,{}^{19}_{Z(y)})]$

(b) $\overline{Q(x,y)} \rightarrow \mathrm{Bew}_k[\mathrm{Neg}(Sb(q^{17}_{Z(x)}\,{}^{19}_{Z(y)}))]$

where $\mathrm{Neg}(Sb(q^{17}_{Z(x)}\,{}^{19}_{Z(y)}))$ is the Gödel number of the negation of the formula with Gödel number $Sb(q^{17}_{Z(x)}\,{}^{19}_{Z(y)})$.

Let *p* = 17Gen *q*; i.e., *p* is the Gödel number of the formula obtained from the formula with Gödel number *q* by generalization with respect to the variable with Gödel number 17.

---

[52] Where *r* is Gödel number of $\mathrm{R}(\mathrm{x}_1,...,\mathrm{x}_\mathrm{n})$ and $u_1,...,u_n$ are the Gödel numbers of $\mathrm{x}_1,...,\mathrm{x}_\mathrm{n}$, respectively. Small capital letters in Gödel, [1986] and *italics* in Gödel [1931], points out this reference to a syntactic object using its Gödel number.

[53] A proof in an extension of Gödel's system *P* with a recursive class *k* of FORMULAS. Since for our argument the system is simply $\mathrm{PA}^{\mathrm{ax}}$, the class $k = \varnothing$ and $x\mathrm{B}_k y$ is simply the proof relation *Pf*(*x*,*y*), with the just mentioned meaning; comp. 7 in the list at the end of 4.1.

[54] What Gödel calls "rekursiv" is actually *primitive recursive*.

Now, if $r = Sb(q^{19}_{Z(p)})$, then $Sb(p^{19}_{Z(p)}) = 17\,Gen\,Sb(q^{19}_{Z(p)}) = 17\,Gen\ r$. 17 *Gen r* is the Gödel number of the famous Gödel's undecidable sentence G. By his Theorem VI, under some assumptions, the sentence 17 *Gen r* is undecidable in his system *P*.

Let us "translate" Gödel's construction in a more "visible" fashion, using directly syntactic objects instead of their Gödel numbers.

Since $Q(x,y)$ is primitive recursive, there is a formula Q(x,y) (with Gödel number *q*), containing just free variables x and y (with Gödel numbers 17 and 19, respectively), such that

(a*) $Q(x,y) \to \vdash \mathrm{Q}(\bar{x},\bar{y})$

(b*) $\sim Q(x,y) \to \vdash \neg\mathrm{Q}(\bar{x},\bar{y})$,

i.e., the formula Q(x,y) *formally expresses* in $\mathrm{PA}^{\mathrm{ax}}$ the primitive recursive relation $Q(x,y)$.

Let now ∀xQ(x,y) be the formula (with Gödel number *p* i.e., 17 *Gen q*) in which only y is free. If $\mathrm{Q}(\mathrm{x},\bar{p})$ is the formula (with Gödel number *r*), in which only x is free, then the formula $\forall\mathrm{xQ}(\mathrm{x},\bar{p})$ (with the Gödel number 17 *Gen r*) is just the Gödel undecidable sentence *G*.

As can be observed, the meaning of this formula is: "For all *x*, *x* is not a proof of the formula obtained from the formula with Gödel number *p* by substituting the numeral for *p* for the free variable y". But this formula is just the formula $\forall\mathrm{xQ}(\mathrm{x},\bar{p})$. So $\forall\mathrm{xQ}(\mathrm{x},\bar{p})$ is self-referential; it asserts simply: *I am not provable*. This formula is the diagonalization of the formula with Gödel number *p*.

Gödel's first incompleteness theorem. *If* $\mathrm{PA}^{\mathrm{ax}}$ *is* ω*-consistent, then* (1) G *is not provable in* PA, *and* (2) ¬G *not provable in* $\mathrm{PA}^{\mathrm{ax}}$.

Proof (1). G is not provable in $\mathrm{PA}^{\mathrm{ax}}$.

(*Reductio*). Suppose G is provable in $\mathrm{PA}^{\mathrm{ax}}$. Then there is a proof of it in $\mathrm{PA}^{\mathrm{ax}}$ with, say, Gödel number *k*. Hence $Q(k,p)$ is false. And then $\vdash \neg\mathrm{Q}(\bar{k},\bar{p})$ (by b*). Since $\vdash$G (by assumption), i.e., $\vdash \forall\mathrm{xQ}(\mathrm{x},\bar{p})$, it follows that $\vdash \mathrm{Q}(\bar{k},\bar{p})$ (by Ax4, MP). And then $\mathrm{PA}^{\mathrm{ax}}$ is inconsistent, and therefore ω-inconsistent, contradicting the hypothesis of the theorem.

(2). ¬G is not provable in $\mathrm{PA}^{\mathrm{ax}}$.

(*Reductio*). Suppose $\vdash\neg$G, i.e., $\vdash \neg\forall\mathrm{xQ}(\mathrm{x},\bar{p})$. By (1) G is not provable in $\mathrm{PA}^{\mathrm{ax}}$; hence for any number *x*, $Q(x,p)$ is true, and therefore for any *x* we have $\vdash Q(\bar{x}, \bar{p})$ (by (a^*)), which together with $\vdash \neg\forall xQ(x, \bar{p})$ destroy the assumed ω-consistency of PA^{ax}.

Remark 1. For the unprovability of G only the assumption of *simple* consistency is needed. The unprovability of ¬G, i.e., $\neg\forall xQ(x, \bar{p})$, can also be argued using Σ_1-Reflection[55]: If α is a Σ_1-sentence then if $\vdash \alpha$, then α is true. $\neg\forall xQ(x, \bar{p})$, equivalent $\exists x\neg Q(x, \bar{p})$ is a *false* Σ_1-sentence (since its negation, G, is true). Whence, by Σ_1-Reflection it is not provable in PA^{ax}.

Remark 2. The same result as to the construction of the sentence G and proving its undecidability in PA^{ax} can be obtained using the relation $R_1(n, y)$:[56] "*n* is the Gödel number of a formula $\alpha(x_1)$, with x_1 free, and *y* is the Gödel number of a proof of its diagonalization: $\alpha(\bar{n})$". Since it is primitive recursive, there is a formula $\beta(x_1,x_2)$ which formally expresses it in PA^{ax}. Let us consider the formula $\forall x_2\neg\beta(x_1,x_2)$ whose Gödel number is *k*. Let $G = \forall x_2\neg\beta(\bar{k}, x_2)$. As can be seen, the meaning of G is "the *diagonalization* of the formula with Gödel number *k*, i.e., G itself, is not provable". Hence G is asserting its own unprovability in PA^{ax}.

This was also the way Gödel constructs its undecidable sentence $G = \forall xQ(x, \bar{p})$. It is the diagonalization of the formula with Gödel number *p*: $\forall xQ(x,y)$ and according to the meaning of the relation $Q(x, y)$, G asserts its own unprovability.

Let us formulate and prove the theorem for this last version.

Gödel's first incompleteness theorem (version). (1) *If* PA^{ax} *is consistent, then* G *is not provable in* PA^{ax}; (2) *If* PA^{ax} *is* ω*-consistent, then* ¬G *is not provable in* PA^{ax}.

Proof (1). (*Reductio*). Assume hypothesis and suppose that G is provable in PA^{ax}. Then there is a proof of it with Gödel number, say, *m*. Hence

[55] Or Σ_1-Soundness. A Σ_1-formula is a formula of the form $\exists xF(x)$ with F(x) decidable. This does not mean that the unprovability of ¬G in PA^{ax} can be proved only on the hypothesis of *simple* consistency of PA^{ax}; comp. Ch. 4, Sect. 4.2.2 (the final *Remark*) and Sect. 4.2.5,**2** Fact 7 (*Lemma*) (below).

[56] The example 10 from the list at the end of 4.1. The 2-place relations $R_1(n,y)$ and $R_2(n,y)$ are used by Kleene [1952], §42, and also by Mendelson [1964], Ch. 3, Sect.5, in order to construct undecidable sentences (Gödel-type and Rosser-type, respectively). As can be observed, $R_1(n,y)$ originates in Gödel [1931].

$R_1(k,m)=1$,[57] and therefore $\vdash \beta(\bar{k},\bar{m})$. Now, since $\vdash$G, i.e., $\vdash \forall x_2 \neg\beta(\bar{k},x_2)$ it follows that $\vdash \neg\beta(\bar{k},\bar{m})$ (by Ax4 and MP), contra consistency of PA^{ax}.

(2) (*Reductio*). Assume hypothesis and $\vdash\neg$G, i.e., $\vdash \neg\forall x_2 \neg\beta(\bar{k},x_2)$, equivalent $\vdash \exists x_2\beta(\bar{k},x_2)$. But, by (1) G is not provable, i.e., for any number *m*, $R_1(k,m)=0$. And then $\vdash \neg\beta(\bar{k},\bar{m})$, for any *m*. This means that PA^{ax} is ω-inconsistent, contra hypothesis.

Remark. An even shorter proof of (1) can be given by deriving $\vdash \exists x_2\beta(\bar{k},x_2)$ directly from $\vdash \beta(\bar{k},\bar{m})$ (by Gen ∃), which together with the assumed provability of G, i.e., $\vdash \forall x_2 \neg\beta(\bar{k},x_2)$, does generate an inconsistency.

Gödel's second incompleteness theorem. *If* PA^{ax} *is consistent, then* $\text{Con}_{PA^{ax}}$ *is not provable in* PA^{ax} (*where* $\text{Con}_{PA^{ax}}$ *is the formula expressing in* L_{PA} *the consistency of* PA^{ax}).[58]

First of all, the part (1) of the first incompleteness theorem can be wholly expressed in L_{PA}. This conditional is

Cond. If PA^{ax} is consistent, then G is not provable in PA^{ax}.

Now, the antecedent of *Cond*, "PA^{ax} is consistent", can be rendered in L_{PA} in the following way. If Π(y,x) is the formula expressing in L_{PA} the proof relation $Pf(y,x)$ (cf. Sect. 4.1) and if "⊥" denotes a logical falsity (a contradiction), then the formula $\text{Con}_{PA^{ax}}$: $\forall y \neg\Pi(y,\text{"}\bot\text{"})$ will express in L_{PA} the consistency of PA^{ax}. Then, the consequent of *Cond*, "G is not provable in PA^{ax}", is just the sentence G (asserting self-referentially that it is not provable in PA^{ax}). And, finally, if for expressing the idea of "if..., then" we use the formal symbol of implication (⊃), then the formula expressing *Cond* is *Impl*: $\forall y\neg\Pi(y,\text{"}\bot\text{"})\supset G$,[59] i.e., $\text{Con}_{PA^{ax}} \supset G$.

Now, the proof of Gödel's second incompleteness theorem is

[57] As in the preceding chapters, we often use "1" and "0" for "true" and "false", respectively.

[58] Theorem XI of Gödel's paper [1931, 1986]. Actually, it is just a corollary of his first incompleteness theorem.

[59] This *Impl* is itself provable in PA^{ax}; its proof was given letter by D. Hilbert and P. Bernays [1939], 283-340. Moreover, the converse of *Impl* is also provable. For a proof of *Impl* and its converse, see Ch. 4, Sect. 4.2.2.

immediate from these considerations. Since if $\mathrm{Con}_{\mathrm{PA}^{ax}}$ were provable in PA^{ax}, then by *Impl* and *modus ponens* it follows that G would be provable in PA^{ax}, and then, by *Cond*, PA^{ax} would be inconsistent.

4.2.1.3. Gödel-Rosser Theorem for PA^{ax}

The undecidability of the Gödelian sentence G, as showed above, requires the hypothesis of ω-consistency of PA^{ax}. B. Rosser[60] constructed a more complex sentence R whose undecidability can be proved under the weaker hypothesis, that of simple consistency.

The sentence R can be obtained using the primitive recursive relations $R_1(n,y)$ and $R_2(n,y)$.[61] Now, if $\beta_1(x_1,x_2)$[62] and $\beta_2(x_1,x_2)$ are the corresponding formulas expressing them in PA^{ax}, the following formula can be constructed:

$$F(x_1):\ \forall x_2(\beta_1(x_1,x_2) \supset \exists x_3(x_3 \le x_2 \wedge \beta_2(x_1,x_3))$$

Let n be its Gödel number. Then the diagonalization of $F(x_1)$, $F(\overline{n})$, is the *Rosser sentence* R, i.e.,

$$R:\ \forall x_2(\beta_1(\overline{n},x_2) \supset \exists x_3(x_3 \le x_2 \wedge \beta_2(\overline{n},x_3)).$$

The meaning of R is: "To any proof of R there exists a proof of ¬R with an equal or smaller Gödel number. By Gödel-Rosser Theorem, if PA^{ax} is consistent, then R is undecidable in PA^{ax}.[63]

4.2.2. Gödel's Theorem [64] (*via* Diagonal Lemma)

As we saw above, the construction of an undecidable sentence G demands the technique of arithmetization and the idea of expressibility of recursive relations and that of representability of recursive functions in PA^{ax}. With these means the construction of an undecidable sentence can

---

[60] B. Rosser [1936].

[61] Comp. the list at the end of Sect. 4.1.

[62] $\beta_1(x_1,x_2)$ was used in the construction of the Gödelian sentence G: $\forall x_2 \neg\beta(\overline{k}, x_2)$; comp. Sect. 4.2.1.2(2) *Remark 2*.

[63] For the proof and details, comp. S.C. Kleene, [1952], §42 and E. Mendelson [1964], Ch. 3, Sect.5.

[64] In what follows by "Gödel's Theorem" we understand Gödel's *first* incompleteness theorem; as we already mentioned, Gödel's second incompleteness theorem is just a corollary of the first.

also be carried out using an ingenious result of mathematical logic called *Diagonal Lemma*.

Let us firstly give an *intuitive argument* for what is meant by such a lemma and then a proof of Gödel's first incompleteness theorem based on this argument.

(1) **Diagonal Lemma** (an intuitive view)[65]

The basic notions implied in this item are: Gödel numbering, expressibility in S, substitution function and diagonalization. Let us review them.

The idea of Gödel numbering is that explained above (Sect. 4.1). As Gödel says, by numbering (or coding) "[t]he metamathematical notions (propositions) thus become notions (propositions) about natural numbers..." and then they "can (at least in part) be expressed by the symbols of the system PM itself". And this means that a formula $F(v)$ can be found "such that $F(v)$ intepreted according to the meaning of the terms of PM, says: v is a provable formula".[66]

In short, as we saw in Sect. 4.1 above, this means:

(1) The "translation" of metamathematical expressions in number-theoretic functions and relations, and

(2) The *expressibility* (*definability*) of these number-theoretic expressions in the *formalism* of the system S.

Since the idea of "expressibility" is central in what we say below, let us define it.

Definition. *Let $R(x)$ be a number-theoretic relation. $R(x)$ is said to be expressible in S[67] if there is a formula of L_S such that for any natural number n holds:*

$R(n)$ holds iff $\alpha(\bar{n})$ is true (in M[68]).

This is the notion of expressibility.[69] But for our intuitive account it is

[65] This argument is inspired by Rosser's [1939], reprinted in M. Davis (ed.) [2004], 223-230.

[66] Cf. K. Gödel [1986], 147.

[67] Let us consider in what follows that $S \supseteq Q$ (i.e., S is an extension of the Robinson system Q).

[68] Where M is the standard model of PA.

[69] Often taken by the authors instead of "definability" (cf. first definition in 4.2.1.2). Do not confound it with the "*formal* expressibility" of a number-theoretical relation within PA^{ax} (cf. Sect. 2 above).

sufficient to consider a more general one: if E is an arbitrary expression, then E is expressible in S if S has a *formal* expression whose meaning is just E. This last definition allows us to speak of the expressibility in S not only of number theoretic expressions but also directly of the metamathematical ones, and then to pass from the metamathematical expression to the number-theoretic expressions and finally to formulas of S, and conversely.

Then, for example, if P symbolizes a *syntactic* property of a formula F_n, whose Gödel number is n, then we have:

(a) The formula F_n has the property P.

And if Q is the corresponding *numerical* (number-theoretical) property of the number n, then we have

(b) n has the property Q.

The expression (b) is got from (a) by arithmetization (*via* Gödel numbering). Finally, (b) is expressible in S by (c) a formula of L_S (i.e., of the language of S).

Now, let us take the (primitive) recursive function $Sub(x,y,v)$.[70] As we saw, $Sub(x,y,v)$ is the Gödel number of the formula resulting from the formula with the Gödel number x by substituting the term with Gödel number y for the variable with Gödel number v.

Example. Let $\alpha_x(x_1)$ be the formula with Gödel number x, that contains the free variable x_1 and that the Gödel number of x_1 is 2^{15}.[71] Let $\overline{w}$ be the numeral denoting the natural number w. Let y be the Gödel number of the term $\overline{w}$. Then $Sub(x,y,v) = Sub(x,y,2^{15})$ = the Gödel number of $\alpha(\overline{w})$. Let us write $Sb(x,y)$ for $Sub(x,y,2^{15})$.

As we saw (Sect. 4.2.1.1), by *the diagonalization* of a formula $\alpha(x)$, whose Gödel number is n, we understand the formula $\alpha(\overline{n})$, i.e., the formula got from $\alpha(x)$ by *substitution* of the numeral of n for the variable x.

Diagonal Lemma (DL). *Let S be a formal system. Suppose that "x has the property Q" is expressible in S. Then a formula* F *can be found such that* F *expresses* "F *has the property P*".

Proof. The proof of DL requests some results, proved by Gödel in his [1931] paper:

(a) A proof that $Sb(x,y)$ is recursive (comp. definition 31 of the list 1-46 in Sect. 2 of [1931]) (and above Sect. 4.1, item 6).

[70] This is the function $Sub(y,u,v)$ defined in the list at the end of Sect. 4.1.

[71] According to the Gödel numbering given in Sect. 4.1.

(b) A proof that $Sb(x,y) = z$ is expressible in S (cf. Theorem I, Sect. 2 of [1931]).

Now, since $Sb(x,y) = z$ is recursive, it is expressible in S (by Gödel's Theorem V). And since "x has the property Q" is expressible in S (by hypothesis), it follows that their conjunction is also expressible in S.

Evidently, since $Sb(x,y)$ is recursive, $Sb(x,x)$ is also recursive. And then "$Sb(x,x)$ has the property Q" is expressible in S. As can be observed, according to the meaning of $Sb(x,y)$, the number $Sb(x,x)$ is the Gödel number of the *diagonalization* of the formula with Gödel number x. Let $\delta(x)$ be $Sb(x,x)$. Since "$\delta(x)$ has the property Q" is expressible in S, let $G_n(x_1)$ be the formula containing x_1 free and whose Gödel number is n which expresses it in S. Let $G_n(\overline{n})$ be the diagonalization of $G_n(x_1)$. Its Gödel number is, clearly, $\delta(n)$. So, $G_n(\overline{n})$ expresses "$\delta(n)$ has the property Q". Let $F_{\delta(n)} = G_n(\overline{n})$. Then $F_{\delta(n)}$ expresses "$\delta(n)$ has the property Q". So, in general, under the above assumption, a formula F_n (with Gödel number n) can be found such that F_n expresses "F_n has the property P".

(2) **Gödel's Theorems**

Now, as we know (Sect. 4.1), $Pf(y,x)$ is the primitive recursive relation "y is a proof of x", expressible in S by a decidable formula $\Pi(y,x)$. According to Sect. 4.2.5, the formula $\exists y\Pi(y,x)$ is a Σ_1-formula expressing[72] the number-theoretic relation $(Ey)Pf(y,x)$. According to Sect. 4.2.5,1, Definition, if a relation is expressible (definable) by a Σ_1-formula, it is a Σ_1-relation.[73] In Gödel's terms $Pf(y,x)$ is the relation yBx (the item 45 of his list, p. 171) and $(Ey)Pf(y,x)$ is the relation $(Ey)yBx$, also called $Bew(x)$, whose meaning is "x is provable"[74] (the item 46 of the same list, p. 171).

As we saw (Sect. 4.2.1.2), a formal system S is ω-consistent iff there is no formula $\alpha(x)$ of L_S, with x free, such that the following hold

(a) For any n: $S \vdash \neg\alpha(\overline{n})$; (i.e. $\vdash \neg\alpha(0)$, $\vdash \neg\alpha(\overline{1})$,...), and

(b) $S \vdash \exists x\alpha(x)$

---

[72] As we mentioned, in Gödel's paper [1931] "expresses" is used as identical with "defines"; comp. K. Gödel [1931], 147.

[73] Also called *semirecursive* or *recursively enumerable*. For details, comp. Sect. 4.2.5 (below).

[74] More about this, Ch. 4, Sect. 4.2.1 (below). Remember, "*Bew* comes from the German "beweisbar" (provable).

otherwise S is ω-inconsistent.

If $\alpha(x)$ is a decidable formula of L_S, then ω-consistency is also called 1-consistency. Evidently, ω-consistency does imply 1-consistency and 1-consistency implies (simple) consistency.

By Fact 7 (Lemma), Sect. 4.2.5 (below), for S the following holds:

(L_1) S is Σ_1-complete

(L_2) S is 1-consistent iff S is Σ_1-sound,

i.e., by (L_1) if a Σ_1-formula α is true, then α is provable in S, and by (L_2) the notions "1-consistency" and "Σ_1-soundness" are equivalent, where Σ_1-soundness means the following: if a Σ_1-formula α is provable in S, then α is true.

Now, in order to apply DL (in the form stated above) for stating and proving Gödel's (first) incompleteness theorem we must chose a suitable property P for the DL. Remember that Gödel's undecidable sentence (with the code 17*Gen r* is a self-referential sentence asserting its own unprovability (comp. Sect. 4.2.1.2). Hence, for the property P Gödel took "is not provable in S", whose number-theoretical counterpart Q is ~*Bew*(x). Since *Bew*(x) is expressible in S (as we saw above), it follows that ~*Bew*(x) is also expressible in S. Hence, by DL a formula F_n can be found (where n is its Gödel number) such that F_n expresses "~*Bew*(n)".

Under the preceding considerations, Gödel's theorem is just around the corner.

First incompleteness theorem. (a) *If S is consistent, then* F_n *is not provable in S, and* (b) *If S is* ω*-consistent, then* $\neg F_n$ *is not provable in S.*

Proof (a) (*reductio*). Assume hypothesis and that F_n is provable in S. Then *Bew*(n) is true. And then the formula (Σ_1) expressing it is S is provable in S (by L_1). But F_n expresses ~*Bew*(n); whence $\neg F_n$ expresses *Bew*(n). Now, since *Bew*(n) is true, it follows that the formula expressing it, i.e., $\neg F_n$, is provable in S. Therefore, $\neg F_n$ is also provable in S, and S is inconsistent (contradicting the hypothesis of (a)). So, F_n is not provable in S.

(b) (*reductio*). Assume the hypothesis of (b) and that $\neg F_n$ is provable in S. But, as above, $\neg F_n$ expresses *Bew*(n), and therefore *Bew*(n) is true (by L_2). And this means that F_n is provable; whence again, it follows that S is inconsistent and therefore S is ω-inconsistent, contradicting the hypothesis of (b). So, $\neg F_n$ is not provable in S.

Gödel's Second Incompleteness Theorem. *If S is consistent, then* Con *is not provable in S* (*where* Con *is the formula of S expressing the consistency of S*).

For Con we can take, for example, the formula expressing ~*Bew*($\perp$), where "$\perp$" is a logical falsity (as in Sect. 4.2.1.2) or, more generally, we can proceed as follows. If α is a provable formula of *S*, then if $\neg\alpha$ were provable in *S*, then, evidently, *S* would be inconsistent. Therefore, if $\neg\alpha$ has, say, the Gödel number *k*, then ~*Bew*(*k*) says that *S* is consistent. And therefore for Con we can take the *formula of S* expressing ~*Bew*(*k*). Hence, the following *formal, number-theoretic* and *metamathematical* levels are so correlated:

(*Eq*) Con iff ~*Bew*(*k*) iff *S* is consistent.

Now, as we saw above, F_n is the formula expressing ~*Bew*(*n*). And then, as the first incompleteness theorem (part (a)) showed, the following conditional holds:

(*Cond*) If ~*Bew*(*k*), then ~*Bew*(*n*),

and therefore the following implication is the *formula* expressing the first incompleteness theorem (part (a)):

(*Impl*)[75] Con $\supset F_n$.

Finally, if Con were provable in *S*, then F_n would be provable in *S* (by MP), contradicting the first incompleteness theorem (part (a)).

In what follows let us proceed more formally in stating and proving DL and Gödel's Theorems *via* DL.

4.2.2.1. Diagonal Lemma (DL)[76]

Diagonal Lemma. *For any formula* $\beta(x_2) \in L_{PA}$ *there is a sentence* G *such that*: $PA^{ax} \vdash G \equiv \beta(\overline{g})$, *where g is the Gödel number of* G.[77]

Proof. Let $\delta(x)$ be the diagonal function. As we know (by Sect. 4.1), it is primitive recursive and therefore formally representable in PA^{ax} (cf. Sect. 3.4) by a formula, say, $\Delta(x_1,x_2)$; that is for any numbers *k*, *m* the following

[75] Moreover, *Impl* is itself provable in *S* (cf. Sect. 4.2.1.2, note 59).

[76] Also called "fixed point lemma" or "self-referential lemma". It is mentioned in K. Gödel [1934], §7 (as a result due to R. Carnap [1934], §35). Later, it also appear in B. Rosser [1939] Lemma 1, S. Feferman [1960], Lemma 5.1, C. Smorynski [1977], 827, G. Boolos [1993], 53-54. For a short history of DL, comp. C. Smorynski [1981].

[77] This result holds for any formal system *S* extending the Robinson system *Q*.

holds:

(*) If $\delta(k) = m$, then $\mathrm{PA}^{\mathrm{ax}} \vdash \forall x_2(\Delta(\bar{k}, x_2) \equiv x_2 = \overline{m})$ (cf. Sect. 2, Def. 4*).

Now, let $\alpha(x_1)$ be the formula $\exists x_2(\Delta(x_1,x_2) \wedge \beta(x_2))$. Let n be its Gödel number, and G be the sentence $\exists x_1(x_1 = \bar{n} \wedge \alpha(x_1))$. This sentence is equivalent to $\alpha(\bar{n})$, since

(**) $\mathrm{FOL}^{\mathrm{ax}} \vdash \exists x_1(x_1 = \bar{n} \wedge \alpha(x_1)) \equiv \alpha(\bar{n})$ (cf. Ch. 2, 4.3, Lemma).

Therefore, we have the following derivations in $\mathrm{PA}^{\mathrm{ax}}$:

(1) $\vdash G \equiv \alpha(\bar{n}) \equiv \exists x_2(\Delta(\bar{n}, x_2) \wedge \beta(x_2))$

Now, since the Gödel number of $\alpha(x_1)$ is n, it follows that the Gödel number of G is $\delta(n)$; let g be such a number, i.e., $\delta(n) = g$. And since $\delta(x)$ is representable in $\mathrm{PA}^{\mathrm{ax}}$ by $\Delta(x_1,x_2)$, it follows that

(2) $\vdash \forall x_2(\Delta(\bar{n}, x_2) \equiv x_2 = \bar{g})$; by (*).

And therefore

(3) $\vdash G \equiv \exists x_2(x_2 = \bar{g} \wedge \beta(x_2))$; (1), (2)

(4) $\vdash G \equiv \beta(\bar{g})$; (3); by (**).

The sentence G is called the *fixed point* of the formula $\beta(x_2)$.

4.2.2.2. Gödel's Theorem (*via* DL)

By the proof above, any formula containing a free variable x does admit of a fixed point.[78] Now, if we take the formula $\Pi(y,x)$ expressing formally in $\mathrm{PA}^{\mathrm{ax}}$ the primitive recursive relation $Pf(y,x)$: "*y* is a proof of *x*", and then construct the formula $\neg\exists y\Pi(y,x)$, by DL there is a sentence G such that $\mathrm{PA}^{\mathrm{ax}} \vdash G \equiv \neg\exists y\Pi(y, \bar{g})$, where g is the Gödel number of G. As can be seen, G is *equivalent* to a sentence asserting "G is not provable".

Gödel's Theorem. (1) *If* $\mathrm{PA}^{\mathrm{ax}}$ *is consistent, then* G *is not provable*; (2) *If* $\mathrm{PA}^{\mathrm{ax}}$ *is* ω*-consistent, then* $\neg$G *is not provable.*

Proof (1) (*Reductio*). Assume hypothesis and that $\vdash$G, and then $\vdash \neg\exists y\Pi(y, \bar{g})$ (by PL). Since it is provable, it follows that there is a proof of it in $\mathrm{PA}^{\mathrm{ax}}$ with Gödel number k. And then $Pf(k,g)$ is true (where g is the

[78] Actually, DL has forms in which the given formula has many free variables (cf. S. Kripke, [1996]) or the form of the *generalized diagonal lemma* (cf. G. Boolos [1993], 53-54).

Gödel number of G). Hence $\vdash \Pi(\bar{k},\bar{g})$. But, from the provability of $\neg\exists y\Pi(y,\bar{g})$, equivalent $\forall y\neg\Pi(y,\bar{g})$, it follows that $\vdash \neg\Pi(\bar{k},\bar{g})$ (by Ax4 and MP); contrary to the assumed consistency of PA^{ax}.

(2) (*Reductio*). Assume hypothesis and that $\vdash\neg$G, and then $\vdash \exists y\Pi(y,\bar{g})$. Since by (1) $\nvdash$ G, it follows that for any *n*, $Pf(n,g)$ is false, and then for any *n*, $\vdash \neg\Pi(\bar{n},\bar{g})$. But $\vdash \exists y\Pi(y,\bar{g})$ and $\vdash \neg\Pi(\bar{n},\bar{g})$ (for any *n*) contradict the assumed ω-consistency of PA^{ax}.

4.2.2.3. Gödel-Rosser Theorem for PA^{ax} (*via* DL)

As we saw, the undecidability of G needs the assumption of ω-consistency.[79] But Rosser[80] has shown, for a more complex sentence R, that the undecidability of R can be proved under the assumption of *simple* consistency of PA^{ax}.

Beside the formula Π(y,x) expressing formally in PA^{ax} the primitive recursive relation $Pf(y,x)$: "*y is a proof of x*", this time we also use the formula Π^{Neg}(y,x) expressing formally in PA^{ax} the primitive recursive relation $Pf(y,\text{neg}(x))$: "*y* is a proof of the *negation* of *x*". Using both formulas, Π(y,x) and Π^{Neg}(y,x), the following formula can be constructed:

FORM: $\forall x_2(\Pi(x_2,x_1) \supset \exists x_3(x_3 < x_2 \wedge \Pi^{Neg}(x_3,x_1)))$.

By Diagonal Lemma, there exists a sentence R such that:

(*) $PA^{ax} \vdash R \equiv \forall x_2(\Pi(x_2,\bar{r}) \supset \exists x_3(x_3 < x_2 \wedge \Pi^{Neg}(x_3,\bar{r}))$,

where *r* is the Gödel number of R.

The sentence R is a *Rosser sentence* for PA^{ax}.

Gödel-Rosser Theorem for PA^{ax}. *If* PA^{ax} *is consistent, then* R *is undecidable in* PA^{ax}.

Proof.

1. R *is not provable in* PA^{ax}.

(*Reductio*). Assume hypothesis of the theorem and suppose that R is provable in PA^{ax}. Hence there exists a proof of R with, say, *k* its Gödel number. Then $Pf(k,r)$ is true and therefore

(1) $PA^{ax} \vdash \Pi(\bar{k},\bar{r})$

---

[79] The undecidability of a sentence G-type cannot be proved under the weaker assumption of *simple* consistency; comp. Ch. 4, Sect. 4.2.2 (final *Remark*).

[80] B. Rosser [1936].

(2) $\mathrm{PA}^{\mathrm{ax}} \vdash \forall x_2(\Pi(x_2,\bar{r}) \supset \exists x_3(x_3 < x_2 \wedge \Pi^{\mathrm{Neg}}(x_3,\bar{r}))$; from the assumption that $\mathrm{PA}^{\mathrm{ax}} \vdash \mathrm{R}$ and (*), by PL.

(3) $\mathrm{PA}^{\mathrm{ax}} \vdash \Pi(\bar{k},\bar{r}) \supset \exists x_3(x_3 < \bar{k} \wedge \Pi^{\mathrm{Neg}}(x_3,\bar{r}))$; (2) Ax4, MP.

(4) $\mathrm{PA}^{\mathrm{ax}} \vdash \exists x_3(x_3 < \bar{k} \wedge \Pi^{\mathrm{Neg}}(x_3,\bar{r}))$; (1), (3), MP.

But $\mathrm{PA}^{\mathrm{ax}}$ is consistent (by hypothesis) and then ¬R is not provable in $\mathrm{PA}^{\mathrm{ax}}$. Therefore, for any number *m*, $Pf(m,\mathrm{neg}(r))$ is false; hence for any number $m<k$, $Pf(m,\mathrm{neg}(r))$ is false. And this implies that for any $m<k$ we have $\mathrm{PA}^{\mathrm{ax}} \vdash \neg\Pi^{\mathrm{Neg}}(\bar{m},\bar{r})$, i.e., $\vdash \neg\Pi^{\mathrm{Neg}}(0,\bar{r})$, $\vdash \neg\Pi^{\mathrm{Neg}}(\bar{1},\bar{r})$,..., $\vdash \neg\Pi^{\mathrm{Neg}}(\overline{k-1},\bar{r})$. And then

$$\mathrm{PA}^{\mathrm{ax}} \vdash \neg\Pi^{\mathrm{Neg}}(0,\bar{r}) \wedge \neg\Pi^{\mathrm{Neg}}(\bar{1},\bar{r}) \wedge ... \wedge \neg\Pi^{\mathrm{Neg}}(\overline{k-1},\bar{r}).$$

Whence

(5) $\mathrm{PA}^{\mathrm{ax}} \vdash \forall x_3(x_3 < \bar{k} \supset \neg\Pi^{\mathrm{Neg}}(x_3,\bar{r}))$; by Sect. 1.4, Th. V.2*; equivalent:

$$\mathrm{PA}^{\mathrm{ax}} \vdash \neg\exists x_3(x_3 < \bar{k} \wedge \Pi^{\mathrm{Neg}}(x_3,\bar{r})),$$ contradicting (4).

2. ¬R *is not provable in* $\mathrm{PA}^{\mathrm{ax}}$.

(*Reductio*). Assume hypothesis of the theorem and suppose that ¬R is provable in $\mathrm{PA}^{\mathrm{ax}}$. So there is a proof of ¬R, let *m* be its Gödel number. Hence $Pf(m,\mathrm{neg}(r))$ is true; it follows that

(1) $\mathrm{PA}^{\mathrm{ax}} \vdash \Pi^{\mathrm{Neg}}(\bar{m},\bar{r})$.

But $\mathrm{PA}^{\mathrm{ax}}$ is consistent (by hypothesis), so R is not provable in $\mathrm{PA}^{\mathrm{ax}}$. Hence there is no number *i*, and therefore no number $i \le m$, such that $Pf(i,r)$ is true. Equivalent, for *any* $i \le m$, $Pf(i,r)$ is false. And this implies that for any $i \le m$, $\mathrm{PA}^{\mathrm{ax}} \vdash \neg\Pi(\bar{i},\bar{r})$, i.e., $\vdash \neg\Pi(0,\bar{r})$, $\vdash \neg\Pi(\bar{1},\bar{r})$,..., $\vdash \neg\Pi(\bar{m},\bar{r})$, and therefore $\mathrm{PA}^{\mathrm{ax}} \vdash \neg\Pi(0,\bar{r}) \wedge \neg\Pi(\bar{1},\bar{r}) \wedge ... \wedge \neg\Pi(\bar{m},\bar{r})$. Whence

(2) $\mathrm{PA}^{\mathrm{ax}} \vdash \forall x_2(x_2 \le \bar{m} \supset \neg\Pi(x_2,\bar{r})$, by Sect. 1.4 Th. V.1*, from which it follows

(3) $\mathrm{PA}^{\mathrm{ax}} \vdash x_2 \le \bar{m} \supset \neg\Pi(x_2,\bar{r})$; (2) Ax4. MP.

Let us now consider the following deduction:

(a) $\bar{m} < x_2$; hyp.

(b) $\Pi^{\mathrm{Neg}}(\bar{m},\bar{r})$; the result (1)

(c) $\bar{m} < x_2 \wedge \Pi^{\mathrm{Neg}}(\bar{m},\bar{r})$; (a), (b), PL

(d) $\exists x_3(x_3 < x_2 \wedge \Pi^{\mathrm{Neg}}(x_3, \bar{r})$; (c), Gen $\exists$

(e) $\bar{m} < x_2 \vdash \exists x_3(x_3 < x_2 \wedge \Pi^{\mathrm{Neg}}(x_3, \bar{r}))$; (a)-(d)

(f) $\mathrm{PA}^{\mathrm{ax}} \vdash \bar{m} < x_2 \supset \exists x_3(x_3 < x_2 \wedge \Pi^{\mathrm{Neg}}(x_3, \bar{r})$; (e) Ded. Th.

(4) $\mathrm{PA}^{\mathrm{ax}} \vdash x_2 \leq \bar{m} \vee \bar{m} < x_2$; by Th. III, 11 of Sect. 1.4.

(5) $\mathrm{PA}^{\mathrm{ax}} \vdash \neg\Pi(x_2, \bar{r}) \vee \exists x_3(x_3 < x_2 \wedge \Pi^{\mathrm{Neg}}(x_3, \bar{r})$; (3), (f), (4), MP, PL[81]

(6) $\mathrm{PA}^{\mathrm{ax}} \vdash \forall x_2(\Pi(x_2, \bar{r}) \supset \exists x_3(x_3 < x_2 \wedge \Pi^{\mathrm{Neg}}(x_3, \bar{r})))$; (5), PL, Gen

And this means that PA$\vdash$R; contradicting the assumed consistency of $\mathrm{PA}^{\mathrm{ax}}$.

Remark. The proof just given of Gödel-Rosser Theorem (*via* DL) is based on the formal expressibility in $\mathrm{PA}^{\mathrm{ax}}$ of the primitive recursive relations $Pf(y,x)$ and $Pf(y,\mathrm{neg}(x))$ by Π(y,x) and Π^{Neg}(y,x), respectively, and of some minimal facts of PA.

Let us give, in what follows, a proof of this theorem, based on the definability in L_{PA} of the relations $Pf(y,x)$ and $Pf(y,\mathrm{neg}(x))$ by the respective formulas and on the Σ_1-completeness of $\mathrm{PA}^{\mathrm{ax}}$.

Gödel-Rosser Theorem. *If* $\mathrm{PA}^{\mathrm{ax}}$ *is consistent, then* R *is undecidable in* $\mathrm{PA}^{\mathrm{ax}}$.

Proof (version). As above, R is the fixed point of FORM, i.e.,

(*) $\mathrm{PA}^{\mathrm{ax}} \vdash \mathrm{R} \equiv \forall x_2(\Pi(x_2, \bar{r}) \supset \exists x_3(x_3 < x_2 \wedge \Pi^{\mathrm{Neg}}(x_3, \bar{r})))$.

1. R *is not provable in* $\mathrm{PA}^{\mathrm{ax}}$.

(*Reductio*). Assume that $\mathrm{PA}^{\mathrm{ax}}$ is consistent and that $\mathrm{PA}^{\mathrm{ax}}\vdash$R. So there is a proof of R with the Gödel number, say, *k*. Hence $Pf(k,r)$ is true and then $\Pi(\bar{k}, \bar{r})$ is true.[82] Since $\mathrm{PA}^{\mathrm{ax}}$ is consistent (by hypothesis), it follows that ¬R is not provable in $\mathrm{PA}^{\mathrm{ax}}$. So there is no number *m*, and then no number $m < k$, such that $Pf(k,\mathrm{neg}(r))$. The formal expression of this assertion is $\neg\exists x_3(x_3 < \bar{k} \wedge \Pi^{\mathrm{Neg}}(x_3, \bar{r})$, a true formula of L_{PA}. It follows that the conjunction of both true formulas is a *true* Σ_0-formula,[83] provable in $\mathrm{PA}^{\mathrm{ax}}$,[84] i.e.,

[81] By $\vdash(p_1 \supset q_1) \supset [(p_2 \supset q_2) \supset ((p_1 \vee p_2) \supset (q_1 \vee q_2))]$.

[82] In the standard model M of L_{PA}.

[83] And then a *true* Σ_1-formula; comp. Sect. 4.2.5, 1 (below)

[84] By Σ_1-completeness of $\mathrm{PA}^{\mathrm{ax}}$; comp. Sect. 4.2.5, 2, Fact 7 (Lemma) (below).

$PA^{ax} \vdash \Pi(\bar{k},\bar{r}) \wedge \neg\exists x_3(x_3 < \bar{k} \wedge \Pi^{Neg}(x_3,\bar{r})$, whence, by Gen $\exists$ it follows that

(**) $PA \vdash \exists x_2(\Pi(x_2,\bar{r}) \wedge \neg\exists x_3(x_3 < x_2 \wedge \Pi^{Neg}(x_3,\bar{r})))$.

But from (*) and the supposition that $PA^{ax} \vdash R$ it follows

(***) $PA \vdash \forall x_2(\Pi(x_2,\bar{r}) \supset \exists x_3(x_3 < x_2 \wedge \Pi^{Neg}(x_3,\bar{r})))$

As can be seen, (**) and (***) are contradictory. So if PA^{ax} is consistent, R is not provable in PA^{ax}.

2. $\neg$R *is not provable in* PA^{ax}.

(*Reductio*). Assume that PA^{ax} is consistent and that $PA^{ax} \vdash \neg R$. So there is an *m* the Gödel number of a proof of $\neg$R in PA^{ax}. Hence $Pf(m,\text{neg}(r))$ is true and therefore $\Pi^{Neg}(\bar{m},\bar{r})$ is true and then provable in PA^{ax}, i.e.,

(a) $PA^{ax} \vdash \Pi^{Neg}(\bar{m},\bar{r})$

Now, since PA^{ax} is assumed to be consistent, R is not provable in PA^{ax}. Hence there is no number *i*, and then no number $i \leq m$, the Gödel number of a proof of R in PA^{ax}. It follows that

(b) $PA^{ax} \vdash \forall x_2(x_2 \leq \bar{m} \supset \neg\Pi(x_2,\bar{r}))$

From (a) follows that

(c) $PA^{ax} \vdash \forall x_2(x_2 > \bar{m} \supset \exists x_3(x_3 < x_2 \wedge \Pi^{Neg}(x_3,\bar{r})))$[85]

A fact of PA^{ax} is

(d) $PA^{ax} \vdash \forall x_2(x_2 \leq \bar{m} \vee x_2 > \bar{m})$

Now, (b), (c) and (d) have the following forms:

(b*) $\forall x_2(\alpha \supset \beta)$

(c*) $\forall x_2(\gamma \supset \delta)$

(d*) $\forall x_2(\alpha \vee \gamma)$

From (b*) and (c*) it follows

(1) $PA^{ax} \vdash \forall x_2(\alpha \supset \beta) \wedge \forall x_2(\gamma \supset \delta)$; by PL

(2) $PA^{ax} \vdash \forall x_2((\alpha \supset \beta) \wedge (\gamma \supset \delta))$; (1) FOL

(3) $PA^{ax} \vdash \forall x_2((\alpha \vee \gamma) \supset (\beta \vee \delta))$; (2) FOL

(4) $PA^{ax} \vdash \forall x_2(\alpha \vee \gamma) \supset \forall x_2(\beta \vee \delta)$; (3) FOL

(5) $PA^{ax} \vdash \forall x_2(\beta \vee \delta)$; (4), (d*), (d), MP.

But (5), *in extenso*, is

---

[85] By $FOL^{ax} \vdash P(\bar{m}) \supset \forall y(y > \bar{m} \supset \exists z(z < y \wedge P(z)))$

(6) $\quad PA^{ax} \vdash \forall x_2(\neg\Pi(x_2,\bar{r}) \vee \exists x_3(x_3 < x_2 \wedge \Pi^{Neg}(x_3,\bar{r})))$, i.e.,

(7) $\quad PA^{ax} \vdash \forall x_2(\Pi(x_2,\bar{r}) \supset \exists x_3(x_3 < x_2 \wedge \Pi^{Neg}(x_3,\bar{r})))$.

That is, $PA^{ax} \vdash R$ and therefore PA^{ax} *is inconsistent*. So, by assumption of consistency of PA^{ax}, it follows that R is *undecidable* in PA^{ax}.

4.2.3. Gödel's Theorem (*via* Paradoxes)

The fact that there is a strong connection between the argument of the existence of an undecidable sentence and the paradoxes was written down by Gödel himself, in the following terms: "The analogy of this argument with the Richard antinomy leaps to the eye. It is closely related to the "Liar" too"; [Footnote] "Any epistemological antinomy could be used for a similar proof of the existence of undecidable propositions."[86]

Let us develop this idea of the relation between Gödel's results and the paradoxes.

4.2.3.1. Paradoxes (examples)

(1) *Epimenides Paradox*[87]

Epimenides the Cretan made the following assertion:

(E) All Cretans are liars.

A variant of (E): "This sentence is false", or equivalently "This sentence is not true".

Is (E) true of false?

(E) is true iff what is stated by (E) is true iff (E) is not true.

Therefore, we are in the impossibility to assign a truth value to the sentence (E). Hence (E) is paradoxical, since we derive:

(E) is true iff (E) is not true.

(2) *Grelling Paradox*[88]

This paradox can be derived in the following way. An adjective is called *autological* if it has the property it denotes, or if it is true of itself (e.g. "short", "polysyllabic", "English"); the adjective is *heterological* if it is

[86] K. Gödel, [1986], 149.

[87] Also known as *Liar Paradox*.

[88] Cf. K. Grelling, L. Nelson, [1908]; it is a semantic paradox, also known as "Heterological Paradox".

not true of itself (e.g. "long", "monosyllabic", "Romanian"). Let us ask: is "heterological" heterological? Any attempt to answer the question does generate a paradox, since "heterological" is heterological iff "heterological" is not true of itself iff "heterological" is not heterological. We have therefore a paradox:
"heterological" is heterological iff "heterological" is not heterological.

(3) *Russell's Paradox*[89]

Usually, a set is a collection of objects or elements. Some sets contain themselves as element; symbolically: $x \in x$. *Example*: "the set of all sets having more than 10 elements". And some sets do not contain themselves as element; symbolically: $x \notin x$. *Example*: "the set of prime numbers"; it is not a prime number.

Let us divide the set of all sets in two disjunctive sets, in the following way:

$M = \{x \mid x \in x\}$; i.e., $x \in M$ iff $x \in x$

$N = \{x \mid x \notin x\}$; i.e., $x \in N$ iff $x \notin x$.

Let us ask: Is *N* an element of itself or not? We derive: $N \in N$ iff $N \notin N$ (by the definition of *N*). But $N \notin N$ iff $N \in N$ (by the same definition of N). Therefore

N ∈ N iff N ∉ N,

and this is Russell's Paradox.

Russell find the resolution of this paradox in the restriction of the notion "set" to the so called well-defined sets. In his *Theory of Types* Russell classified the sets according to their type or level; type 1: are the individual objects, type 2: sets of objects of type 1, type 3: sets of sets of type 1 or type 2. Generally, if a set is of the type *n*, then its elements are of type $n-1$ or lower. By such construction, the self-reference is excluded and then the sets like *M* and *N* are ruled out.[90]

This is also the solution to Epimenides Paradox, since the sentence (E): "All Cretans are liars" must have a higher type than other sentences made by Cretans, and then the self-reference disappears.[91]. In this case the distinction "object-language" – "metalanguage" is a key one in defining the

[89] Mentioned for the first time in *Gottlob Freges Briefwechsel*, Felix Meiner Verlag, Hamburg, 1980, 59-60 (Russell an Frege, 16.6.1902); is a set-theoretical paradox.
[90] As we saw, for the construction of the sentence G the essential means are: *self-reference* and *negation*.
[91] Comp. A. Tarski, [1936].

notion of truth.

Russell used his paradox to show that the set-theoretic *principle of comprehension*[92] is inconsistent.

Compr $\exists y \forall x(x \in y \equiv A(x))$,

where $A(x)$ is any formula of the language of set theory, with x free variable and in which y does not occur free.

If $A(x)$ is the formula $x \notin x$, then from *Compr* we derive

$\forall x(x \in y \equiv x \notin x)$,

whence for $x = y$ we have:

$y \in y \equiv y \notin y$.

Remark 1. The inconsistency of *Compr* does not depend on the interpretation of the symbol "$\in$". Let us suppose that $x \in y$ means "y is true of x", equivalent "x satisfies y", where y is the Gödel number of a formula $A(x_1)$; symbolic: $\mathrm{Sat}(A(x_1), x)$ or $\mathrm{Sat}(y, x)$. In this interpretation from *Compr* we derive:

$\forall x(\mathrm{Sat}(y, x) \equiv A(x))$.

Let $A(x)$ be: $\sim\mathrm{Sat}(x, x)$; then

$\forall x(\mathrm{Sat}(y, x) \equiv\sim \mathrm{Sat}(x, x))$,

whence, for $x = y$:

$\mathrm{Sat}(y, y) \equiv\sim \mathrm{Sat}(y, y)$.

And this is just the *Grelling Paradox*, and shows the following fact: the language of PA, $\mathrm{L_{PA}}$, does not admit of the semantic predicate of satisfiability (and therefore neither the semantic predicate of truth).[93]

Remark 2. The Grelling Paradox is just another way to state Epimenides Paradox. Respectively, from the last paradox the Grelling Paradox can be derived.

(a) (E) "This sentence is false".

(b) " "Yields a falsehood when appended to its own quotation" yields a falsehood when appended to its own quotation".

The sentence (b) is equivalent to the sentence (a). Since by appending the sentence *mentionated* in (a) to its own quotation is just the sentence (b). Hence (b) says of itself that it is false.

(c) " "Is not true of itself" is not true of itself".[94]

---

[92] "Unrestricted comprehension scheme" or "Naive comprehension scheme".

[93] This fact is just the content of *Tarski's Theorem* (see below).

[94] The form (3) of the sentence (1) is given by Quine [1966], 9.

The sentence (c) is an abbreviation of (b) and as can be seen (c) is a form of Grelling Paradox.

4.2.3.2. Paradoxes and Gödel's Theorem

The Gödel's reference to Liar (Epimenides) Paradox as a source of his undecidable sentence G is wholly justified, if we compare the two self-referential sentences:

(E) This sentence is not true.

G. This sentence is not provable.

G results from (E) by replacing (essential fact!) the semantic notion "true" with the syntactic one "provable".

Now, since G is derivable from (E) and Grelling Paradox is just another form of (E), then, as can be expected, G can be derived from Grelling Paradox. Let us detail this fact. Firstly, we prove a particular form of Gödel's Theorem, from which afterwards we derive the undecidable sentence G *via* Grelling Paradox, and imitate the proof of this theorem in the context of this paradox.

Let $Prv(y,x)$: "y is provable of x".

Let us consider the following *metamathematical* expression: $R(y,x,z)$: "y is the Gödel number of a formula $\alpha(x_1)$, with x_1 free, and z the Gödel number of a proof of the formula $\alpha(\bar{x})$". Its numerical expression, given by arithmetization, is the following:

$$R(y,x,z): \quad Fml(y) \wedge Fr(y,2^{15}) \wedge Pf(z, Sub(y, Num(x), 2^{15}).$$

Since all notions from the definition of $R(y,x,z)$ are primitive recursive it follows that $R(y,x,z)$ is a primitive recursive relation. And then it is *formally* expressible in PA^{ax} by a formula of L_{PA}, say PRV(y,x,z). Let ∃zPrv(y,x,z) be the formula defining in L_{PA} the recursive enumerable relation $(Ez)R(y,x,z)$,[95] whose meaning is "y is provable of x", i.e., $Prv(y,x)$. Then the formula ¬∃zPRV(y,x,z) defines in L_{PA} the numerical relation $\sim Prv(y,x)$, i.e., the metamathematical relation "y is not provable of x". Finally, let us consider the formula ¬∃zPRV(x_1,x_1,z), let m be its Gödel number, and $G: \neg\exists z\mathrm{PRV}(\overline{m},\overline{m},z)$, whose meaning is: "$m$ is not provable of m".

---

[95] On the "recursive enumerable relations"; comp. 4.2.5 (below).

Theorem. *If* $\mathrm{PA}^{\mathrm{ax}}$ *is* ω*-consistent, then* G *is undecidable in* $\mathrm{PA}^{\mathrm{ax}}$.
Proof.
(1) G *is not provable in* $\mathrm{PA}^{\mathrm{ax}}$.
Reductio. Assume the hypothesis of theorem and suppose that $\mathrm{PA}^{\mathrm{ax}} \vdash \mathrm{G}$, i.e., $\mathrm{PA}^{\mathrm{ax}} \vdash \neg\exists z\mathrm{PRV}(\overline{m},\overline{m},z)$. So, there is a number k such that $R(m,m,k)$ is true. And then $\mathrm{PA}^{\mathrm{ax}} \vdash \mathrm{PRV}(\overline{m},\overline{m},\overline{k})$ (since PRV formally expresses R in $\mathrm{PA}^{\mathrm{ax}}$). Therefore $\mathrm{PA}^{\mathrm{ax}} \vdash \exists z\mathrm{PRV}(\overline{m},\overline{m},z)$ (by Gen $\exists$), contradicting the consistency[96] (and therefore ω-consistency) of $\mathrm{PA}^{\mathrm{ax}}$. Hence (1) holds.
(2) $\neg$G *is not provable in* $\mathrm{PA}^{\mathrm{ax}}$.
Reductio. Assume hypothesis and suppose that $\mathrm{PA}^{\mathrm{ax}} \vdash \neg\mathrm{G}$, i.e., $\mathrm{PA}^{\mathrm{ax}} \vdash \exists z\mathrm{PRV}(\overline{m},\overline{m},z)$. By (1), G is not provable in $\mathrm{PA}^{\mathrm{ax}}$, and this means that for *any n*, $R(m,m,n)$ is false. Hence for *any n*, $\mathrm{PA}^{\mathrm{ax}} \vdash \neg\mathrm{PRV}(\overline{m},\overline{m},\overline{n})$, contradicting the assumed ω-consistency of $\mathrm{PA}^{\mathrm{ax}}$.
Remark. (2) can also be proved in the following way. By (1) G is not provable in $\mathrm{PA}^{\mathrm{ax}}$. But G says "*m* is not provable of *m*", hence G is true. It follows that its negation, $\neg$G: $\exists z\mathrm{PRV}(\overline{m},\overline{m},z)$, is a *false* Σ_1-formula. Now, since ω-consistency does imply *1-consistency* and 1-consistency is equivalent to Σ_1-soundness,[97] it follows that since $\mathrm{PA}^{\mathrm{ax}}$ is Σ_1-sound, $\neg$G is not provable in $\mathrm{PA}^{\mathrm{ax}}$.

Now, the argument from this theorem can be imitated *via* Grelling Paradox in the following way. A formula is called *Gödel heterological* if it is not provable of itself.[98] As we saw above, $\neg\exists z\mathrm{PRV}(x_1,x_1,z)$ is such a formula. Let now $\mathrm{GHet}(x_1)$: $\neg\exists z\mathrm{PRV}(x_1,x_1,z)$, and $\mathrm{GHet}(\overline{m})$: $\neg\exists z\mathrm{PRV}(\overline{m},\overline{m},z)$, whose meaning is:

"Gödel heterological" is Gödel heterological.

The formula $\mathrm{GHet}(x_1)$ *defines* the set *GH* of Gödel numbers of Gödel heterological formulas, i.e., $GH = \{x \mid \mathrm{GHet}(x)\}$; $m \in GH$ iff $\mathrm{GHet}(m)$, respectively.

Let us ask: Is "Gödel heterological" Gödel heterological?

The sentence "*Gödel heterological*" *is Gödel heterological* is the sentence $\mathrm{GHet}(\overline{m})$.

---

[96] As can be seen, for the proof of (1) only the consistency of $\mathrm{PA}^{\mathrm{ax}}$ is really needed.
[97] Comp. Sect. 4.2.5, Fact 7. Lemma (below).
[98] Obtainable from the definition of "heterological" by replacing "true" with "provable"; comp. 4.2.3.1(2).

The answer is in the affirmative.

Gödel's Theorem. "*Gödel heterological*" *is Gödel heterological*.

Proof (*Reductio*). Suppose that GHet(x_1) is not Gödel heterological, i.e., GHet(x_1) is provable of itself, i.e., $\mathrm{PA}^{ax} \vdash \mathrm{GHet}(\overline{m})$. Being provable of itself, there is a number k such that $R(m,m,k)$ (from the theorem above) is true, and then $\mathrm{PA}^{ax} \vdash \mathrm{PRV}(\overline{m},\overline{m},\overline{k})$. Whence, by Gen $\exists$, $\mathrm{PA}^{ax} \vdash \exists z\mathrm{PRV}(\overline{m},\overline{m},z)$, contra consistency of PA^{ax}.

Variant (*Reductio*). Suppose that GHet(x_1) is not Gödel heterological. This means that $\mathrm{PA}^{ax} \vdash \neg\exists z\mathrm{PRV}(\overline{m},\overline{m},z)$. On the other hand, since GHet(x_1) is not Gödel heterological it follows that its Gödel number $m \notin GH$. So $(Ez)R(m,m,z)$ is true, whence $\exists z\mathrm{PRV}(\overline{m},\overline{m},z)$ is a true Σ_1-formula, and therefore $\mathrm{PA}^{ax} \vdash \exists z\mathrm{PRV}(\overline{m},\overline{m},z)$ (by Σ_1-completeness of PA^{ax}), contradicting the consistency of PA^{ax}.

From this argument it follows that GHet(x_1) is not provable of itself;[99] i.e., GHet(x_1) is Gödel heterological. And this means that $\mathrm{GHet}(\overline{m})$ is *true* but *not provable in* PA^{ax}. Since its negation, $\neg\mathrm{GHet}(\overline{m})$, is false, neither it is provable in PA^{ax} (by Σ_1-soundness of PA^{ax}, since $\neg\mathrm{GHet}(\overline{m})$ is a false Σ_1-sentence). Hence $\mathrm{GHet}(\overline{m})$ is undecidable in PA^{ax}.

Since from Grelling (Liar, Epimenides) Paradox the Gödel's Theorem can be derived, replacing "true" with "provable", this fact suggests that from Grelling Paradox another important result can be derived: Tarski's Theorem.

Tarski's Theorem. $\mathrm{L_{PA}}$ *does not admit of a satisfiability predicate*.

Proof (*Reductio*). Let $Sat(y,x)$ be the relation: "*y* is the Gödel number of a formula $\alpha(x_1)$ and *x* does satisfy $\alpha(x_1)$, or "*y* is true of *x*". Suppose that $\mathrm{L_{PA}}$ has a formula SAT(y,x) defining it, i.e., for any *m, n*:

$Sat(m,n)$ iff $\mathrm{SAT}(\overline{m},\overline{n})$ is true (in M).

Let HET(x) be the formula ¬SAT(x,x). This formula defines the set *H* of Gödel numbers of the heterological formulas, i.e.,

$H = \{x \mid \neg\mathrm{SAT}(\overline{x},\overline{x})\}$; $n \in H$ iff $\mathrm{Het}(\overline{n})$ is true, respectively.

Let *n* be the Gödel number of HET(x). Is HET(x) heterological?

---

[99] Observe that the following expressions are equivalent: "*Gödel heterological*" *is Gödel heterological*, $\mathrm{GHet}(\overline{m})$ and "Is not provable of itself" is not provable of itself.

We can derive: HET(x) is heterological iff $Sat(n,n)$ iff SAT $(\overline{n},\overline{n})$ is true (by definability). But HET(x) is heterological iff $n \in H$ iff $\neg$SAT $(\overline{n},\overline{n})$ is true; impossible.

Therefore, there is no formula of L_{PA} defining the set of Gödel numbers of heterological formulas.[100] And since the notion of satisfiability can be equivalently defined in terms of truth, it follows that the notion of truth is also not definable in L_{PA}; and this is the standard form of Tarski's Theorem.

4.2.4. Kleene's generalizations of Gödel's and Rosser's Theorems[101]

4.2.4.1. Kleene's T-predicate

The construction of the predicate $T(z,x,y)$ [102] is based on the idea of defining the computable functions in terms of the systems of equations.[103] An equation is a formula of the form $r = s$, where r and s are terms.[104] A system Z of equations is a finite sequence of equations $r_1 = s_1$, $r_2 = s_2$,..., $r_m = s_m$, where r_m has the form $f_j^n(t_1,...,t_n)$ (where f_j^n is called the *principal function letter* of *Z*). The system *Z* has two rules, R_1 (substitution), a rule with one premise, and R_2 (replacement), a rule with two premises. A *deduction* of an equation *e* from a system *Z* is a finite sequence of equations, $e_1,...,e_k$, where $e_k = e$, and each equation of the sequence either is an equation of *Z*, or is a consequence of a preceding equation using R_1, or is a consequence of two preceding equations using R_2. If there exist a deduction of *e* from *Z*, then *e* is *deducible* from *Z* (symbolically: $Z \vdash e$).

Definition.*Let* $\varphi(x_1,...,x_n)$ [105] *be an n-ary number-theoretic partial function.* φ *is computable by a system Z of equations iff* f_j^n *is the principal function*

[100] This fact also follows from the *inconsistency* of *Compr*; comp. 4.2.3.1(3).

[101] The sections 4.2.4 and 4.2.6 (5.1 and 5.2) are an account of Kleene's remarkable results in the analysis of Gödelian incompleteness phenomenon.

[102] It is the key notion of a lot of results, including the generalized forms of incompleteness theorems (Gödel, Rosser).

[103] A Herbrand's idea, developed by Gödel [1934] (in vol. K. Gödel [1986], 346-371).

[104] For details, comp. S.C. Kleene, [1952], §54.

[105] In Kleene's notation small Greek letters φ, ψ, χ,... denote the number-theoretic functions.

letter of Z and for any natural numbers $k_1,...,k_n,k$ *the following holds:*

$$Z \vdash f_j^n(\bar{k}_1,...,\bar{k}_n) = \bar{k} \text{ iff } \varphi(k_1,...,k_n) = k.$$

If there is such a system of equations defining φ, then φ is called Herbrand- Gödel computable.

Now, similar to the arithmetization of the syntax of PA[ax], the *metamathematical* notions of Kleene's calculus of equations can also be *arithmetized*.[106] E.g. $Eq(x)$: "*x* is the Gödel number of an equation"; $Z(x)$: "*x* is the Gödel number of a system of equations"; $\mathfrak{S}_n(Z,x_1,...,x_n,Y)$: "*Y* is a deduction from *Z* of an equation of the form $f_j^n(\bar{x}_1,...,\bar{x}_n) = \bar{x}$, where f_j^n is the principal function symbol of *Z*, $\bar{x}_1,...,\bar{x}_n,\bar{x}$ are the corresponding numerals"; $\mathfrak{U}(Y) = x$, where *Y* is a deduction of an equation of the form $r = \bar{x}$. By arithmetization, the metamathematical notions of this kind have their corresponding numerical notions (i.e., number-theoretic functions and relations). Moreover, all of them are *primitive recursive*.[107]

The numerical expression of the metamathematical predicate $\mathfrak{S}_n(Z,x_1,...,x_n,Y)$ is the primitive recursive predicate $S_n(z,x_1,...,x_n,y)$. Using this predicate, Kleene introduces the numerical (number-theoretic) predicate $T_n(z,x_1,...,x_n,y)$:[108]

$$T_n(z,x_1,...,x_n,y) =_{df} S_n(z,x_1,...,x_n,y) \wedge (t)_{t<y} \widetilde{S}(z,x_1,...,x_n,t).$$

Evidently, for *S* and *T* the following holds:

(1) $T_n(z,x_1,...,x_n,y) \rightarrow S_n(z,x_1,...,x_n,y)$

(2) $(Ey)\, T_n(z,x_1,...,x_n,y) \leftrightarrow (Ey) S_n(z,x_1,...,x_n,y)$.[109]

4.2.4.2. Theorems

Using the predicate $T(z,x_1,...,x_n,y)$ Kleene states and proves the following theorems.

---

[106] Comp. S.C. Kleene, [1952], §56.

[107] Comp. S.C. Kleene, [1952], §56, Lemma III.

[108] Comp. S.C. Kleene, [1952], §57.

[109] Here "↔" and "~" are our intuitive symbols for *equivalence* and *negation*, respectively (cf. Sect. 2 above). They correspond to Kleene's intuitive symbols "≡" and "–", respectively.

Enumeration Theorem.[110] *For each* $n \geq 0$, *given any recursive predicate* $R(x_1,...,x_n, y)$, *numbers f and g can be found such that*

(1) $(Ey)R(x_1,...,x_n, y) \leftrightarrow (Ey)T_n(f, x_1,...,x_n, y)$

(2) $(y)R(x_1,...,x_n, y) \leftrightarrow (y)\widetilde{T}_n(g, x_1,...,x_n, y)$.[111]

As can be seen, for $z = 0,1,2,...$ the $n+1$-place predicate $(Ey)T_n(z, x_1,...,x_n, y)$ *enumerates* the predicates of the form $(Ey)R(x_1,...,x_n, y)$, with $R(x_1,...,x_n, y)$ recursive. Similarly, $(y)\widetilde{T}_n(z, x_1,...,x_n, y)$ enumerates the predicates of the form $(y)R(x_1,...,x_n, y)$, with $R(x_1,...,x_n, y)$ recursive.

With the idea of diagonalization, the enumeration theorem leads to the following result.

Theorem.[112] *Let* $R(x, y)$ *be an arbitrary recursive predicate. Then the numbers f and g can be found such that*

(1) $(Ey)R(f, y) \not\leftrightarrow (y)\widetilde{T}(f, f, y)$

(2) $(y)R(g, y) \not\leftrightarrow (Ey)T(g, g, y)$,

where $T(z, x, y)$ *is the predicate* $T_n(z, x_1,...,x_n, y)$ *for* $n = 1$,[113] *and* "$\not\leftrightarrow$" *is the negation of the intuitive equivalence.*

Proof (1). We have the following derivations:

(a) $(Ey)R(x, y) \leftrightarrow (Ey)T(f, x, y)$; by Enum. Th. (1)

(b) $(Ey)R(f, y) \leftrightarrow (Ey)T(f, f, y)$; (a) Subst. f / x

(c) $(Ey)R(f, y) \not\leftrightarrow (\widetilde{E}y)T(f, f, y)$; (b) PL

(d) $(Ey)R(f, y) \not\leftrightarrow (y)\widetilde{T}(f, f, y)$; (c) FOL

Hence (1) holds.

(2). Again, we have the following derivations:

(a) $(y)R(x, y) \leftrightarrow (y)\widetilde{T}(g, x, y)$; by Enum. Th. (2)

(b) $(y)R(g, y) \leftrightarrow (y)\widetilde{T}(g, g, y)$; (a) Subst. g / x

(c) $(y)R(g, y) \not\leftrightarrow (Ey)T(g, g, y)$; (b) FOL

Hence (2) holds.

---

[110] Comp. S.C. Kleene, [1943], [1952], §57.

[111] For a proof of this theorem; comp. S.C. Kleene, [1952], 280-2.

[112] Cf. S.C. Kleene, [1952], 283, Theorem V (Part I).

[113] In what follows we omit the attachment of the subscript "1" to *T*.

Remark 1. If the recursive predicate is $R(x)$, then, by the theorem above, the numbers *f* and *g* can be found such that:

(3) $R(f) \nleftrightarrow (y)\widetilde{T}(f,f,y)$

(4) $R(g) \nleftrightarrow (Ey)T(g,g,y)$.

For the proof of (3) and (4), given $R(x)$ recursive, we take instead the recursive predicate $R(x,y)$ defined as $R(x) \wedge y = y$, for which evidently the following holds: $R(x) \leftrightarrow R(x,y)$.

Remark 2. Since $T(x,x,y)$ is recursive, so is $\widetilde{T}(x,x,y)$ (by Sect. 3.2). And then, by the above theorem, the predicate $(y)\widetilde{T}(x,x,y)$ is a predicate of the form $(y)R(x,y)$, with $R(x,y)$ recursive, not expressible in the *dual* form $(Ey)R(x,y)$,[114] with $R(x,y)$ recursive.

As we saw, by Enumeration Theorem the predicate $(Ey)T(z,x,y)$ $(z = 0,1,2,...)$ enumerates (with repetitions) all predicates of the form $(Ey)R(x,y)$, with $R(x,y)$ recursive. By the preceding theorem, the predicate $(y)\widetilde{T}(x,x,y)$, constructed by diagonalization, is not a predicate in this enumeration. Therefore, $(y)\widetilde{T}(x,x,y)$ is not recursively enumerable.[115]

Intermezzo. "Recursive" – "Computable/Decidable"

(a) *Any recursive function* φ^n *is effectively computable.*

Argument. Let *Z* be a system of equations defining the recursive function φ^n. Then for the given arguments $x_1,...,x_n$ the value of φ can be found by deducing the equations from *Z* until the derivation of the equation $f_j^n(\bar{x}_1,...,\bar{x}_n) = \bar{x}$, where *x* is just the value of φ^n; i.e., $\varphi(x_1,...,x_n) = x$.[116]

(b) *Any recursive predicate* R_n *is effectively decidable.*

Argument. If R^n is recursive,[117] then for the given arguments $x_1,...,x_n$ we

[114] Any recursive predicate is expressible in both forms; cf. S.C. Kleene, [1952], §57, Theorem VI.

[115] For a simple proof, by *reductio*, comp. M. Davis, [1958], 68, Theorem 1.6. For the "recursive enumerable relations", comp. Sect. 4.2.5 below.

[116] Such a system always exists and the deduction of such an equation is always possible; cf. S.C. Kleene, [1952], §§54, 55, 58.

[117] R^n is recursive iff its characteristic function C_{R^n} is recursive; comp. Sect. 3.2.

can decide whether it is true or false. For this fact we compute the value of its characteristic function C_{R^n} (in the way indicated in (a)) and then read such a value, 0 or 1.

The converses of (a) and (b) seems to be true, as is stated by *Church's Thesis*:

(a*) Any effectively computable function is recursive.

(b*) Any effectively decidable predicate is recursive.

The sentences (a), (b), (a*), (b*) together connect the notions "recursive" – "computable/decidable" and give a definition to the notion of "*algorithm*".[118] That is, "To give a decision procedure for a predicate $P(x)$ thus means to give a general recursive predicate $R(x)$ such that $P(x) \leftrightarrow R(x)$".[119]

4.2.4.3. Church's Theorem[120]

There is no algorithm for either of the predicates $(y)\widetilde{T}(x,x,y)$ or $(Ey)T(x,x,y)$.

Proof. By (3) and (4) of the Remark 1 to the *Theorem* (above).

4.2.4.4. Generalized form of Gödel's Theorem

Kleene's provability predicate

In Kleene [1952] the provability predicate is the predicate $(Ey)R(x,y)$, where $R(x,y)$ is primitive recursive and means: "*y* is a proof of the formula $\mathrm{A}(\bar{x})$". Hence $(Ey)R(x,y)$ means: "$\mathrm{A}(\bar{x})$ is provable", and therefore the following equivalence holds:

(*) $(Ey)R(x,y) \leftrightarrow \vdash \mathrm{A}(\bar{x})$.[121]

The construction of such predicate runs along the following lines.[122]

Let $\mathfrak{R}(x,Y)$ be the following metamathematical predicate:

"Y is a proof of $\mathrm{A}(\bar{x})$". Then $(EY)\mathfrak{R}(x,Y)$ will be the

[118] Equivalent "*computable* procedure for a numerical function", "*decision* procedure for a numerical predicate", respectively.

[119] Cf. S.C. Kleene, [1952], 301.

[120] Cf. A. Church, [1936(b)]; S.C. Kleene, [1952], §60, Theorem XII.

[121] A system *S* formalizes the theory of the predicate $P(x)$ if A(x) expresses $P(x)$ and (*) holds.

[122] Comp. S.C. Kleene, [1952], §60.

metamathematical predicate "$\mathrm{A}(\bar{x})$ is provable", and therefore $(Ey)R(x,y)$ is its *numerical* counterpart; whence (*).

Let $P(x)$ be a predicate, let A(x) be a formula expressing it in *S*. Then S is a *correct* (*sound*) and *complete* formalization for $P(x)$ if the following holds:

(**) $\quad \vdash \mathrm{A}(\bar{x}) \leftrightarrow P(x)$

(where the left-right part means correctness and its converse means completeness of *S* for $P(x)$).

By (*) and (**) we have

(***) $\quad (Ey)R(x,y) \leftrightarrow P(x)$, i.e., *S* is a formal system correct for $P(x)$ (left-right implication) and complete for $P(x)$ (right-left implication) if there exists a recursive $R(x,y)$ such that (***) holds.

Gödel's Theorem (I).[123] *There is no correct and complete formal system for the predicate* $(y)\widetilde{T}(x,x,y)$.

Proof. There is no recursive $R(x,y)$ such that (***) holds for the predicate $(y)\widetilde{T}(x,x,y)$, i.e., $(Ey)R(x,y) \leftrightarrow (y)\widetilde{T}(x,x,y)$ does not hold for all *x*, since by Theorem (4.2.4.2, part (1)) for any recursive $R(x,y)$ a number *f* can be found such that $(Ey)R(f,y) \not\leftrightarrow (y)\widetilde{T}(f,f,y)$, i.e., for $x = f$ (***) does not hold.

Gödel's Theorem (II).[124] *Let S be a formal system in which* A(x) *expresses the predicate* $(y)\widetilde{T}(x,x,y)$. *Then a number f can be found such that: if S is correct for* $(y)\widetilde{T}(x,x,y)$, *then* $(y)\widetilde{T}(f,f,y)$ *is true but* $\mathrm{A}(\bar{f})$ *is not provable in S.*

Proof. Assume hypothesis of the theorem and let $R(x,y)$ be a recursive predicate for which (*) holds, i.e.,

(1) $\quad (Ey)R(x,y) \leftrightarrow \vdash \mathrm{A}(\bar{x})$.

For $R(x,y)$, by Enumeration Theorem, there is an *f* such that

(2) $\quad (Ey)R(x,y) \leftrightarrow (Ey)T(f,x,y)$, and therefore

(3) $\quad (Ey)R(f,y) \leftrightarrow (Ey)T(f,f,y)$; (2) Subst. f/x

[123] Comp. S.C. Kleene, [1952], §60 Theorem XIII (Part I) (a generalized form of Gödel's Theorem).

[124] Cf. S.C. Kleene, [1952], §60, Theorem XIII (Part II).

(4) $\vdash \mathrm{A}(\bar{f}) \to (y)\widetilde{T}(f,f,y)$; by the assumed correctness of *S*

(5) $(y)\widetilde{T}(f,f,y) \leftrightarrow (\widetilde{E}y)T(f,f,y) \leftrightarrow \nvdash \mathrm{A}(\bar{f})$; by (1), (3)

(6) $\vdash \mathrm{A}(\bar{f})$; hyp.

(7) $(y)\widetilde{T}(f,f,y)$; (4), (6) MP

(8) $\nvdash \mathrm{A}(\bar{f})$; by (5)

(9) $\nvdash \mathrm{A}(\bar{f})$; (6)-(8) (reductio)

(10) $(y)\widetilde{T}(f,f,y)$; by (5).

Hence $(y)\widetilde{T}(f,f,y)$ is true and $\mathrm{A}(\bar{f})$ is not provable in *S*.

Remark. The idea of "generalized form" of Gödel's Theorem is given by its *abstract* formulation, i.e., the only requirements on *S* is that it has a formula A(x) expressing the predicate $(y)\widetilde{T}(x,x,y)$, is correct for this predicate and that it satisfies (*): $(Ey)R(x,y) \leftrightarrow \vdash \mathrm{A}(\bar{x})$.

In the usual way of construction of an axiomatic system the requirement on *S* to be correct for the predicate $(y)\widetilde{T}(x,x,y)$, expressed by the formula A(x), is replaced by the weaker condition of ω-consistency. This is the case with the following form of Gödel's Theorem.

Gödel's Theorem (III).[125] *Let S be a formal system, let* θ(x,y) *be the formula expressing formally in S the predicate* $T(x,x,y)$, *let* A(x): ∀y¬θ(x,y). *Suppose that for some recursive* $R(x,y)$ *the equivalence* (*) *holds. Then a number f can be found such that:*

(1) *If S is consistent, then* $\nvdash \mathrm{A}(\bar{f})$.

(2) *If S is* ω*-consistent, then* $\nvdash \neg\mathrm{A}(\bar{f})$.

Proof. Since θ(x,y) formally expresses $T(x,x,y)$ in *S* we have:

(a) If $T(x,x,y)$, then $\vdash \theta(\bar{x},\bar{y})$.

(b) If $\widetilde{T}(x,x,y)$, then $\vdash \neg\theta(\bar{x},\bar{y})$.

(1) *Reductio*. Assume hypothesis of (1) and that $\vdash \mathrm{A}(\bar{f})$. Then, by (*) $(Ey)R(f,y)$, and therefore $(Ey)T(f,f,y)$ (*via* Enumeration Theorem). So, there is an *y* such that $T(f,f,y)$ holds, whence by (a) $\vdash \theta(\bar{f},\bar{y})$. But $\mathrm{A}(\bar{f})$ is the formula $\forall y \neg\theta(\bar{f},y)$, and therefore since $\vdash \mathrm{A}(\bar{f})$ (by

[125] Cf. S.C. Kleene, [1952], §60, Theorem XIII.

hypothesis), it follows that $\vdash \neg\theta(\bar{f},\bar{y})$ (by Ax4 and MP). So S is inconsistent. Therefore $\nvdash \mathrm{A}(\bar{f})$.

(2) *Reductio*. Assume hypothesis of (2) and that $\vdash \neg\mathrm{A}(\bar{f})$, i.e., $\vdash \exists \mathrm{y}\theta(\bar{f},\mathrm{y})$. But, by (1), $\nvdash \mathrm{A}(\bar{f})$. And then, by (*), $(\widetilde{E}y)R(f,y)$, equivalent $(y)\widetilde{T}(f,f,y)$, whence for *any* y, $\widetilde{T}(f,f,y)$, and therefore $\vdash \neg\theta(\bar{f},\bar{y})$ (by b); contradicting the assumed ω-consistency of *S*. Therefore $\nvdash \neg\mathrm{A}(\bar{f})$.

4.2.4.5. A symmetric form of Gödel's Theorem[126]

This time Kleene uses the *T*-predicate in order to construct the following predicate

$$(y)[\widetilde{T}((x)_1,x,y) \vee (Ez)_{z\leq y} T((x)_0,x,z)],$$

equivalently $(\widetilde{E}y)[T((x)_1,x,y) \wedge (z)_{z\leq y}\widetilde{T}((x)_0,x,z)]$.

Abbreviations:

$W_0(x,y)$: $T((x)_1,x,y) \wedge (z)_{z\leq y}\widetilde{T}((x)_0,x,z)$

$W_1(x,y)$: $T((x)_0,x,y) \wedge (z)_{z\leq y}\widetilde{T}((x)_1,x,z)$

The following holds:

(1) $(Ey)W_1(x,y) \to (\widetilde{E}y)W_0(x,y)$

Argument (*reductio*). Suppose that for an x fixed there is an y_1 such that $W_1(x,y_1)$ holds, i.e.,

(a) $T((x)_0,x,y_1)$ and

(b) $(z)_{z\leq y_1}\widetilde{T}((x)_1,x,z)$.

And suppose that there is an y_0 such that $W_0(x,y_0)$ holds, i.e.,

(c) $T((x)_1,x,y_0)$ and

(d) $(z)_{z\leq y_0}\widetilde{T}((x)_0,x,z)$.

Now, from (a) and (d) follows that $y_1 > y_0$; and from (b) and (c) follows that $y_0 > y_1$.

---

[126] Cf. S.C. Kleene, [1952], §61, Theorem XV, also called a generalization of Rosser's form of Gödel's Theorem.

Suppose *S* is a formal system whose language contains the formulas B(x) and ¬B(x) and that

(2) $(Ey)W_0(x,y) \to \vdash \mathrm{B}(\bar{x})$ and

(3) $(Ey)W_1(x,y) \to \vdash \neg\mathrm{B}(\bar{x})$.

But, by (*), the provability of formulas B(x) and ¬B(x) means that there exist the recursive predicates $R_0(x,y)$ and $R_1(x,y)$ such that

(4) $(Ey)R_0(x,y) \leftrightarrow \vdash \mathrm{B}(\bar{x})$

(5) $(Ey)R_1(x,y) \leftrightarrow \vdash \neg\mathrm{B}(\bar{x})$.

Gödel's Theorem (a symmetric form). *There is no simply consistent and complete formal system satisfying* (2)-(5).

Proof. Suppose *S* is consistent and satisfies (2)-(5). By Enumeration Theorem (1) the numbers f_0 and f_1 can be found such that (setting $f = 2^{f_0} \cdot 3^{f_1}$)

(6) $(Ey)R_0(x,y) \leftrightarrow (Ey)T(f_0,x,y) \leftrightarrow (Ey)T((f)_0,x,y)$,

(7) $(Ey)R_1(x,y) \leftrightarrow (Ey)T(f_1,x,y) \leftrightarrow (Ey)T((f)_1,x,y)$.

The proof must show that neither $\mathrm{B}(\bar{f})$, nor $\neg\mathrm{B}(\bar{f})$ is provable in *S*.

Reductio.

1. $\vdash \mathrm{B}(\bar{f})$; hyp.
2. $(Ey)R_0(f,y)$; 1, (4)
3. $(Ey)T((f)_0,f,y)$; 2, (6)
4. $\nvdash \neg\mathrm{B}(\bar{f})$; since *S* is consistent (by hyp.)
5. $(\widetilde{E}y)R_1(f,y)$, 4, (5)
6. $(\widetilde{E}y)T((f)_1,f,y)$; 5, (7)
7. $(y)\widetilde{T}((f)_1,f,y)$; 6, FOL
8. $(Ey)T((f)_0,f,y) \wedge (z)_{z\le y}\widetilde{T}((f)_1,f,z)$; 3, 7, i.e., $(Ey)W_1(f,y)$
9. $\vdash \neg\mathrm{B}(\bar{f})$; by (3); contrary to 4.

Hence $\nvdash \mathrm{B}(\bar{f})$.

In order to prove that $\nvdash \neg\mathrm{B}(\bar{f})$ we have to proceed similarly, respectively

1. $\vdash \neg\mathrm{B}(\bar{f})$; hyp.
2. $(Ey)R_1(f,y)$; 1, (5)

3. $(Ey)T((f)_1, f, y)$; 2, (7)

4. $\not\vdash B(\bar{f})$; since S is consistent (by hyp.)

5. $(\widetilde{E}y)R_0(f, y)$, 4, (4)

6. $(\widetilde{E}y)T((f)_0, f, y)$; 5, (6)

7. $(y)\widetilde{T}((f)_0, f, y)$; 6, FOL

8. $(Ey)T((f)_1, f, y) \wedge (z)_{z \leq y}\widetilde{T}((f)_0, f, z)$, 3, 7, i.e., $(Ey)W_0(f, y)$

9. $\vdash \mathrm{B}(\bar{f})$; 8, (2).

Hence $\not\vdash \neg\mathrm{B}(\bar{f})$.

The use of label "symmetric form" for the generalized form of Rosser's form of Gödel's Theorem is motivated by the author in the following way. Since the predicates $T((x)_1, x, y)$ and $T((x)_0, x, z)$ are primitive recursive, they are formally expressible in S by the formulas α(x,y) and β(x,z), respectively.

Let B(x): $\exists \mathrm{y}[(\alpha(\mathrm{x},\mathrm{y}) \wedge \forall \mathrm{z}(\mathrm{z} \leq \mathrm{y} \supset \neg\beta(\mathrm{x},\mathrm{z}))]$

Then ¬B(x): $\neg\exists \mathrm{y}[(\alpha(\mathrm{x},\mathrm{y}) \wedge \forall \mathrm{z}(\mathrm{z} \leq \mathrm{y} \supset \neg\beta(\mathrm{x},\mathrm{z}))]$.

Now, if B(x) is *interpreted* as expressing the predicate $(Ey)W_0(x, y)$ and ¬B(x) as expressing $(\widetilde{E}y)W_0(x, y)$, then $\neg\mathrm{B}(\bar{f})$, i.e.,

$\neg\exists \mathrm{y}[\alpha(\bar{f},\mathrm{y}) \wedge \forall \mathrm{z}(\mathrm{z} \leq \mathrm{y} \supset \neg\beta(\bar{f},\mathrm{z})]$,

is the formula corresponding to the Rosser formula R (comp. Sect. 4.2.1.3) true and not provable.

But, as Kleene says, nothing prevent us to consider some *other interpretation*, taking "entirely symmetrical"[127] ¬B(x) as expressing $(Ey)W_1(x, y)$ and B(x) to be $(\widetilde{E}y)W_1(x, y)$. In this case $\mathrm{B}(\bar{f})$ correspond to Rosser's true but not provable formula R.

4.2.4.6. Kleene's Normal Form Theorem

(1) **Normal form theorem**[128]

In order to define a (total) recursive function, Kleene uses the metamathematical predicate $\mathfrak{S}_n(E, x_1, \ldots, x_n, Y)$: "*E* is a system of equations

[127] S.C. Kleene, [1952], 310.

[128] Cf. Kleene, [1952], §58.

and Y is a deduction from E of an equation of the form $f(\bar{x}_1,...,x_n) = \bar{x}$" (where f is the principal symbol of Z and $\bar{x}_1,...,\bar{x}_n,\bar{x}$ are the numerals for the corresponding natural numbers), and the function $\mathfrak{U}(Y)$ with the following meaning: $\mathfrak{U}(Y) = x$: "Y is a deduction of an equation of the form $r = \bar{x}$ (where $\bar{x}$ is a numeral).[129] Now, $\varphi(x_1,...,x_n)$ is (total) recursive if and only if there is a system E of equations such that

(1) $(x_1)...(x_n)(EY)\mathfrak{S}_n(E,x_1,...,x_n,Y)$

(2) $(x_1)...(x_n)(Y)\,[\mathfrak{S}_n(E,x_1,...,x_n,Y) \to \mathfrak{U}(Y) = \varphi(x_1,...,x_n)]$.

And then the arithmetic forms of (1) and (2) are

(1*) $(x_1)...(x_n)(Ey)S_n(e,x_1,...,x_n,y)$

(2*) $(x_1)...(x_n)(y)[S_n(e,x_1,...,x_n,y) \to U(y) = \varphi(x_1,...,x_n)]$,

where S_n and U are recursive.

From (1*) and (2*) it follows that $\varphi(x_1,...,x_n)$ can be expressed in the following form

(3*) $\varphi(x_1,...,x_n) = U(\mu y S_n(e,x_1,...,x_n,y))$.

Since (1*), (2*) and (3*) hold if S_n is replaced by T_n [130] the following Theorem holds.

Normal form theorem. *If* $\varphi(x_1,...,x_n)$ $(n \geq 0)$ *is a (total) recursive function, a number e can be found such that*

(4) $(x_1)...(x_n)(Ey)T_n(e,x_1,...,x_n,y)$

(5) $\varphi(x_1,...,x_n) = U(\mu y T_n(e,x_1,...,x_n,y))$

(6) $(x_1)...(x_n)(y)[T_n(e,x_1,...,x_n,y) \to U(y) = \varphi(x_1,...,x_n)]$,

with T_n *the primitive recursive predicate and U the primitive recursive function defined above.*

The number *e* for which (4)-(6) hold is called the *Gödel number* of the (total) recursive function $\varphi(x_1,...,x_n)$.

---

129 Cf. Kleene, [1952], 278, 288.

130 Comp. and (1) and (2) of 4.2.4.1 (above).

(2) **Partial recursive functions**

According to Church's Thesis, every effective calculable function (every effective decidable predicate) is recursive. Together with its converse, this thesis gives a definition of the idea of algorithm (or calculation/ decision procedure) for a number-theoretic function/ predicate. As we saw (4.2.4.2 Intermezzo), if $P(x)$ is such a predicate, then to give an algorithm for it means to give a recursive predicate $R(x)$ such that $P(x) \leftrightarrow R(x)$. By a well-known result, [131] for the predicates $(Ey)T(x,x,y)$ and $(y)\widetilde{T}(x,x,y)$ there exists no algorithm.

Let us now consider an effectively decidable predicate $R(x,y)$ and a procedure for finding the least *y* such that $R(x,y)$, for an *x* fixed. Clearly, such a procedure leads to the sought *y* iff $(Ey)R(x,y)$. And then it can be understood as an algorithm for calculation a number-theoretical function whose domain is the class of such numbers *x* for which $(Ey)R(x,y)$. This function is $\mu yR(x,y)$: the least *y* such that $R(x,y)$. And then $\mu yR(x,y)$ is effectively calculable iff $(Ey)R(x,y)$ is effectively decidable. And since, by the Church's result, for the predicate $(Ey)T(x,x,y)$ there is no algorithm, it follows that it is not recursive (effectively decidable) and therefore $\mu yT(x,x,y)$ is not a recursive function. I.e., there is no algorithm for deciding, for *x* fixed, whether the function $\mu yT(x,x,y)$ is defined or not. Such a function is an example of *partial* recursive function (as an incompletely defined function).

Generally, a *partial function* $\varphi(x_1,...,x_n)$ is a function from any subset, *proper* or *improper*, of *n*-tuples of natural numbers to the natural numbers. If for a given *n*-tuple $x_1,...,x_n$, $\varphi(x_1,...,x_n)$ has a value, then φ is *defined*, otherwise it is *undefined*. The range of definition of a partial function is the set of *n*-tuples $x_1,...,x_n$ for which it is defined. If $\varphi(x_1,...,x_n)$ is defined for all *n*-tuples $x_1,...,x_n$ of natural numbers, then φ is *total*. Hence the partial functions include the total functions, and therefore the partial recursive functions include the total recursive functions. So, if $R(x_1,...,x_n,y)$ is a recursive predicate, then the function $\mu yR(x_1,...,x_n,y)$ is *partial recursive*[132] (it is defined iff $(Ey)R(x_1,...,x_n,y)$). And

[131] Cf. A. Church [1936(b)]; comp. and 4.2.4.2 and 4.2.4.3 (above).

[132] Hence, using μ-operator, from a completely defined predicate $R(x_1,...,x_n,y)$ a partial

$\mu yR(x_1,...,x_n,y)$ is *recursive* iff the following holds: $(x_1)...(x_n)(Ey)R(x_1,...,x_n,y)$.

As in the case of recursive functions, every partial recursive function, obtained *via* I-VI (of Sect. 3.1), is partial recursive.

(a) **Normal form theorem for partial recursive functions**

If $\varphi(x_1,...,x_n)$ is partial recursive, then (1) and (1*) (from the normal form theorem for total recursive functions) do not necessarily hold. So, from (1*), for given $x_1,...,x_n$, only remains $(Ey)S_n(e,x_1,...,x_n,y)$ as the condition for definability of $\varphi(x_1,...,x_n)$. Since φ is partial recursive, (2) and (2*) hold if "=" is replaced with "$\simeq$", where "$\simeq$" is the "complete equality",[133] and this means: $\varphi(x_1,...,x_n)\simeq\psi(x_1,...,x_n)$ if for given $x_1,...,x_n$, if either of φ and ψ is defined, so is the other and have the same value. And in (3*) $U(\mu yS_n(e,x_1,...,x_n,y))$ is defined only if $(Ey)S_n(e,x_1,...,x_n,y)$, and therefore both members in (3*) have the same range of definition, and then (3*) holds for all $x_1,...,x_n$, when "=" is replaced by "$\simeq$".

Theorem.[134] (a) *For any partial recursive function* $\varphi(x_1,...,x_n)$, *a number e can be found such that* $\varphi(x_1,...,x_n)\simeq U(\mu yT_n(e,x_1,...,x_n,y))$ (*where* $(Ey)T_n(x_1,...,x_n,y)$ *is the condition of definition of* φ).

(b) $(x_1)...(x_n)(y)[T_n(e,x_1,...,x_n,y)\rightarrow U(y)\simeq\varphi(x_1,...,x_n)]$.

As in the preceding case, *e* is the *Gödel number of* φ (or its *index*).

(b) **Enumeration theorem for partial recursive functions**[135]

Kleene abbreviates $U(\mu yT_n(z,x_1,...,x_n,y))$ by $\Phi_n(z,x_1,...,x_n)$. As we saw above, $\Phi_n(z,x_1,...,x_n)$ is an $n+1$ partial recursive function. If *z* is fixed, Φ_n is an *n*-place partial recursive function. Whence, by (a) of the normal form theorem for partial recursive functions, if $\varphi_n(x_1,...,x_n)$ is any partial recursive function, then there is a *e* such that

function $\mu yR(x_1,...,x_n,y)$ is constructed.

[133] Cf. S.C. Kleene, [1952], 328.

[134] I.e. normal form theorem for partial recursive functions; cf. S.C. Kleene, [1952], §63, Theorem XIX.

[135] S.C. Kleene, [1952], §65, Theorem XXII.

$$\varphi(x_1,...,x_n) \simeq \Phi_n(e,x_1,...,x_n),$$

where e is the Gödel number of φ, and just this is the idea of enumeration theorem.

Enumeration theorem. *For* $z = 0,1,2,...$, *the* $n+1$*-place partial recursive function* $\Phi(z,x_1,...,x_n)$ *gives an enumeration* (*with repetitions*) *of the n-place partial recursive functions*.

Remark. This theorem does not hold for *total* recursive functions. This can be proved by a diagonal argument. Suppose, for example, that $\Phi(z,x)$ does enumerate the total recursive functions. Let φ be the total recursive function $\Phi(x,x)+1$. Then, by theorem, there were an e such that $\varphi(x) = \Phi(e,x)$ for all x. Hence for $x = e$, $\varphi(e) = \Phi(e,e)$. But $\varphi(e) = \Phi(e,e)+1$. Therefore $\Phi(e,e) = \Phi(e,e)+1$, which is impossible. This contradiction is derived for $z = x = e$, where e is the Gödel number of $\Phi(x,x)+1$ (i.e., of $\varphi(x)$).

For partial recursive functions, instead, the theorem holds, since for some arguments a partial recursive function may be undefined. So $\Phi(e,e) \simeq \Phi(e,e)+1$ is not impossible, since this means that for $z = x = e$, the function $\Phi(z,x)$ must be undefined.

(c) **Diagonal Lemma** (*via* recursion theory)

The Diagonal Lemma can be proved in a variety of ways. As we saw (comp. 4.2.2.1) this lemma can be proved using the Gödel numbering and the diagonal function, $\delta(x)$. A proof of this lemma can also be obtained using the recursion theory, *via* Kleene's function $\Phi(x,y)$.

Diagonal Lemma. *For any formula* β(y) *of* $\mathrm{L_{PA}}$[136] *there is a sentence* G *such that:* $\mathrm{PA^{ax}} \vdash \mathrm{G} \equiv \beta(\overline{g})$, *where g is the Gödel number of* G.

Proof. Let β(y) be any formula of $\mathrm{L_{PA}}$, let $\Phi(x_1,x_2)$ be the Kleene's partial recursive function that enumerates the 1-place partial recursive functions, let $\mathrm{F(x_1,x_2,y)}$ be the formula of $\mathrm{L_{PA}}$ which represents it in $\mathrm{PA^{ax}}$. Then $\mathrm{F(x_2,x_2,y)}$ will represent $\Phi(x_2,x_2)$ in $\mathrm{PA^{ax}}$, i.e., for any m, n:

If $\Phi(m,m) = n$, then $\mathrm{PA^{ax}} \vdash \forall \mathrm{y}(\mathrm{F}(\overline{m},\overline{m},\mathrm{y}) \equiv \mathrm{y} = \overline{n})$, i.e.,

(a) $\mathrm{PA^{ax}} \vdash \forall \mathrm{y}(\mathrm{F}(\overline{m},\overline{m},\mathrm{y}) \supset \mathrm{y} = \overline{n})$, and

(b) $\mathrm{PA^{ax}} \vdash \forall \mathrm{y}(\mathrm{y} = \overline{n} \supset \mathrm{F}(\overline{m},\overline{m},\mathrm{y}))$, equivalently

[136] The result holds for any consistent system $S \supseteq Q$.

$$PA^{ax} \vdash F(\overline{m},\overline{m},\overline{n}).$$

Let *Form*: $\exists y(F(x_2, x_2, y) \wedge \beta(y))$, and $\varphi(\overline{m})$ be the Gödel number of *Form**: $\exists y(F(\overline{m},\overline{m}, y) \wedge \beta(y))$.

Let k be the index of $\varphi(x)$, i.e., $\varphi(x) = \Phi(k,x)$. Let G: $\exists y(F(\overline{k},\overline{k}, y) \wedge \beta(y))$, whose Gödel number is $\varphi(k)$; let $\varphi(k) = g$.

1. $PA^{ax} \vdash G \supset \beta(\overline{g})$

   (1) $PA^{ax} \vdash \forall y(F(\overline{k},\overline{k}, y) \supset y = \overline{g})$; by (a), since $\Phi(k,k) = \varphi(k) = g$

   (2) $PA^{ax} \vdash \forall y(F(\overline{k},\overline{k}, y) \supset y = \overline{g}) \supset [G \supset (F(\overline{k},\overline{k},\overline{g}) \wedge \beta(\overline{g}))]$; FOL^{ax}_{id}

   (3) $PA^{ax} \vdash G \supset (F(\overline{k},\overline{k}, g) \wedge \beta(\overline{g}))$; (1), (2), MP

   (4) $PA^{ax} \vdash G \supset \beta(\overline{g})$; (3) PL

2. $PA^{ax} \vdash \beta(\overline{g}) \supset G$

   (1) $PA^{ax} \vdash F(\overline{k},\overline{k},\overline{g})$; by (b) and Ch. 2, Sect. 4.3, Lemma

   (2) $PA^{ax} \vdash F(\overline{k},\overline{k},\overline{g}) \supset [\beta(\overline{g}) \supset (F(\overline{k},\overline{k},\overline{g}) \wedge \beta(\overline{g}))]$; PL

   (3) $PA^{ax} \vdash \beta(\overline{g}) \supset (F(\overline{k},\overline{k},\overline{g}) \wedge \beta(\overline{g}))$; (1), (2), MP

   (4) $PA^{ax} \vdash (F(\overline{k},\overline{k},\overline{g}) \wedge \beta(\overline{g})) \supset \exists y(F(\overline{k},\overline{k}, y) \wedge \beta(y))$; FOL

   (5) $PA^{ax} \vdash \beta(\overline{g}) \supset \exists y(F(\overline{k},\overline{k}, y) \wedge \beta(y))$; (3), (4), PL

   i.e., $PA^{ax} \vdash \beta(\overline{g}) \supset G$.

The Diagonal Lemma follows from 1 and 2, by PL.

Remark. In the formulation of DL β(y) is a formula containing only one free variable y. But DL holds for formulas β containing more free variables and then G will be a formula with free variables.[137] If β(y,z) is a formula of $S \supseteq Q$, for example, then by DL there is a formula G(z) of *S* such that $S \vdash \forall z(G(z) \equiv \beta(\overline{g}, z))$, where g is the Gödel number of G(z).

4.2.5. Recursive enumerability and incompleteness

As we saw above, Kleene's generalized forms of Gödel's and Rosser's theorems are essentially based on a fundamental result of recursion theory: Enumeration Theorem. By this theorem the predicate $(Ey)T(z,x,y)$ is a rec. en. 2-place relation which enumerates the rec. en. predicates of the form $(Ey)R(x,y)$ (with R recursive). The formulation of this result is in no way unique. The relation which enumerates rec. en. relations depends on the

[137] Comp. Sect. 4.2.2.2 footnote 78.

formalism adopted for the developing the recursion theory, and then for different formalisms this relation will be different. Kleene's predicate $T(z,x,y)$, for example, can be interpreted in terms of Turing machines. It is a primitive recursive 3-place relation[138] which can express the following fact: "y is the Gödel number of a terminating computation at input x on the Turing machine with Gödel number z". Since $T(z,x,y)$ is recursive, by prefixing it with (Ey) the result will be the rec. en. relation (or Σ_1-relation)[139] $(Ey)T(z,x,y)$, expressing the following fact: "the Turing machine with Gödel number z halts at input x". And then the relation $(y)\widetilde{T}(z,x,y)$ means "the Turing machine z does not halt at input x". So, the sentence $(y)\widetilde{T}(k,k,y)$ says, simply, that "the Turing machine k does not halt at input k".[140] So, if the recursion theory is developed in terms of Turing machines, then the 2-place relation which enumerates the rec. en. sets, for example, is the following: "z codes a Turing machine which gives the output "yes" on the input x".

Another strategy for constructing the rec. en. relation which enumerates rec. en. sets/ relations is that of Smullyan.[141] It takes as formalism for presenting the recursion theory the formal system Q, this strategy being motatívated by the fact that *the rec. en. relations are exactly those relations representable in Q*.[142]

Some other way to define the rec. en. relation which enumerates the rec. en. sets/ relations is that of Kripke.[143] Here the rec. en. enumerating relation $W(e,n)$ is the relation $Sat(e,n)$, where e is the Gödel number of a formula $\alpha(\mathrm{x})$ (with x free) and n satisfies α.[144] More exactly, if $\alpha(\mathrm{x})$ is a formula of $\mathrm{L_{PA}}$ whose Gödel number is e (symbolically, $\alpha_e(\mathrm{x})$) then $Sat(e,n)$ iff $\alpha_e(\overline{n})$ is true in M (where M is the standard model for $\mathrm{L_{PA}}$). In order to avoid all the problems concerning the presence of semantic notions in the language of arithmetic Kripke choses a fragment of $\mathrm{L_{PA}}$, the language *RE* (it is a Σ_1-language, containing only atomic Σ_1-formulas, the connectives $\wedge$ and

---

[138] and then a *decidable* relation (cf. 4.2.4.2, Intermezzo).
[139] If R is recursive, then $(Ey)R$ is rec. en.; for proofs comp. e.g. R. Smullyan [1992], Ch. 4, §1; M. Davis [1958], 66-67, Theorem 1.2-Theorem 1.4; H. Rogers [1967], 66 Corollary XI.
[140] A short comparison with Gödel's Theorem (III) (Sect. 4.2.4.4) will be instructive.
[141] Cf. R. Smullyan [1993], Ch. III, §1.
[142] Or R or $\mathrm{PA^{ax}}$ (if $\mathrm{PA^{ax}}$ is consistent).
[143] Cf. S. Kripke [1996], Lecture VII.
[144] Evidently, this definition can be extended to formulas with n free variables.

$\vee$, $\exists x\alpha$ and $\forall x(x < t \supset \alpha)$ (where t does not contain x). Note that it does not contain negation, implication or unbounded universal quantifier). In such a language a formula $\alpha_e(x)$ is said to *define* a set *S* if for any *n*: $n \in S$ iff *Sat*(*e*,*n*) iff $\alpha_e(\bar{n})$ is true (in *M*). This choice is motivated by the following fact: *a relation R^n is recursively enumerable iff it is defined in the language RE*.

Besides the Kleene's result concerning Enumeration Theorem we also mentioned Smullyan's and Kripke's since we'll use these strategies in what follows in exposing the subject of recursive enumerability and incompleteness.

To begin with, following Smullyan, let us see how the celebrated Gödel's result can be derived *via* recursive enumerability.

First of all, let us define the Σ_1-formulas and Σ_1-relations and present some general facts implied in Smullyan's account.

1. Σ_1-formulas and relations

Definition. *A formula of* L_{PA} *is called* Σ_1 *if it is a member of the smallest class that contains all the formulas of the following forms:*

(a) $x_1 = x_2$, $0 = x_1$, $x_1' = x_2$, $x_1 + x_2 = x_3$, $x_1 \cdot x_2 = x_3$

(b) $\alpha \wedge \beta$; $\alpha \vee \beta$, $\exists x\alpha$, $\forall x(x < y \supset \alpha)$, *whenever* α *and* β *are members of this class.*

Remarks.

(1) A Σ_1-formula begins with *one* unbounded existential quantifier, all the other quantifiers of this formula being bounded. There are formulas containing *many* unbounded existential quantifiers but in which all universal quantifiers are bounded. These formulas are usually called Σ-formulas. But, as can be shown, the Σ-formulas are provably equivalent to Σ_1-formulas.

(2) The formulas from (a) are called *atomic* formulas. All these formulas are Σ_1, since they are equivalent to a formula constructed from them using existential quantification and conjunction. For example, $x + y' = z$ can be written as $\exists u \exists z(y' = u \wedge x + u = z)$.

(3) If α is a Σ_1-formula, then $\exists x(x < y \wedge \alpha)$ is a Σ_1-formula; and then the

class of Σ_1-formulas is closed under *bounded* quantification (universal and existential). But if α is Σ_1 *is not always* the case that its negation, $\neg\alpha$, is also a Σ_1-formula!

(4) The class of Σ_0-formulas contains the atomic formulas (list (a)), the formulas $\neg\alpha$, $\alpha \wedge \beta$, $\alpha \vee \beta$ whenever α and β are Σ_0 and the formulas $\forall x(x < y \supset \alpha)$, $\exists x(x < y \wedge \alpha)$ if α is Σ_0.

Definition. *A relation (or set) is Σ_1 iff it is expressible (definable) by a Σ_1-formula.*

Since a Σ_1-formula α begins with one unbounded existential quantifier and all the other quantifiers are bounded, such a formula has the form α: $\exists x_{n+1}\beta(x_1,...,x_n,x_{n+1})$, where $\beta(x_1,...,x_n,x_{n+1})$ is a Σ_0-formula.

Therefore, a relation $R(x_1,...,x_n)$ is Σ_1 iff there is a Σ_1-formula $\alpha(x_1,...,x_n)$ of L_{PA} that expresses it, i.e., such that for any numbers $k_1,...,k_n$ holds: $R(k_1,...,k_n)$ holds iff $\alpha(\bar{k}_1,...,\bar{k}_n)$ is true (in M[145]).

Accordingly, a relation is called Σ_0 iff it is definable by a Σ_0-formula.[146]

The Σ_1-relations are also called *recursively enumerable relations*.[147]

Definition. *A relation R is recursive iff it and its complement are rec. en.*[148]

As we saw above, a Σ_1-formula of L_{PA} is a formula of the form $\exists x\alpha$, where α is a Σ_0-formula (and then decidable). Now, a Π_1-formula of L_{PA} is a formula of the form $\forall x\alpha$, where α is a Σ_0-formula (and then decidable).

A formal system *S* is Σ_1-*sound* if any Σ_1-sentence (closed formula) of L_S, provable in *S*, is true (in *M*). And it is Σ_1-*complete* if the converse holds. *S* is called Π_1-*sound* if any Π_1-sentence of L_S, provable in *S*, is true (in *M*). A system *S* is 1-*consistent* if there is no decidable formula $\alpha(x)$ such that both $S \vdash \neg\alpha(\bar{n})$, for every *n*, and $S \vdash \exists x\alpha(x)$. Evidently, the following

[145] Where *M* is the standard model of L_{PA}.

[146] The Σ_0-relations are also called *constructive arithmetic relations*.

[147] In what follows we use "Σ_1-relation" and "recursively enumerable relation" interchangeably; for them we also use abbreviated *rec. en*. They are also named *semi-recursive relations* (cf. Boolos, Burgess and Jeffrey [2002], Sect. 7.2) or semi-computable relations (cf. M. Davis [1958], Ch. 5, S. Kripke [1996], 10).

[148] The above definitions for "rec. en." and "recursive relations" are not unique. Equivalent definitions can be found in Kleene [1952], Turing [1936-37], Post [1944], Smullyan [1961], [1993], Markov [1961], Rogers [1967]. If you want a *proof* for the equivalence of this definitions, comp. *inter alia* Boolos, Burges and Jeffrey [2002], 82 (Propoz. 7.16); M. Davis [1958], 67 (Theorem 1.5); H. Rogers [1967], 58 (Theorem II).

holds: If S is ω-consistent, then S is 1-consistent and therefore S is consistent.

A formal system S is called *recursively axiomatizable*[149] (axiomatizable, for short) if the set *Th* of Gödel numbers of its theorems is rec. en.

A set S is *representable* in a formal system S if there is a formula $\alpha(x)$ of L_S such that for any n:

$n \in S$ iff $S \vdash \alpha(\bar{n})$.

2. Some facts

Fact 1. The class of rec. en. relations is closed under existential quantifications, boolean operations $\wedge$ and $\vee$, and bounded quantifications (universal and existential).[150]

Fact 2. If $f(x,y)$ is a recursive function and $R(x)$ is a rec. en. relation, then $R^*(f(x,y))$ is rec. en.

Proof. $R^*(f(x,y))$ holds iff $(Ez)(f(x,y)=z \wedge R(z))$. By Fact 1 it follows that R^* is rec. en.

Remark. If $R(x)$ is the rec. en. set S[151] and $f(x)$ is recursive, it follows, by Fact 2, that the relation $f(x) \in S$ is an rec. en. relation.

If in Fact 2 R is recursive, then R^* is also recursive (argue!).

Fact 3. If $f(x_1,...,x_n)$ is recursive and $R(x_1,...,x_n)$ is rec. en., then the set $S = \{x \mid f(x_1,...,x_n) = x \wedge R(x_1,...,x_n)\}$ is an rec. en. set.

Proof. $x \in S$ iff $E(x_1)...(Ex_n)(R(x_1,...,x_n) \wedge x = f(x_1,...,x_n))$. So, by Fact 1, S is rec. en.

Fact 4. (1) If $R(x_1,...,x_n,y_1,...,y_m)$ is rec. en. and $k_1,...,k_m$ are arbitrary numbers, then the n-place relation $R(x_1,...,x_n,k_1,...,k_m)$ is rec. en.

Since, for example, if $R(x,y)$ is rec. en., then $R(k,y)$ is also rec. en., since $R(k,y)$ is the relation $\exists z(z = k \wedge R(z,y))$, plus Fact. 1.

(2) If $f(x_1,...,x_n,y_1,...,y_m)$ is recursive and $k_1,...,k_m$ are arbitrary

---

[149] Or "formal system", Σ_1-system, respectively.

[150] This fact can be argued using the above considerations on Σ_1-formulas; but for a detailed proof, comp. R.M. Smullyan, [1992], 50-53. The facts 1-6 are basic facts in the recursion theory. The form given here is that of Smullyan [1993], §§1-3.

[151] Since the sets are 1-place relations.

numbers, then the n-ary function $f(x_1,...,x_n,k_1,...,k_m)$ is recursive. (Similarly if the numbers $m_1,...,m_n$ replace the variables $x_1,...,x_n$) (by Substitution).

Fact 5. Let $R(x_1,...,x_m,y_1,...,y_n)$ $(m,n>0)$ be an rec. en. relation. Then there is a rec. en. relation $Q(x,y)$ such that for any $x_1,...,x_m,y_1,...,y_n$ holds:

$$R(x_1,...,x_m,y_1,...,y_n) \leftrightarrow Q(J_m(x_1,...,x_n),J_n(y_1,...,y_n)).$$

Proof. Using J-functions (comp. Sect. 4.1), define the following equivalence:

$Q(x,y)$ iff $x=J_m(x_1,...,x_m)$, $y=J_n(y_1,...,y_n)$ and $R(x_1,...,x_m,y_1,...,y_n)$.

I.e., $Q(x,y)$ is the following rec. en. relation:

$$(Ex_1)...(Ex_m)(Ey_1)...(Ey_n)(x=J_m(x_1,...,x_m)\wedge$$
$$\wedge\, y=J_n(y_1,...,y_n)\wedge R(x_1,...,x_m,y_1,...,y_n)).$$

Fact 6. If $\alpha(x_1)$ is a formula containing x_1 free and y is its Gödel number, then the Gödel number of the formula $\alpha(\bar{x})$, where x is any number, depends of y and x, and then it is $r(y,x)$, where $r(y,x)$ is a recursive function.

Argument. As can be seen, $r(y,x)$ is the Gödel number of the formula derived from the formula with Gödel number y by substituting all free occurrences of the variable x_1, with, say, Gödel number 2^{15},[152] with the numeral for x. The arithmetic form of this operation is the recursive function $Sub(y,Num(x),2^{15})$, i.e., just $r(y,x)$.

Another way of arguing (in order to avoid substitution) is that adopted by Smullyan.[153] It is based on the Tarski's remark[154] that for $\alpha(x_1)$ and any number n, the following equivalence holds:

$$\alpha(\bar{n}) \equiv \forall x_1(x_1=\bar{n}\supset\alpha(x_1))$$ [155]

The right member of this equivalence is abbreviated by $\alpha[\bar{n}]$, and then $\vdash\alpha(\bar{n})\equiv\alpha[\bar{n}]$. In this case, if y is the Gödel number of $\alpha(x_1)$, then $r(y,n)$ is the Gödel number of $\alpha[\bar{n}]$, and can be uniquely determined, using the

[152] Comp. Sect. 4.1. Remember, the Gödel number of a symbol, e.g. $g(x_1)=15$ is different from the Gödel number of an *expression* consisting only of the respective symbol, i.e., 2^{15} (cf. Sect. 4.1).

[153] Comp, R. Smullyan, [1992], Ch. II, §6.

[154] Comp. A Tarski, A. Mostowski, R. Robinson [1953], 44.

[155] Comp. Ch. 2, 4.3, Lemma.

primitive recursive function of concatenation.[156]

Remark. Since $r(y,x)$ is recursive, $r(x,x)$ is also recursive and it is just the *diagonal function*.[157]

Fact 7. Lemma.[158] (1) PA^{ax} is Σ_1-complete.
(2) PA^{ax} is 1-consistent iff PA^{ax} is Σ_1-sound.
(3) PA^{ax} is consistent iff PA^{ax} is Π_1-sound.

Proof. (1) Let *S* be a true Σ_1-sentence, i.e., *S* has the form $\exists x\alpha(x)$ with $\alpha(x)$ decidable. Being true, there is a number *n* such that $\alpha(\overline{n})$ is true (in *M*). And then $PA^{ax} \vdash \alpha(\overline{n})$; therefore $PA^{ax} \vdash \exists x\alpha(x)$ (by Gen$\exists$).

(2) We have to prove: If PA^{ax} is 1-consistent, then [if $PA^{ax} \vdash \exists x\alpha(x)$, then $\exists x\alpha(x)$ is true (in *M*)]; equivalently: If PA^{ax} is 1-consistent and $PA^{ax} \vdash \exists x\alpha(x)$, then $\exists x\alpha(x)$ is true (in *M*).

Reductio. Assume the antecedent of the conditional and that $\exists x\alpha(x)$ is false. It follows that for any *n*, $\alpha(\overline{n})$ is false, and then for any *n*, $PA^{ax} \vdash \neg\alpha(\overline{n})$. But since $PA^{ax} \vdash \exists x\alpha(x)$ (by assumption), it follows that PA^{ax} is 1-inconsistent. The converse follows from the definition of 1-consistency.

(3) We must prove: If PA^{ax} is consistent, then [if $PA^{ax} \vdash \forall x\alpha(x)$, then $\forall x\alpha(x)$ is true (in *M*)]; equivalent: If PA^{ax} is consistent and $PA^{ax} \vdash \forall x\alpha(x)$, then $\forall x\alpha(x)$ is true (in *M*).

Reductio. Assume the antecedent and that $\forall x\alpha(x)$ is false. Then there is an *n* such that $\alpha(\overline{n})$ is false. It follows that $PA^{ax} \vdash \neg\alpha(\overline{n})$. But since $PA^{ax} \vdash \forall x\alpha(x)$ follows that for *any n*: $PA^{ax} \vdash \alpha(\overline{n})$ (by Ax4, MP), contradicting the consistency of PA^{ax}. The converse is immediate.

Remark. *Via* this Lemma, the Gödel's Theorem can be expressed in the following terms:

Gödel's Theorem. 1. *If* PA^{ax} *is* Π_1-*sound, then* $PA^{ax} \nvdash G$.
2. *If* PA^{ax} *is* Σ_1-*sound, then* $PA^{ax} \nvdash \neg G$.

---

[156] Smullyan's argument for the recursivity of the function $r(y,x)$ is based on the following idea: $r(y,x) = z$ is an rec. en. relation (cf. Smullyan [1992], 25) and then $r(y,x)$ is recursive (by the following result of recursion theory: a function $f(x_1,...,x_n)$ is recursive iff the relation $f(x_1,...,x_n) = y$ is recursive iff $f(x_1,...,x_n) = y$ is rec. en.; cf. Smullyan [1993], 10; comp. also H. Rogers [1967], 65, Theorem IX(a)).

[157] Comp. the item 9 from the list at the end of Sect. 4.1.

[158] This result holds for any extension of PA^{ax}.

As we know, the sentence G has the form $\forall x\alpha(x)$ with $\alpha(x)$ decidable, hence it is a Π_1-sentence.

3. Theorems

Theorem 1.[159] *If* S *is an axiomatizable system and the set S is representable in* S, *then S is rec. en.*

Proof. Let $\alpha(x_1)$ be the formula representing S in S, let y be its Gödel number. Then for any x:

$$x \in S \text{ iff } \vdash \alpha_y[\bar{x}].$$

The Gödel number of the formula $\alpha_y[\bar{x}]$ is, evidently, a function on y and x, i.e., $r(y,x)$, where r is recursive (by Fact 6 above). And then we have

$$x \in S \text{ iff } r(y,x) \in Th.$$

Now, since Th is rec. en. (since S is axiomatizable), it follows that $r(y,x) \in Th$ is also rec. en. (by Fact 2, above). Given $\alpha(x_1)$ with Gödel number k, the Gödel number of $\alpha[\bar{x}]$ only depends on x, and will be $r(k,x)$. Since $r(y,x)$ is recursive, so is $r(k,x)$ (by Fact 4 above). And then $x \in S$ iff $r(k,x) \in Th$. Since $S = r^{-1}(Th)$, it follows that S is rec. en.

Theorem 2.[160] *If* S $\supseteq Q$ *is any consistent axiomatizable system and S is rec. en. then S is representable in* S.

By Theorem 1 and Theorem 2, if S $\supseteq Q$ is consistent and axiomatizable then the rec. en. relations are *exactly* the representable relations in S.

As we saw by Kleene's Enumeration Theorem (comp. 4.2.4.2) the predicate $(Ey)T(z,x,y)$ enumerates the predicate of the form $(Ey)R(x,y)$ with R recursive. In other words, the rec. en. relation $(Ey)T(z,x,y)$ enumerates the rec. en. sets $S = \{x \mid (Ey)R(x,y)\}$.[161] This idea of enumeration can also be given, in a slight different form, in the following terms: by taking $W(z,x)$ be an rec. en. relation enumerating the rec. en. sets in the way indicated by the two theorems above. Respectively, if S $\supseteq Q$ is

[159] With minor modifications this is the Theorem 1 in Smullyan [1993], 35.

[160] Due to J. Shepherdson [1961], A. Ehrenfeucht and S. Feferman [1960]. An elegant proof is given in R. Smullyan [1992], Ch. VII, §1, or [1993], 85.

[161] The sets of natural numbers can be taken as special cases of predicates/relations (they are predicates/relations of one argument).

consistent and axiomatizable, then the sets rec. en. are exactly the sets represented in S.[162]

Enumeration Theorem.[163] *There is a 2-place rec. en. relation such that for any rec. en. set S there is an z such that* $S = \{x \mid W(z,x)\}$.

Proof. Let $S \supseteq Q$ be any consistent axiomatizable system. Define $W_z = \{x \mid \vdash \alpha_z[\bar{x}]\}$, where *z* is the Gödel number of the formula $\alpha(x_1)$. Then the Gödel number of $\alpha_z[\bar{x}]$ depends on *z* and *x*; let $r(z,x)$ be this number, where $r(z,x)$ is recursive (by Fact 6 above). Since S is axiomatizable, the set *Th* (of Gödel numbers of theorems of S) is rec. en. And since $r(z,x)$ is recursive it follows that the relation $r(z,x) \in Th$ is rec. en. (by Fact 2 above). But this means nothing else than $\vdash \alpha_z(\bar{x})$, i.e., $x \in W_z$, a relation we write symbolically as $W(z,x)$.

On the other hand, if *S* is a rec. en. set, then it is representable in S by a formula $\alpha(x_1)$ with, say, *z* its Gödel number. Hence we set $S = W_z$ for this *z*, and therefore *S* appear in the enumeration.

The number *z* is called the *index* of *S* in the enumeration and then we write equivalently $S = W_z = \{x \mid W(z,x)\}$.

For the *n*-place $(n \geq 2)$ rec. en. relations the enumeration can proceed similarly, i.e., by defining $W_z^n = \{\langle x_1,...,x_n\rangle \mid \vdash \alpha_z(\bar{x}_1,...,\bar{x}_n)\}$ and arguing as above.

Alternatively, the enumeration of *n*-place rec. en. relations can be given using the pairing function $J(x,y)$ and its corresponding forms for $n \geq 2$, i.e., $J_n(x_1,...,x_n)$ (comp. Sect. 4.1). In this case we simply define an $n+1$-place rec. en. relation W^{n+1} in the following way:

$W^{n+1}(z,x_1,...,x_n)$ iff $W^2(z,u)$, where $u = J_n(x_1,...,x_n)$, and thus this case is reducible to the case for sets.

---

[162] As we mentioned, this is Smullyan's strategy. Originally, the notation *W*(*e*,*x*) is Kleene's. It also appears in Rogers [1967], 68 (Theorem XII), Kripke [1996] and is also used in the present account.

[163] Essentially, this is Kleene's Enumeration Theorem, but uses a different schema of indexing the rec. en. sets (comp. R. Smullyan [1993], Ch. 3, §1).

4. Recursive enumerability and incompleteness

With the apparatus of recursive enumerability the Gödel's Theorem can be given in following terms.

Theorem 1.[164] *If* S *is a consistent and axiomatizable system and S is a rec. en. and non-recursive set, then* S *is incomplete.*

Proof. Let $\alpha(x)$ be the formula representing S in S. Let S^* be the set represented by $\neg\alpha(x)$. Both sets, S and S^*, are rec. en. (since S is axiomatizable), but since S is not recursive S^* is not the complement of S (i.e., $S^* \neq \widetilde{S}$). It follows that there is a number n such that $n \in \widetilde{S}$, and then $n \notin S$, and $n \notin S^*$. This means that $\nvdash \alpha(\overline{n})$ and $\nvdash \neg\alpha(\overline{n})$. Therefore $\alpha(\overline{n})$ is undecidable in S.

Another form of Gödel's incompleteness theorem can also be given under the stronger assumption of soundness of S.

Theorem 2.[165] *Let* S *be an aziomatizable system, let S be a non rec. en. set and* $\beta(x)$ *be the formula expressing it in* L_S. *If* S *is sound, then there is a formula* $\beta(\overline{n})$ *true and not provable in* S.

Proof. Let $S_0 = \{n \mid \vdash \beta(\overline{n})\}$, i.e., $\beta(x)$ represents S_0 in S and then S_0 is rec. en. (since S is axiomatizable). Now, since $\beta(x)$ expresses S in L_S and S is not rec. en., it follows that there is a number n: $n \in S$ and $n \notin S_0$; and since S is sound, $S_0 \subseteq S$. For this n we have: $\beta(\overline{n})$ is true and $\beta(\overline{n})$ is not provable in S. Since $\beta(\overline{n})$ is true, its negation $\neg\beta(\overline{n})$ is false, and therefore not provable in S (by soundness of S).

Remark. As can be seen this theorem is a form of Kleene's Theorem XIII (II) (comp. Sect. 4.2.4.4). In a sense, it is more general, since S in this proof is *any* not rec. en. set and not just the *anti-diagonal set* $\widetilde{K} = \{x \mid (y)\widetilde{T}(x,x,y)\}$. But, in contrast to Kleene's proof, the above proof is not constructive.[166]

Now, using the notions of recursive enumerability and the Fact 7 (*Lemma*), the following form of Gödel's Theorem can be given.[167]

---

[164] This is *Theorem 7* of Smullyan [1993], 41.

[165] Comp. and Smullyan [1993], 42, *Theorem 8*.

[166] I.e., it asserts the *existence* of a true but not provable formula of L_S, but gives no method how to discover it.

[167] This also imitates Kleene's Theorem XIII (II).

Theorem 3. *Let* $S \supseteq Q$ *any axiomatizable* ω*-consistent system. Then there is a true but undecidable formula of* L_S.

Proof. Let W^2 be a 2-place rec. en. relation enumerating the rec. en. sets. Let ∃yβ(z,x,y) the Σ_1-formula expressing it in L_S.

Let us define the following sets:

(1) $K = \{x \mid \exists y\beta(\bar{x}, \bar{x}, y)\}$; equivalent: $x \in K$ iff $\exists y\beta(\bar{x}, \bar{x}, y) = 1$

(2) $\widetilde{K} = \{x \mid \forall y\neg\beta(\bar{x}, \bar{x}, y)\}$; equivalent $x \in \widetilde{K}$ iff $\forall y\neg\beta(\bar{x}, \bar{x}, y) = 1$

(3) $\widetilde{K}_0 = \{x \mid \vdash \forall y\neg\beta(\bar{x}, \bar{x}, y)\}$; equivalent $x \in \widetilde{K}_0$ iff $\vdash \forall y\neg\beta(\bar{x}, \bar{x}, y)$.

As can be seen, ∃yβ(x,x,y) expresses *K* in L_S and its negation, ¬∃yβ(x,x,y), equivalent ∀y¬β(x,x,y), expresses $\widetilde{K}$ in L_S and represents the set $\widetilde{K}_0$ in S. Hence *K* and $\widetilde{K}_0$ are rec. en. Being rec. en., $\widetilde{K}_0$ has an index *e* in enumeration and then

(4) $\widetilde{K}_0 = \{x \mid \exists y\beta(\bar{e}, \bar{x}, y)\}$; equival. $x \in \widetilde{K}_0$ iff $\exists y\beta(\bar{e}, \bar{x}, y) = 1$.

Therefore, for any *x*:

$x \in \widetilde{K}_0$ iff $\exists y\beta(\bar{e}, \bar{x}, y) = 1$ (by (4)) iff $\vdash \forall y\neg\beta(\bar{x}, \bar{x}, y)$ (by (3)).

Whence for $x = e$:

(*Eq*) $e \in \widetilde{K}_0$ iff $\exists y\beta(\bar{e}, \bar{e}, y) = 1$ iff $\vdash \forall y\neg\beta(\bar{e}, \bar{e}, y)$.

Now, since S is ω-consistent, it is consistent and therefore Π_1-sound (by Fact 7). And then if $\forall y\neg\beta(\bar{e}, \bar{e}, y)$ were provable then it will be true. But this is impossible, since by (*Eq*) its negation, $\exists y\beta(\bar{e}, \bar{e}, y)$ is also true. Hence $\forall y\neg\beta(\bar{e}, \bar{e}, y)$ is not provable in S. Whence, by (*Eq*), $e \notin \widetilde{K}_0$ and therefore $\exists y\beta(\bar{e}, \bar{e}, y) = 0$, hence $\forall y\neg\beta(\bar{e}, \bar{e}, y) = 1$.

On the other hand, since S is ω-consistent, (by Fact 7) it is Σ_1-sound and therefore $\exists y\beta(\bar{e}, \bar{e}, y)$, being false, it is not provable in S.

Therefore, $\forall y\neg\beta(\bar{e}, \bar{e}, y)$ is a formula of L_S true but *undecidable* in S.

4.2.6. Separability, Recursive Inseparability, Incompleteness and Undecidability

The ideas of separability and recursive inseparability give new ways of proving the incompleteness and undecidability of formal systems.

1. Separability in S

Definition 1. *Let* S *be a formal system,* S_1 *and* S_2 *be two number-theoretic sets. A formula* α(x) *of* Ls ***weakly separates*** S_1 *from* S_2 *in* S *if the following holds:*

(a) *If* $n \in S_1$, *then* $\vdash \alpha(\overline{n})$.

(b) *If* $n \in S_2$, *then* $\nvdash \alpha(\overline{n})$.

Lemma 1. *If* α(x) *weakly separates* S_1 *from* S_2 *in* S *and* S *is consistent, then* α(x) *represents some superset*[168] S_1' *of* S_1 *disjoint from* S_2.

Proof. Let S_1' be the set represented in S by the formula α(x), i.e.,

(*) $n \in S_1'$ iff $\vdash \alpha(\overline{n})$.

Then $S_1 \subseteq S_1^*$ (by (a) and (*)). Moreover, $S_1' \cap S_2 = \varnothing$ (i.e., S_1' and S_2 are disjoint), since if there were an *n* such that $n \in S_1'$ and $n \in S_2$, then $\vdash \alpha(\overline{n})$ (by (*)) and $\nvdash \alpha(\overline{n})$ (by (b)); contradiction.

Definition 2. *Let* S *be a formal system,* S_1 *and* S_2 *be two number-theoretic sets. A formula* α(x) *of* Ls ***strongly separates*** S_1 *from* S_2 *if the following holds:*

(a) *If* $n \in S_1$, *then* $\vdash \alpha(\overline{n})$.

(b) *If* $n \in S_2$, *then* $\vdash \neg\alpha(\overline{n})$.

Lemma 2. *If* α(x) *strongly separates* S_1 *from* S_2 *in* S *and* S *is consistent, then*

(1) α(x) *represents some superset* S_1' *of* S_1 *disjoint from* S_2.

(2) ¬α(x) *represents some superset* S_2' *of* S_2 *disjoint from* S_1.

Proof. (1) Firstly, as in Lemma 1, let S_1' be the set represented in S by α(x), i.e.,

(*) $n \in S_1'$ iff $\vdash \alpha(\overline{n})$.

Then, evidently $S_1 \subseteq S_1'$. Moreover, $S_1' \cap S_2 = \varnothing$; since if there were an *n* such that $n \in S_1'$ and $n \in S_2$, then we would have $\vdash \alpha(\overline{n})$ (by (*)) and $\vdash \neg\alpha(\overline{n})$ (by (b)), and therefore S would be inconsistent.

(2) Let S_2' be the set represented by ¬α(x) in S, i.e.,

(**) $n \in S_2'$ iff $\vdash \neg\alpha(\overline{n})$.

---

[168] *S* is a superset of *T* if *T* is a subset of *S*; i.e., $T \subseteq S$.

Obviously, $S_2 \subseteq S_2'$ (by (b) and (**)). And $S_2' \cap S_1 = \varnothing$; otherwise there would be an n such that $n \in S_2'$ and $n \in S_1$. Whence by (**) and (a) we derive: $\vdash \neg\alpha(\bar{n})$ and $\vdash \alpha(\bar{n})$, contradicting the consistency of S.

Remark. If S is consistent, then any formula which strongly separates S_1 from S_2 in S also weakly separates S_1 from S_2 in S.

Notations.[169] If S is a number-theoretic set, then S^* is defined as follows:

(Eq) $n \in S^*$ iff $\delta(n) \in S$,

where $\delta(x)$ is the diagonal function. Intuitively, if n is the Gödel number of a formula, then $\delta(n)$ is the Gödel number of its diagonalization. As can be observed, $S^* = \delta^{-1}(S)$.

Let P be the set of Gödel numbers of *provable* formulas of a system S, and R be the set of Gödel numbers of *refutable* formulas of S.

Theorem.[170] *If* α(x), *with Gödel number k, weakly separates* R^* *from* P^* *in* S, *then* $\alpha(\bar{k})$ *is undecidable in* S.

Proof. Since α(x) weakly separates R^* from P^* in S, then for any n, and therefore also for k:

(a) If $k \in R^*$, then $\vdash \alpha(\bar{k})$.

(b) If $k \in P^*$, then $\nvdash \alpha(\bar{k})$.

Now, if $\alpha(\bar{k})$ were provable, then $\delta(k) \in P$ and therefore $k \in P^*$ (by (Eq)). But by (b), if $k \in P^*$, then $\nvdash \alpha(\bar{k})$. So, a contradiction is derived. Hence $\alpha(\bar{k})$ *is not provable in* S.

On the other hand, if $\neg\alpha(\bar{k})$ were provable, then $\delta(k) \in R$ and therefore $k \in R^*$ (by (Eq)). Whence, by (a), $\alpha(\bar{k})$ would be provable; again a contradiction. So, under hypothesis of theorem, $\alpha(\bar{k})$ is undecidable in S.

Corollary. *If* α(x), *with Gödel number k, strongly separates* R^* *from* P^*, *in* S *and* S *is consistent, then* $\alpha(\bar{k})$ *is undecidable in* S.

Proof (by Theorem 1 and Remark above).

[169] These notations are borrowed from R. Smullyan, [1992], Ch. I.
[170] Cf. R. Smullyan [1993], 7, *Theorem 3*.

2. Recursive inseparability

Definition 1. *Two disjoint sets* S_1 *and* S_2 *are called* ***recursively separable*** *if there is a recursive set S such that* $S_1 \subseteq S$ *and* $S_2 \subseteq \widetilde{S}$. *S is said to separate* S_1 *and* S_2.

Definition 2. *The sets* S_1 *and* S_2 *are called* ***recursively inseparable*** (*abrev. rec. insep.*) *iff they are disjoint and not recursively separable.*

Definition 3. *A formal system* S *is called recursively inseparable* (*rec. insep., for short*) *if the sets P and R are rec. insep.*

Theorem 1. *Let* $\mathrm{S} \supseteq Q$ *be a consistent formal system extending Q, let S be a recursive set such that* $R^* \subseteq S$. *Then there is a k such that* $k \in S$ *and* $k \in P^*$.

Proof. Since S is recursive, it is formally expressible in S. Let $\alpha(x)$, with Gödel number k, be the formula expressing it in S, i.e., for any n, and then also for k:

(a) If $k \in S$, then $\vdash \alpha(\bar{k})$.

(b) If $k \notin S$, then $\vdash \neg\alpha(\bar{k})$.

Now, suppose $k \notin S$, then $\vdash \neg\alpha(\bar{k})$ (by (b)). But $\vdash \neg\alpha(\bar{k})$ iff $\delta(k) \in R$ iff $k \in R^*$. This is impossible, since $R^* \subseteq S$. So k must be in S. Whence, by (a), $\vdash \alpha(\bar{k})$. But $\vdash \alpha(\bar{k})$ iff $\delta(k) \in P$ iff $k \in P^*$. Therefore, $k \in S$ and $k \in P^*$.

Theorem 2. *If* S *is consistent, then* S *is rec. insep.*

Proof. Since S is consistent, the pairs (P, R) and (P^*, R^*) are disjoint. Now, suppose (for *reductio*) that P^* and R^* are recursively separable. Then there is a recursive set S such that $R^* \subseteq S$ and $P^* \subseteq \widetilde{S}$. But since S is recursive and $R^* \subseteq S$ it follows that there is a k such that $k \in S$ and $k \in P^*$ (by Theorem 1). Since $P^* \subseteq \widetilde{S}$, $k \in \widetilde{S}$. So $k \in S$ and $k \in \widetilde{S}$. Impossible! Hence P^* and R^* are rec. insep. And then P and R are rec. insep. (since P and R are disjoint, $\delta(x)$ is recursive and $P^* = \delta^{-1}(P)$ and $R^* = \delta^{-1}(R)$ are rec. insep.) (detail!).

3. Incompleteness and Undecidability

Theorem.[171] *Let* S *be any consistent axiomatizable system, let* (S_1,S_2) *be a rec. insep. pair of sets. If* (S_1,S_2) *is strongly separable in* S, *then* S *is incomplete.*

Proof. Assume hypothesis. If α(x) is a formula that strongly separates (S_1,S_2) in S, then, by Lemma 2, α(x) represents some superset S_1' of S_1 in S and ¬α(x) represents some superset S_2' of S_2 in S. The sets S_1' and S_2' are disjoint (since S is consistent), and are rec. en. (since S is axiomatizable). But S_1' and S_2' are not the complement of each other; otherwise S_1' would be a recursive set such that $S_1 \subseteq S_1'$ and $S_2 \subseteq S_2'$ and therefore (S_1,S_2) would be recursively separable (contra hypothesis). Hence there is a number *n* such that $n \notin S_1'$ and $n \notin S_2'$. Whence, by (*) and (**) of Lemma 2, $\not\vdash \alpha(\overline{n})$ and $\not\vdash \neg\alpha(\overline{n})$.

Remark. As we know (comp. Ch. 2, Sect. 5.2), a formal system S is called *recursivelly undecidable* if and only if the set *P* of Gödel numbers of its provable formulas is not recursive. And S is called *essentially recursively undecidable* if and only if S and every consistent extension $\mathrm{S}^{\mathrm{ext}}$ of S (i.e. $\mathrm{S}^{\mathrm{ext}} \supseteq \mathrm{S}$) is recursively undecidable. As can be argued, *if* S *is rec. insep., then S is essentially recursively undecidable*. Since otherwise there were a consistent extension $\mathrm{S}^{\mathrm{ext}}$ of S which would be recursively decidable. And then the sets P^{ext} and R^{ext} of Gödel numbers of provable formulas of $\mathrm{S}^{\mathrm{ext}}$ and of refutable formulas of $\mathrm{S}^{\mathrm{ext}}$, respectively, would be both rec. en. such that $\widetilde{P}^{\mathrm{ext}} = R^{\mathrm{ext}}$ and therefore they would be recursive. But in this case since $P \subseteq P^{\mathrm{ext}}$ and $R \subseteq R^{\mathrm{ext}}$ it follows that *P* and *R* would be rec. sep., i.e. S would be recursively separable.

4. Rosser's Theorem (*via* rec. insep.)

4.1. Kleene's rec. insep. pair (S_1,S_2) of rec. en. sets

The existence of such a pair was given by Kleene, using the partial recursive function $\Phi(z,x)$. Let (S_1,S_2) a pair of sets, defined as follows:

(1) $S_1 = \{n \mid \Phi(n,n) = 0\}$

(2) $S_2 = \{n \mid \Phi(n,n)$ is defined and $> 0\}$.

Firstly, let us observe that S_1 and S_2 are disjoint (by their

[171] Cf. R. Smullyan [1993], Ch. II, §6, Theorem 15.

definitions) and both are rec. en. (since a function Φ is partial recursive iff its graph is rec. en.).

Reductio. Suppose (S_1, S_2) is recursively separable. Then there exists a recursive set S, such that

(3) $S_1 \subseteq S$ and

(4) $S_2 \subseteq \widetilde{S}$.

Let $\varphi(x)$ be the characteristic function of S. Using it we write:

(5) If $n \in S$, then $\varphi(n) = 1$.

(6) If $n \in \widetilde{S}$, then $\varphi(n) = 0$.

It follows that

(7) If $\Phi(n,n) = 0$, then $\varphi(n) = 1$ (by (1), (3) and (5)).

(8) If $\Phi(n,n) > 0$, then $\varphi(n) = 0$ (by (2), (4) and (6)).

Now, since S is recursive, $\varphi(x)$ is recursive, and then partial recursive. By the *enumeration theorem for partial recursive functions*, there is an e such that for any x, $\varphi(x) = \Phi(e,x)$. Whence for $x = e$

(9) $\varphi(e) = \Phi(e,e)$.

And then we derive:

(10) If $\Phi(e,e) = 0$, then $\varphi(e) = 1$ (by (7)), and therefore $\Phi(e,e) = 1$ (by (9)).

(11) If $\Phi(e,e) > 0$, then $\varphi(e) = 0$ (by (8)), and therefore, $\Phi(e,e) = 0$ (contradiction). Hence (S_1, S_2) is rec. insep.

4.2. Rosser's Theorem

The existence of a rec. insep. pair (S_1, S_2) of rec. en. sets can also be used for a proof of Rosser's theorem. Kleene's pair, constructed *via* the partial recursive function Φ, fits in best this purpose.

Rosser's Theorem. *If* $\mathrm{S} \supseteq Q$ *is consistent axiomatizable system, then* S *is incomplete.*

Proof.[172] Let (S_1, S_2) be the Kleene's rec. insep. and rec. en. pair of sets, i.e.,

(1) $S_1 = \{m \mid \Phi(m,m) = 0\}$

(2) $S_2 = \{m \mid \Phi(m,m)$ is defined and $> 0\}$.

[172] This version of proof of Rosser's Theorem, using Kleene's rec. insep. pair (S_1,S_2), is due to S. Kripke [1996], 91.

Let $\alpha(\mathrm{x}_1,\mathrm{x}_2)$ be the formula which represents the function $\Phi(x,x)$ in S, i.e.,

(*) If $\Phi(m,m)=n$, then $\vdash \forall \mathrm{x}_2(\alpha(\overline{m},\mathrm{x}_2)\equiv \mathrm{x}_2=\overline{n})$.

If $n=0$, then $\vdash \forall \mathrm{x}_2(\alpha(\overline{m},\mathrm{x}_2)\equiv \mathrm{x}_2=0)$, and then

$\vdash \forall \mathrm{x}_2(\mathrm{x}_2=0\supset \alpha(\overline{m},\mathrm{x}_2)$ (by PL) equivalently,

(3) $\vdash \alpha(\overline{m},0)$ (by Ch. II, Sect. 4.3, Lemma).

If $n>0$, then $\Phi(m,m)=n$, and then, by (*) and PL

$\vdash \forall \mathrm{x}_2(\alpha(\overline{m},\mathrm{x}_2)\supset \mathrm{x}_2=\overline{n})$, equivalently $\vdash \forall \mathrm{x}_2(\mathrm{x}_2\neq \overline{n}\supset \neg\alpha(\overline{m},\mathrm{x}_2))$, and then (since for any x_2, $x_2\neq n$ (where $n>0$) it follows that $n=0$).

$\vdash \forall \mathrm{x}_2(\mathrm{x}_2=0\supset \neg\alpha(\overline{m},\mathrm{x}_2))$, equivalently

(4) $\vdash \neg\alpha(\overline{m},0)$.

Now, since S is consistent

(5) Non($\vdash \alpha(\overline{m},0)$ and $\vdash \neg\alpha(\overline{m},0)$).

Let (6) $M_1=\{m\mid \vdash \alpha(\overline{m},0)\}$ and (7) $M_2=\{m\mid \vdash\neg\alpha(\overline{m},0)\}$. By consistency of S, M_1 and M_2 are disjoint. Moreover, since S is axiomatizable, these sets are rec. en.

By (1), (3) and (6) it follows that $S_1\subseteq M_1$ and by (2) and (4) and (7) it follows that $S_2\subseteq M_2$.

Now, if S were complete, the sets M_1 and M_2 would be the complement of each other (with $M_1\cup M_2=\mathbf{N}$), and then they would be recursive. So, M_1 would be a recursive set that would separate S_1 and S_2, contrary to the rec. inseparability of them.

5. Basic tools of recursion theory

5.1. S_n^m-Theorem[173]

The idea of this fundamental result of recursive function theory is the following: if $R(x,y)$ is an rec. en. relation, how to determine the index of the relation $R(k,y)$, obtained by substituting a number k for x in $R(x,y)$? More exactly, this idea is given in the following terms.

[173] A Kleene's discovery; cf. S.C. Kleene [1952], §65, Theorem XXIII. Sometimes it is known as Iteration Theorem (comp. R. Smullyan [1993], 51).

Theorem 1(a). *For any 2-place relation* $R(x,y)$ *there is a recursive function* φ *such that for any number k*, $W_{\varphi(k)} = R_k = \{y \mid R(k,y)\}$.

Proof. Let $R(x,y)$ be an rec. en. relation, let *e* be its Gödel number. I.e., *e* is the Gödel number of a formula $\alpha(\mathrm{x}_1,\mathrm{x}_2)$ representing $R(x,y)$ in *Q*. Then, as can be observed, the formula $\alpha(\bar{k},x_2)$ represents in *Q* the set $R_k = \{y \mid R(k,y)\} = \{y \mid \vdash \alpha(\bar{k},\bar{y})\}$, i.e., $y \in R_k$ iff $\vdash \alpha(\bar{k},\bar{y})$. Since each time the formula $\alpha(\mathrm{x}_1,\mathrm{x}_2)$ (with Gödel number *e*) is the same, the variable substituted, x_1 (with Gödel number, say 2^{15})[174] is the same, the Gödel number of the formula $\alpha(\bar{k},\mathrm{x}_2)$ only depends on *k*, i.e., it is $\varphi(k)$, and can be determined each time. This number is the Gödel number of the formula obtained from the formula with Gödel number *e*, by substituting the numeral for *k* for the variable with Gödel number 2^{15}. I.e., $\varphi(k) = Sub(e, Num(k), 2^{15})$, where *Sub* is the recursive function defined in 4.1.[175]

Remark. In Smullyan's account,[176] which avoid substitution, the number $\varphi(k)$ is the Gödel number of the formula $\alpha[\bar{k},\mathrm{x}_2]$: $\forall \mathrm{x}_1(\mathrm{x}_1 = \bar{k} \supset \alpha(\mathrm{x}_1,\mathrm{x}_2))$, *provably equivalent* to $\alpha(\bar{k},\mathrm{x}_2)$ (comp. Sect. 4.2.5, Fact 6).

For arbitrary rec. en. relations R^{m+n} the theorem can be derived from the form already given, using *J*-functions.

Theorem 1(b). *For any* $m+n$*-place rec. en. relation* $(m,n>0)$ $R(x_1,\dots,x_m,y_1,\dots,y_n)$ *there is a recursive function* φ *such that for any numbers* $k_1,\dots,k_m$, $W_{\varphi(k_1,\dots,k_m)} = R_{k_1,\dots,k_m} = \{\langle y_1,\dots,y_n\rangle \mid R(k_1,\dots,k_m,y_1,\dots,y_n)\}$.

Proof. By 4.2.5,2, Fact 5[177] and Theorem 1(a).

Theorem 2. *Let* R^{m+n} *an* $m+n$*-place rec. en. relation, let e be its index, and* $k_1,\dots,k_m$ *be any numbers. Then there exists a recursive function* φ^{m+1} *such that* $W_{\varphi(e,k_1,\dots,k_m)} = \{\langle y_1,\dots,y_n\rangle \mid R(k_1,\dots,k_m,y_1,\dots,y_n)\}$.

---

[174] In the arithmetization given in 4.1.

[175] Comp. the item 6 in the list at the end of 4.1.

[176] Comp. R. Smullyan, [1993], 51.

[177] For details, comp. R. Smullyan [1993], 52; S. Kripke [1996], 96-97.

Proof. Since e is the index of R^{m+n}, it follows that

$$R(x_1,...,x_m,y_1,...,y_n) = W(e,x_1,...,x_m,y_1,...,y_n).$$

By the preceding form of S_n^m Theorem, 1(b), applied to $W(e,x_1,...,x_m,y_1,...,y_n)$, there is a recursive function φ^{m+1} such that

$$W_{\varphi(e,k_1,...,k_m)} = \{\langle y_1,...,y_n\rangle \mid W(e,k_1,...,k_m,y_1,...,y_n)\}$$
$$= \{\langle y_1,...,y_n\rangle \mid R(k_1,...,k_m,y_1,...,y_n)\}$$

Theorem 3.[178] *Let* $\varphi(x_1,...,x_m,y_1,...,y_n)$ *be an* $m+n$ *partial recursive function, let e be its index. Then there is a recursive function* ψ^{m+1} *such that* $\psi(e,k_1,...,k_m)$ *is an index of* $\varphi^n(k_1,...,k_m,y_1,...,y_n)$.

Proof. Since e is an index of φ^{m+n}, it follows that

$$\varphi(x_1,...,x_m,y_1,...,y_n) = \Phi^{m+n+1}(e,x_1,...,x_m,y_1,...,y_n).$$

Let $W^{m+n+2}(e,x_1,...,x_m,y_1,...,y_n,z)$ be the graph of Φ^{m+n+1}. Then, by the previous form of the theorem, there is a recursive function ψ^{m+1} such that for any $k_1,...,k_m$,

$$W_{\psi(e,k_1,...,k_m)} = \{\langle y_1,...,y_n,z\rangle \mid W(e,k_1,...,k_m,y_1,...,y_n,z)\}$$
$$= \{\langle y_1,...,y_n,z\rangle \mid \Phi(e,k_1,...,k_m,y_1,...,y_n) = z\}$$
$$= \{\langle y_1,...,y_n,z\rangle \mid \varphi(k_1,...,k_m,y_1,...,y_n) = z\}$$

It follows that $\psi(e,k_1,...,k_m)$ is the index of the *n*-place partial recursive function $\varphi(k_1,...,k_m,y_1,...,y_n)$.

Kleene's form of Gödel's Theorem (*via* S_n^m-Theorem)

As is well-known we may refer to a formal system S either by specifying its set of axioms, or by referring, simply, to the set of its theorems. If S is axiomatizable, then *Th* (the set of Gödel numbers of its theorems) is rec. en., and then it has an index (by Enum. Th.), i.e., $Th = W_e$.

As follows from the Kleene's Theorem,[179] if $S \supseteq Q$ is ω-consistent

---

[178] Due to S.C. Kleene [1952], §65, Theorem XXIII. The function ψ of this theorem is the function S_n^m in Kleene's original notation of Theorem XXIII. Comp. also H. Rogers [1967], 23, Theorem V. The notation adopted here is that of Kripke [1996], 96-7.

[179] Kleene's Theorem XIII (II); comp. and Sect. 4.2.4.4 (Gödel's Theorem II).

and axiomatizable, and $A(x): \forall y \neg\beta(x,x,y)$ expresses the set

$$\widetilde{K} = \{x \mid (y)\widetilde{T}(x,x,y)\},$$ [180]

then a number f can be found such that $A(\bar{f})$ is true, but $A(\bar{f})$ is undecidable, where f is the index of the set

$$\widetilde{K}^* = \{x \mid \vdash \forall y \neg\beta(\bar{x},\bar{x},y)\}.$$

If we refer to S as being W_e, then the number f depends on the number e.

Using S_n^m-Theorem (for $m=n=1$), Kleene's Theorem can be expressed and proved in the following form.

Theorem. *Let* W_e *be any* ω*-consistent and axiomatizable system* $S \supseteq Q$ *with a formula* A(x) *expressing the set* $\widetilde{K}$. *Then there is a recursive function* ψ *such that for all e,* $A(\overline{\psi(e)})$ *is true but undecidable in* W_e.

Proof.[181] Let A(x) be the Π_1-formula expressing the set $\widetilde{K}$, let $R(e,m) = W_e \vdash A(\overline{m})$ (i.e., R is the relation "$A(\overline{m})$ is a theorem of W_e"). R is an rec. en. relation. By S_1^1-Theorem there is (and can be found) a recursive function ψ such that $W_{\psi(e)} = R_e = \{m \mid R(e,m)\} = \{m \mid W_e \vdash A(\overline{m})\}$, for all e. So, $A(\bar{f})$, i.e., $A(\overline{\psi(e)})$ is true and undecidable (if W_e is ω-consistent).

Using S_n^m-*Theorem,* another fundamental result of recursion theory can be derived: the recursion theorem.

5.2. Recursion Theorem.[182] *If* $R(x,y)$ *is any rec. en. relation, then there is a number e such that* $W_e = R_e$.

Proof. Since $R(x,y)$ is rec. en., the relation $R(\Phi(x,x),y)$ is also rec. en. (by 4.2.5, 2 Fact 2). By S_1^1-Theorem there is a recursive function ψ such that for all m, $W_{\psi(m)} = \{y \mid R(\Phi(m,m),y)\}$. Since ψ is recursive, it has an index f. And then $W_{\psi(f)} = \{y \mid R(\Phi(f,f),y)\} = \{y \mid R(\psi(f)),y\} = R_{\psi(f)}$. Let $e = \psi(f)$. Whence $W_e = R_e$.

---

[180] A(x) expresses the predicate $(y)\widetilde{T}(x,x,y)$.

[181] This is a form of Kripke's proof; cf. S. Kripke [1996], 97-98.

[182] Also called the Kleene's second recursion theorem. For this form of the theorem comp. R. Smullyan [1993], Ch. VIII, §1 and S. Kripke [1996], 107-8.

4.2.7. Creative sets and incompleteness

A new fashion in the investigation of incompleteness phenomenon is achieved by using the idea of *creative sets*, a topic whose starting point is due to E. Post.[183]

Post set *C*

As we saw, the diagonal set $K = \{x \mid W(x,x)\}$ is rec. en. but its complement, $\widetilde{K} = \{x \mid \widetilde{W}(x,x)\}$ is not rec. en. And this means that for any rec. en. subset K^* of $\widetilde{K}$ there is a number x such that $K^* \subset \widetilde{K} \to x \in \widetilde{K} - K^*$. Moreover, such a number *x* can be found from the index of any rec. en. subset of $\widetilde{K}$.

Definition. *A set* S *is* ***productive***[184] *if there exists a recursive function* $\varphi(x)$ *such that for any number x:* $W_x \subseteq S \to \varphi(x) \in S - W_x$. $\varphi(x)$ *is called the* ***productive function*** *of S.*

Definition. *A set S is* ***creative***[185] *if* (1) *S is rec. en., and* (2) $\widetilde{S}$ *is productive.*

The set *K* is an example of a creative set, since it is rec. en. and its complement, $\widetilde{K}$, is productive (with identity function $I(x)$ productive function for $\widetilde{K}$).

Definition. *A formal system* S *is* ***creative*** *if the set P of Gödel numbers of its theorems is creative.*

The system $\mathrm{PA}^{\mathrm{ax}}$ is an example of an axiomatizable, undecidable and creative system.

A notable example of a creative set is Post's set C, defined as follows:

$C = \{x \mid x \in W_x\}$, i.e., $x \in C$ iff $x \in W_x$.

As can be observed, if the sets C and W_x are disjoint, then $x \notin C \cup W_x$, equivalent $x \notin C$ and $x \notin W_x$. And then $I(x)$ is a productive function for $\widetilde{C}$ (since $x \in \widetilde{C} - W_x$, and therefore $I(x) \in \widetilde{C} - W_x$).

Theorem. *Let* S *be a correct and axiomatizable formal system. If* α(x) *expresses the set* $\widetilde{C}$ *in* Ls *and represents the set* W_f *in* S, *then* $\alpha(\bar{f})$ *is true*

[183] E. Post [1944].
[184] The name is due to J. Dekker [1955].
[185] The name given by Post [1944], §3; comp. also H. Rogers [1967], §§7.2, 7.3.

and not provable in S.[186]

Proof. By definition of *C*: $x \in C$ iff $x \in W_x$, and if *C* and W_x are disjoint, then $x \notin C$ and $x \notin W_x$. Since S is correct $W_f \subseteq \widetilde{C}$. And then $C \cap W_x = \varnothing$ (i.e., they are disjoint); hence for the number *f* (the index of W_f) we have: $f \notin C$ and $f \notin W_f$, i.e., $f \in \widetilde{C}$ and $f \notin W_f$. Since α(x) expresses $\widetilde{C}$, it follows that $\alpha(\bar{f})$ is true, and since α(x) represents W_f in S it follows that $\alpha(\bar{f})$ is not provable in S.

Remark. In order to obtain the same result of incompleteness (and undecidability) using the set *C* we are looking for a number *n* for which a formula $\alpha(\overline{n})$ is neither provable, nor refutable. And this happens, simply, when the sets *C* and W_n are *disjoint*, i.e., when S is consistent and axiomatizable, α(x) represents the set *C* in S and ¬α(x) represents the set W_n in S. (Detail the proof!).

4.2.8. Complete effective inseparability and incompleteness

As we saw (Sect. 4.2.4.5), by Kleene's symmetric form of Gödel's theorem[187] the undecidability of a sentence in a formal system S can be proved under the assumption of simple consistency of S. This form of Gödel's Theorem also leads directly to another approach of incompleteness and undecidability,[188] that uses the so-called complete effective inseparability (defined below in **2**), *via* Kleene's function.

1. Kleene's symmetric form of Gödel's theorem (review)

Using the rec. en. predicates $(Ey)W_0(x,y)$ and $(Ey)W_1(x,y)$[189] Kleene defines the following rec. en. sets:

$$C_0 = \{x \mid (Ey)W_0(x,y)\} \text{ and } C_1 = \{x \mid (Ey)W_1(x,y)\}.$$

---

[186] Essentially, in the original form α(x) expresses $\widetilde{K} = \{x \mid (y)\widetilde{T}(x,x,y)\}$ and represents the set $\widetilde{K}^* = \{x \mid \vdash \alpha(\bar{x})\}$, and *f* is the index of the rec. en. set $\widetilde{K}^*$.

[187] Actually a generalization of Rosser's Theorem; comp. S.C. Kleene, [1952], §61.

[188] As it is sketched in Kleene [1952], 311-312.

[189] Comp. S.C. Kleene [1952], §61; comp. Sect. 4.2.4.5.

By the construction of these predicates (comp. 4.2.4.5(1)) the sets C_0 and C_1 are disjoint.

Now, using the rec. en. predicates $(Ey)R_0(x,y)$ and $(Ey)R_1(x,y)$ Kleene defines the following rec. en. sets:

$D_2 = \{x \mid (Ey)R_0(x,y)\}$ and $D_3 = \{x \mid (Ey)R_1(x,y)\}$.

Let B(x) be the formula that strongly separates C_0 from C_1 in S and that represents D_2 in S; and ¬B(x) be the formula that represents D_3 in S (comp. 4.2.4.5, (2)-(3) and (4)-(5)). Then, obviously $C_0 \subseteq D_2$, $C_1 \subseteq D_3$, and D_2 and D_3 are disjoint.

By his theorem, a number *f* can be found such that $f \notin D_2$ and $f \notin D_3$, from which it follows that $\nvdash \mathrm{B}(\bar{f})$ and $\nvdash \neg\mathrm{B}(\bar{f})$. This number $f = 2^{f_0} \cdot 3^{f_1}$, where f_0 is the index of D_2 and f_1 is the index of D_3 in an enumeration. So, *f* is the value of a recursive function $\psi(x,y)$ (i.e., $2^x \cdot 3^y$), for *x* and *y* specified.

By his *reductio*, if $\mathrm{B}(\bar{f})$ were provable, then $f \in D_2$ (since B(x) represents D_2 in S). And since S is consistent, then $f \notin D_3$. Now, by the construction of the predicate $(Ey)W_1(x,y)$, it follows that $(Ey)W_1(f,y)$ holds and then $f \in C_1$. But $C_1 \subseteq D_3$ and therefore $f \in D_3$, whence $\neg\mathrm{B}(\bar{f})$ would be provable (since ¬B(x) represents the set D_3 in S). But this means that S would be inconsistent. So, if S is consistent, $\nvdash \mathrm{B}(\bar{f})$.

By a similar argument we can show that if S is consistent, then $\nvdash \neg\mathrm{B}(\bar{f})$. Therefore, if S is consistent, $\mathrm{B}(\bar{f})$ is undecidable in S.

Let us resume! From the preceding argument it follows that for the disjoint pair (C_0, C_1) the recursive function $\psi(x,y)$ is such that

(1) If $\psi(f_0, f_1) \in W_{f_1} - W_{f_0}$, then $\psi(f_0, f_1) \in C_0$.

(2) If $\psi(f_0, f_1) \in W_{f_0} - W_{f_1}$, then $\psi(f_0, f_1) \in C_1$,

where W_{f_0} and W_{f_1} are the sets D_2 and D_3.

Such a function, $\psi(x,y)$, is called a *Kleene function*[190] for the disjoint pair (C_0, C_1).

[190] The name given by R. Smullyan, [1993], 68.

Using $J(x,y)$ as a Kleene function for an rec. en. pair (K_1,K_2) of sets, *via* the notion of completely effectively inseparable sets, Smullyan gives an interesting proof of Kleene's symmetric form of Gödel's Theorem.[191] Let us sketch the ideas.

2. Effective and complete effective inseparability

As we saw (comp. 4.2.6, 2), a disjoint pair of sets (S_1,S_2) is rec. insep. if there is no recursive set S such that $S_1 \subseteq S$ and $S_2 \subseteq \widetilde{S}$. And this means that for arbitrary disjoint rec. en. sets W_i, W_j such that $S_1 \subseteq W_i$ and $S_2 \subseteq W_j$: $W_i \neq \widetilde{W}_j$; and therefore there is an *n* such that $n \notin W_i \cup W_j$.

Definition 1. *A disjoint pair of sets* (S_1,S_2) *is effectively inseparable* (*abrev. eff. insep.*) *if there is a recursive function* $\varphi(x,y)$ *such that for any numbers i, j, with* $S_1 \subseteq W_i$, $S_2 \subseteq W_j$ *and* W_i *disjoint from* W_j, *the number* $\varphi(i,j) \notin W_i \cup W_j$. *The function* $\varphi(x,y)$ *is called an eff. insep. function for the pair* (S_1,S_2).

Definition 2. *A disjoint pair of sets* (S_1,S_2) *is completely effectively inseparable*[192] (*abrev. compl. eff. insep.*) *if there is a recursive function* $\varphi(x,y)$ *such that for any numbers i, j, with* $S_1 \subseteq W_i$ *and* $S_2 \subseteq W_j$ *the following holds*:

$$\varphi(i,j) \in W_i \leftrightarrow \varphi(i,j) \in W_j .$$

The function $\varphi(x,y)$ is called a compl. eff. insep. function for the pair (S_1,S_2).

As can be remarked, if $W_i \cap W_j = \varnothing$, the number $\varphi(i,j) \notin W_i \cup W_j$. Hence if $\varphi(x,y)$ is a compl. eff. insep. function for a pair (S_1,S_2), then it is also an eff. insep. function for (S_1,S_2).

Theorem.[193] *Let the following be given:*

(a) S *is a consistent and axiomatizable system.*

(b) α(x) *strongly separates the pair* (K_1,K_2) *in* S.

---

[191] Comp. R. Smullyan, [1993], Cap. V, I, §2.

[192] The terms "effectively inseparable" and "completely effectively inseparable" are those of R. Smullyan [1993], Cap. V, I.

[193] This is Smullyan's version of Kleene's symmetric form of Gödel's Theorem, proved *via* compl. eff. insep. of an rec. en. pair (K_1,K_2); cf. R. Smullyan [1993], Ch. V, I, Theorem 3.

(c) W_k *is the set represented by* α(x).

(d) W_m *is the set represented by* $\neg\alpha$(x).

Then $\nvdash \alpha(\overline{J(k,m)})$ *and* $\nvdash \neg\alpha(\overline{J(k,m)})$ (*i.e., the sentence* $\alpha(\overline{J(k,m)})$ *is undecidable in* S, *and therefore* S *is incomplete*).

Proof. By (b) for any n:

(1) $n \in K_1 \to \vdash \alpha(\overline{n})$

(2) $n \in K_2 \to \vdash \neg\alpha(\overline{n})$.

By (c) and (d), we have, accordingly

(3) $n \in W_k \leftrightarrow \vdash \alpha(\overline{n})$

(4) $n \in W_m \leftrightarrow \vdash \neg\alpha(\overline{n})$.

Whence, by (1) and (3): $K_1 \subseteq W_k$, and by (2) and (4): $K_2 \subseteq W_m$. Obviously, $W_k \cap W_m = \varnothing$ (since S is consistent).

Now, since $J(x,y)$ is a Kleene function for (K_1,K_2), $J(x,y)$ is a compl. eff. insep. function for (K_1,K_2) and since (W_k,W_m) are disjoint, $J(x,y)$ is an eff. insep. function for (K_1,K_2) (comp. Def 1). So, the number $J(k,m) \notin W_k$ and $J(k,m) \notin W_m$. Whence, by (3) and (4): $\nvdash \alpha(\overline{J(k,m)})$ and $\nvdash \neg\alpha(\overline{J(k,m)})$. Therefore $\alpha(\overline{J(k,m)})$ is undecidable in S.

Remark. As we saw (comp. Sect. 4.2.6,3, Theorem) the undecidability of a sentence $\alpha(\overline{n})$ in S follows from the following assumptions:[194]

(a) S is consistent and axiomatizable.

(b) α(x) strongly separates (S_1,S_2) in S.

(c) (S_1,S_2) are rec. insep.

From (b) are deducible:

(1) α(x) represents a set S_1' such that $S_1 \subseteq S_1'$ in S.

(2) $\neg\alpha$(x) represents a set S_2' such that $S_2 \subseteq S_2'$ in S.

By (a) S_1' and S_2' are disjoint and rec. en. And by (c) they are not the complement of each other. Hence there is an n such that $n \notin S_1'$ and $n \notin S_2'$. Whence, by (1) and (2), $\nvdash \alpha(\overline{n})$ and $\nvdash \neg\alpha(\overline{n})$.

[194] By a *non-constructive* proof, since it gives no clue how to find this n.

In the *constructive* version of this theorem[195] (the theorem above) the number n for which $\alpha(\overline{n})$ is undecidable in S is delivered (constructively) by a Kleene's function $J(x,y)$, this being an effective insep. function for (K_1,K_2), in which case the number $n = J(k,m)$ is outside both sets W_k and W_m. And therefore $\not\vdash \alpha(\overline{n})$ and $\not\vdash \neg\alpha(\overline{n})$.

[195] The correspondence between the non-constructive form and the constructive form is easy to grasp: S_1, S_2, S_1', S'_2 become K_1, K_2, W_k, W_m, respectively; and α(x) has the corresponding role.

Chapter 4. MODAL LOGIC OF PROVABILITY

The modal logic concerns basically the study of notions "necessity" and "possibility". The modal logic of provability is that special part of modal logic, as it is applied to the study of mathematical reasoning, in which case the idea of *necessity* becomes the mathematical idea of *provability*. What is properly called the modal logic of provability is the propositional modal system $\mathcal{GL}$ (Gödel-Löb).

4.1. Systems of propositional modal logic

4.1.1. Syntax

The propositional *modal* logic PML is an extension of the propositional *classical* logic. The language of modal logic (L_{PML}) consists of the following classes of primitive (undefined) symbols: p, q, r,... (propositional variables); $\neg$, $\supset$ (propositional connectives); $\Box$ (modal operator) and auxiliary symbols: (,) (parentheses).

The notion of formula of L_{PML} is inductively defined strictly by the following rules:

(1) p, q, r,... are (atomic) formulas of L_{PML}.

(2) If α is a formula of L_{PML}, then $\neg\alpha$ and $\Box\alpha$ are formulas of L_{PML}.

(3) If α and β are formulas of L_{PML}, then $\alpha\supset\beta$ is a formula of L_{PML} (where α, β,... denote arbitrary formulas of L_{PML}).

As usually, we also use some other connectives of propositional classical logic, e.g., $\wedge$, $\vee$, $\equiv$. They will be introduced in the well-known way, using the primitive symbols $\neg$ and $\supset$.

Remark. Sometimes[1], besides the "$\neg$" we use, as primitive symbol, "$\perp$". It is 0-ary connective and denotes the *logical falsity*. Then, its relative, "$\top$", will denote the *logical truth*. Using "$\perp$", evidently, the symbol of negation is dispensable, since $\neg$p and $\neg\Box$p, for example, can be expressed by $p\supset\perp$ and $\Box p\supset\perp$, respectively. Even the symbol "$\top$" is eliminable, since it can be defined by $\perp\supset\perp$. The reason for this minor change is the following. The idea of provability in PA^{ax} is shaped by $\mathcal{GL}$, and since many notable

[1] For the language of $\mathcal{GL}$.

formulas of L_{PA} regarding the provability are *constant sentences*[2], the corresponding modal formulas not containing symbols for propositional variables (also called letterless sentences, e.g., $\bot$, $\Box\bot\supset\bot$, $\neg\Box\bot$) will play a key role.

The system $\mathcal{K}$ (Kripke)

To set out a modal system in the language L_{PML} means to specify its axioms, rules of deduction and (eventually) the definitions.

Axioms of $\mathcal{K}$

(1) Any valid formula α of propositional classical[3] logic (PL).

(2) The modal axiom **K**: $\Box(p\supset q)\supset(\Box p\supset\Box q)$.

Rules of deduction

Modus ponens (MP) $\dfrac{\alpha,\ \alpha\supset\beta}{\beta}$

Substitution (Subst) $\dfrac{\vdash\alpha(p_1,\ldots,p_n)}{\vdash\alpha(\beta_1/p_1,\ldots,\beta_n/p_n)}$

Necessitation (N) $\dfrac{\vdash\alpha}{\vdash\Box\alpha}$

Def$\Diamond$. $\Diamond\alpha=_{df}\neg\Box\neg\alpha$,

where "$\Diamond$" is the modal operator "possibility", dual to the operator "$\Box$".

In Subst $\alpha(p_1,\ldots,p_n)$ is a formula of L_{PML} containing the propositional variables $p_1,\ldots,p_n$, and $\alpha(\beta_1/p_1,\ldots,\beta_n/p_n)$ is the formula of L_{PML} obtained from $\alpha(p_1,\ldots,p_n)$ by substituting the formulas of L_{PML} $\beta_1,\ldots,\beta_n$ for $p_1,\ldots,p_n$, respectively (where $\beta_1,\ldots,\beta_n$ are arbitrary).

As evident, the first two rules (MP and Subst) are not specifically modal.[4]

All the other propositional modal systems are *proper* extensions of the system $\mathcal{K}$, as follows:

---

[2] As we'll see below, such a formula α^* is the realization (interpretation) of some modal *letterless* sentence.

[3] As can be seen (1) is an axiom *schema*, i.e., a prescription (recipe) indicating what can we take as being an axiom of $\mathcal{K}$, it is not a formula of the object language. In what follows this use of (1) is often indicated by PL.

[4] Comp. their use in PL^{ax} in Ch. 1 (3.1, 3.2).

$\mathcal{K}4 = \mathcal{K} + \mathbf{4}$: $\Box p \supset \Box\Box p$
$\mathcal{T} = \mathcal{K} + \mathbf{T}$: $\Box p \supset p$
$\mathcal{S}4 = \mathcal{K} + \mathbf{T} + \mathbf{4}$
$\mathcal{S}5 = \mathcal{K} + \mathbf{T} + \mathbf{5}$: $\Diamond p \supset \Box\Diamond p$
$\mathcal{B} = \mathcal{K} + \mathbf{T} + \mathbf{B}$: $p \supset \Box\Diamond p$
$\mathcal{GL} = \mathcal{K} + \mathbf{W}$: $\Box(\Box p \supset p) \supset \Box p$

If $\mathcal{S}$ is a system, then $\mathcal{S}'$ is an extension of $\mathcal{S}$ if $\mathcal{S} \subseteq \mathcal{S}'$ (i.e., all theorems of $\mathcal{S}$ are also theorems of $\mathcal{S}'$); and if $\mathcal{S}'$ has theorems which are not theorems of $\mathcal{S}$, then $\mathcal{S}'$ is a *proper* extension of $\mathcal{S}$.

A formal system is called *normal* if it is an extension (proper or not) of the system $\mathcal{K}$, i.e., if the set of its theorems contains all valid formulas of propositional classical logic, the (distribution) axiom **K** and is closed under MP, Subst and N.

Let $\mathcal{S}$ be a normal system. Then the following results hold:

Theorems of $\mathcal{K}$

Deriv. 1[5] $\dfrac{\vdash \alpha \supset \beta}{\vdash \Box\alpha \supset \Box\beta}$

(1) $\vdash \alpha \supset \beta$; hyp.
(2) $\vdash \Box(\alpha \supset \beta)$; (1) N
(3) $\vdash \Box(\alpha \supset \beta) \supset (\Box\alpha \supset \Box\beta)$; **K**, Subst α/p, β/q
(4) $\vdash \Box\alpha \supset \Box\beta$; (2), (3), MP

Deriv. 2 $\dfrac{\vdash \alpha \equiv \beta}{\vdash \Box\alpha \equiv \Box\beta}$

(1) $\vdash \alpha \equiv \beta$; hyp.
(2) $\vdash (\alpha \equiv \beta) \supset (\alpha \supset \beta)$; PL
(3) $\vdash \alpha \supset \beta$; (1), (2), MP

[5] Deriv. 1 – Deriv. 4 are called *derived rules of deduction*; they are provable in $\mathcal{S}$, are not necessary for $\mathcal{S}$ but whose application makes the proofs easier.

(4) $\vdash \Box\alpha \supset \Box\beta$: (3) Deriv. 1

(5) $\vdash (\alpha \equiv \beta) \supset (\beta \supset \alpha)$; PL

(6) $\vdash \beta \supset \alpha$; (1), (5) MP

(7) $\vdash \Box\beta \supset \Box\alpha$; (6) Deriv. 1

(8) $\vdash \Box\alpha \equiv \Box\beta$; (4), (7) PL: $(p \supset q) \supset ((q \supset p) \supset (p \equiv q))$ + Subst, MP

Deriv. 3 $\dfrac{\vdash \alpha \supset \beta}{\vdash \Diamond\alpha \supset \Diamond\beta}$

(1) $\vdash \alpha \supset \beta$; hyp.

(2) $\vdash (\alpha \supset \beta) \supset (\neg\beta \supset \neg\alpha)$; PL

(3) $\vdash \neg\beta \supset \neg\alpha$; (1), (2) MP

(4) $\vdash \Box\neg\beta \supset \Box\neg\alpha$; (3) Deriv. 1

(5) $\vdash \neg\Box\neg\alpha \supset \neg\Box\neg\beta$; (4) PL (contrapoz. $\supset$)

(6) $\vdash \Diamond\alpha \supset \Diamond\beta$; (5) Def. $\Diamond$

Deriv. 4 $\dfrac{\vdash \alpha \equiv \beta}{\vdash \Diamond\alpha \equiv \Diamond\beta}$

(1) $\vdash \alpha \equiv \beta$; hyp.

(2) $\vdash \alpha \supset \beta$; (1) PL

(3) $\vdash \Diamond\alpha \supset \Diamond\beta$; (2) Deriv. 3

(4) $\vdash \beta \supset \alpha$; (1) PL

(5) $\vdash \Diamond\beta \supset \Diamond\alpha$; (4) Deriv. 3

(6) $\vdash \Diamond\alpha \equiv \Diamond\beta$; (3), (5) PL

Replacement Theorem (Repl).[6] *Let* $\alpha(\beta)$ *be a formula of* $\mathrm{L}_{\mathrm{PML}}$ *containing as subformula* (*proper or not*) *the formula* β. *Let* $\alpha(\gamma)$ *be the formula obtained from* $\alpha(\beta)$ *by replacing an arbitrary number of occurrences of* β *with* γ. *Then*

If $\vdash \beta \equiv \gamma$, *then* $\vdash \alpha(\beta) \equiv \alpha(\gamma)$.

[6] This theorem is an extension of the corresponding theorem from PL (with a slight notational modification); comp. Ch. 1, 3.2.1.2.

Proof (induction on the complexity of α).

(1) $\alpha(\beta)$=p. In this case $\alpha=\beta=p$. If γ is a formula such that $\vdash p\equiv\gamma$ (e.g., $\gamma=\neg\neg p$, $\gamma=p\wedge(q\vee\neg q)$), then $\vdash\alpha(\beta)\equiv\alpha(\gamma)$ evidently holds.

(2) $\alpha(\beta) = \neg\delta$.
We have: if $\vdash\beta\equiv\gamma$, then $\vdash\delta(\beta)\equiv\delta(\gamma)$ (ind. hyp.). It follows that $\vdash\neg\delta(\beta)\equiv\neg\delta(\gamma)$ (by PL), i.e., $\vdash\alpha(\beta)\equiv\alpha(\gamma)$.

(3) $\alpha(\beta)=\delta\supset\varepsilon$, and the theorem holds for δ and ε, i.e.,
(a) If $\vdash\beta\equiv\gamma$, then $\delta(\beta)\equiv\delta(\gamma)$, and
(b) If $\vdash\beta\equiv\gamma$, then $\vdash\varepsilon(\beta)=\varepsilon(\gamma)$; and therefore, by PL
(c) If $\vdash\beta\equiv\gamma$, then $\vdash(\delta(\beta)\equiv\delta(\gamma))\wedge(\varepsilon(\beta)\equiv\varepsilon(\gamma))$.
(d) If $\vdash\beta\equiv\gamma$, then $\vdash(\delta(\beta)\supset\varepsilon(\beta))\equiv(\delta(\gamma)\supset\varepsilon(\gamma))$; from (c) by PL:
$\vdash(p\equiv q)\wedge(r\equiv s)\supset(p\supset r)\equiv(q\supset s)$.
I.e., If $\vdash\beta\equiv\gamma$, then $\vdash\alpha(\beta)\equiv\alpha(\gamma)$.

(4) $\alpha(\beta)=\Box\delta$, and the theorem holds for δ, i.e., If $\vdash\beta\equiv\gamma$, then $\vdash\delta(\beta)\equiv\delta(\gamma)$. But then $\vdash\Box\delta(\beta)\equiv\Box\delta(\gamma)$ (by Deriv. 2), i.e., $\vdash\alpha(\beta)\equiv\alpha(\gamma)$.

K$_1$ $\vdash\Box(\alpha\wedge\beta)\equiv(\Box\alpha\wedge\Box\beta)$

(1) $\vdash(\alpha\wedge\beta)\supset\alpha$; PL
(2) $\vdash\Box(\alpha\wedge\beta)\supset\Box\alpha$; (1) Deriv. 1
(3) $\vdash(\alpha\wedge\beta)\supset\beta$; PL
(4) $\vdash\Box(\alpha\wedge\beta)\supset\Box\beta$; (3) Deriv. 1
(5) $\vdash\Box(\alpha\wedge\beta)\supset(\Box\alpha\wedge\Box\beta)$; (2), (4) PL
by $\vdash(p\supset q)\supset((p\supset r)\supset(p\supset(q\wedge r))$
(6) $\vdash\alpha\supset(\beta\supset(\alpha\wedge\beta))$; PL
(7) $\vdash\Box\alpha\supset\Box(\beta\supset(\alpha\wedge\beta))$; (6) Deriv. 1
(8) $\vdash\Box(\beta\supset(\alpha\wedge\beta))\supset(\Box\beta\supset\Box(\alpha\wedge\beta))$; by **K**
(9) $\vdash\Box\alpha\supset(\Box\beta\supset\Box(\alpha\wedge\beta))$; (7), (8) PL
(10) $\vdash(\Box\alpha\wedge\Box\beta)\supset\Box(\alpha\wedge\beta)$; (9) PL
(11) $\vdash\Box(\alpha\wedge\beta)\equiv(\Box\alpha\wedge\Box\beta)$; (5), (10) PL

The theorem **K**$_1$ can be generalized in the following form:
$\vdash\Box(\alpha_1\wedge...\wedge\alpha_n)\equiv(\Box\alpha_1\wedge...\wedge\Box\alpha_n)$ (argue!).

$\mathbf{K}_2$ $\vdash (\Box\alpha\vee\Box\beta)\supset\Box(\alpha\vee\beta)$

(1) $\vdash \alpha\supset(\alpha\vee\beta)$; PL

(2) $\vdash \beta\supset(\alpha\vee\beta)$; PL

(3) $\vdash \Box\alpha\supset\Box(\alpha\vee\beta)$; (1) Deriv. 1

(4) $\vdash \Box\beta\supset\Box(\alpha\vee\beta)$; (2) Deriv. 2

(5) $\vdash (\Box\alpha\vee\Box\beta)\supset\Box(\alpha\vee\beta)$; (3), (4) PL

by $\vdash (p\supset r)\supset((q\supset r)\supset((p\vee q)\supset r))$

$\mathbf{K}_3$ $\vdash \Box\alpha\equiv\neg\Diamond\neg\alpha$

(1) $\alpha\equiv\neg\neg\alpha$; PL

(2) $\Box\alpha\equiv\Box\neg\neg\alpha$; (1) Deriv. 1

(3) $\Box\neg\neg\alpha\equiv\neg\neg\Box\neg\neg\alpha$; PL,Subst

(4) $\Box\alpha\equiv\neg\neg\Box\neg\neg\alpha$; (2), (3) PL

(5) $\Box\alpha\equiv\neg\Diamond\neg\alpha$; (4) Def. $\Diamond$

This theorem together with Def $\Diamond$ indicate a simple procedure of *interchanging* the modal operators $\Box$ and $\Diamond$ in an adjacent sequence of such operators: a negation sign is put before the sequence and such a sign is put after it, and then by replacing $\Box$ with $\Diamond$ and $\Diamond$ with $\Box$ throughout the sequence (and, finally, if it is the case, by deleting all the double negation signs).

Examples. $\Box\Diamond=\neg\Diamond\Box\neg$; $\Box\Box\Diamond=\neg\Diamond\Diamond\Box\neg$,

$\neg\Diamond\Box\Diamond=\neg\neg\Box\Diamond\Box\neg=\Box\Diamond\Box\neg$.

In what follows we refer to this procedure by **Interch**.

$\mathbf{K}_4$ $\vdash \Diamond(\alpha\vee\beta)\equiv(\Diamond\alpha\vee\Diamond\beta)$

(1) $\vdash \Box(\neg\alpha\wedge\neg\beta)\equiv(\Box\neg\alpha\wedge\Box\neg\beta)$; $\mathbf{K}_1$

(2) $\vdash \neg\Diamond\neg(\neg\alpha\wedge\neg\beta)\equiv(\Box\neg\alpha\wedge\Box\neg\beta)$; (1) Interch.

(3) $\vdash \neg\Diamond(\alpha\vee\beta)\equiv(\neg\Diamond\alpha\wedge\neg\Diamond\beta)$; (2) PL, Interch.

(4) $\vdash \Diamond(\alpha\vee\beta)\equiv(\Diamond\alpha\vee\Diamond\beta)$; (3) PL

$\mathbf{K}_5$ $\vdash (\Diamond\alpha\wedge\Box\beta)\supset\Diamond(\alpha\wedge\beta)$

Let us give a more convenient form of $\mathbf{K}_5$, equivalent to $\mathbf{K}_5$.

(1) $\neg\Diamond(\alpha\wedge\beta)\supset\neg(\Diamond\alpha\wedge\Box\beta)$; $\mathbf{K}_5$, PL

(2) $\Box\neg(\alpha\wedge\beta)\supset(\Box\beta\supset\Box\neg\alpha)$; (1) Interch., PL

Now, to prove $\mathbf{K}_5$ means to prove (2).

(a) $\vdash \Box\neg(\alpha\wedge\beta)\equiv\Box(\neg\alpha\vee\neg\beta)\equiv\Box(\beta\supset\neg\alpha)$; PL

(b) $\vdash \Box(\beta\supset\neg\alpha)\supset(\Box\beta\supset\Box\neg\alpha)$; **K**, Subst.

(c) $\vdash \Box\neg(\alpha\wedge\beta)\supset(\Box\beta\supset\Box\neg\alpha)$; (a), (b) PL

The following formulas are also the theorems in $\mathcal{K}$ (show that!):

K$_6$ $\vdash \Diamond(\alpha\supset\beta)\equiv(\Box\alpha\supset\Diamond\beta)$

K$_7$ $\vdash \Box(\alpha\supset\beta)\supset(\Diamond\alpha\supset\Diamond\beta)$

K$_8$ $\vdash \Diamond(\alpha\wedge\beta)\supset(\Diamond\alpha\wedge\Diamond\beta)$

K$_9$ $\vdash (\Box\alpha\wedge\Diamond\beta)\supset\Diamond(\alpha\wedge\beta)$

K$_{10}$ $\vdash \Box(\alpha\vee\beta)\supset(\Box\alpha\vee\Diamond\beta)$.

Since all the other normal systems are extensions of $\mathcal{K}$, all the above theorems of $\mathcal{K}$ will be theorems in these systems.

First Substitution Theorem. *Let* $\alpha(p)$ *be a formula of* L_{PML} *containing the variable* p, $\alpha(\beta)$ *and* $\alpha(\gamma)$ *be the formulas obtained from* $\alpha(p)$ *by substituting* β *and* γ, *respectively for* p. *Then the following holds:*

If $\vdash \beta\equiv\gamma$, *then* $\vdash \alpha(\beta)\equiv\alpha(\gamma)$.

Proof (induction on the complexity of α).

α=p. Then if $\vdash \beta\equiv\gamma$, then $\vdash \beta\equiv\gamma$.

$\alpha=\neg\delta(p)$. Since the theorem holds for $\delta(p)$ (by ind. hyp.) we have: if $\vdash \beta\equiv\gamma$, then $\vdash \delta(\beta)\equiv\delta(\gamma)$ and therefore $\vdash \neg\delta(\beta)\equiv\neg\delta(\gamma)$ (by PL), i.e., $\vdash \alpha(\beta)\equiv\alpha(\gamma)$.

$\alpha(p)=\delta(p)\supset\varepsilon(p)$. Since the theorem holds for $\delta(p)$ and $\varepsilon(p)$ we have:

(a) If $\vdash \beta\equiv\gamma$, then $\vdash \delta(\beta)\equiv\delta(\gamma)$.

(b) If $\vdash \beta\equiv\gamma$, then $\vdash \varepsilon(\beta)\equiv\varepsilon(\gamma)$.

And therefore: If $\vdash \beta\equiv\gamma$, then $\vdash (\delta(\beta)\equiv\delta(\gamma)\wedge(\varepsilon(\beta)\equiv\varepsilon(\gamma))$ by PL

But then $\vdash (\delta(\beta)\supset\varepsilon(\beta))\equiv(\delta(\gamma)\supset\varepsilon(\gamma))$ (by PL), i.e.,

$\vdash \alpha(\beta)\supset\alpha(\gamma)$; by $\vdash [(p\equiv q)\wedge(r\equiv s)]\supset[(p\supset r)\equiv(q\supset s)]$ + Subst + MP

As in the case of Replacement Theorem in PL (comp. Ch. 1, Sect. 2.4.2), if β occurs only in δ or only in γ, the proof can be constructed accordingly (exercise).

$\alpha(p)=\Box\delta(p)$. Since the theorem holds for $\delta(p)$ we have: If $\vdash \beta\equiv\gamma$, then $\vdash \delta(\beta)\equiv\delta(\gamma)$. But then $\vdash \Box\delta(\beta)\equiv\Box\delta(\gamma)$ (by Deriv. 2), i.e., $\vdash \alpha(\beta)\equiv\alpha(\gamma)$.

Definition. *The modal operator* ***strong box*** $\boxdot$ *is defined as follows:*

$$\boxdot\alpha=_{df}\Box\alpha\wedge\alpha.$$

Theorem 1. (a) $\mathcal{K}4 \vdash \boxdot\Box\alpha\equiv\Box\alpha$

(b) $\mathcal{K}4 \vdash \boxdot\Box\alpha\equiv\Box\boxdot\alpha$

(c) $\mathcal{K}4 \vdash \boxdot\alpha \equiv \boxdot\boxdot\alpha$

Proof. (a) $\mathcal{K}4 \vdash \boxdot\Box\alpha \equiv \Box\alpha$

(1) $\mathcal{K}4 \vdash \Box\alpha \supset \Box\Box\alpha$; by **4**

And since $\boxdot\Box\alpha$ is $\Box\Box\alpha \wedge \Box\alpha$, we have

(2) $\mathcal{K}4 \vdash (\Box\alpha \supset \Box\Box\alpha) \supset ((\Box\Box\alpha \wedge \Box\alpha) \equiv \Box\alpha)$,

by PL: $\vdash (p \supset q) \supset ((q \wedge p) \equiv p)$.

(3) $\mathcal{K}4 \vdash (\Box\Box\alpha \wedge \Box\alpha) \equiv \Box\alpha$; (1), (2), MP

(b) $\mathcal{K}4 \vdash \boxdot\Box\alpha \equiv \Box\boxdot\alpha$

Since $\mathcal{K}4 \vdash \boxdot\Box\alpha \equiv \Box\Box\alpha \wedge \Box\alpha$ (by Def) $\equiv \Box(\Box\alpha \wedge \alpha)$ (by $\mathbf{K}_1$) $\equiv \Box\boxdot\alpha$ (by Def.)

(c) $\mathcal{K}4 \vdash \boxdot\alpha \equiv \boxdot\boxdot\alpha$

By Def. $\boxdot\boxdot\alpha$ is $\Box\boxdot\alpha \wedge \boxdot\alpha$, equivalent $\Box\alpha \wedge \boxdot\alpha$ (by (a) and (b), equivalent $\Box\alpha \wedge \Box\alpha \wedge \alpha$ (by Def), that is $\Box\alpha \wedge \alpha$, i.e. $\boxdot\alpha$.

Theorem 2. *If* $\mathcal{K}4 \vdash \boxdot\alpha \supset \beta$, *then*

(a) $\mathcal{K}4 \vdash \Box\alpha \supset \Box\beta$ *and*

(b) $\mathcal{K}4 \vdash \boxdot\alpha \supset \boxdot\beta$.

Proof. (a)

(1) $\mathcal{K}4 \vdash \boxdot\alpha \supset \beta$; hyp.

(2) $\mathcal{K}4 \vdash \Box\boxdot\alpha \supset \Box\beta$; (1) Deriv. 1

(3) $\mathcal{K}4 \vdash \Box\alpha \supset \Box\beta$; (2) Theorem 1, (a) and (b).

(b) (exercise).

Theorem 3. *If* K4 $\vdash \Box\alpha \supset \beta$, *then* K4 $\vdash \Box\alpha \supset \Box\beta$.

Proof. (1) $\mathcal{K}4 \vdash \Box\alpha \supset \beta$; hyp.

(2) $\mathcal{K}4 \vdash \Box\Box\alpha \supset \Box\beta$; (1) Deriv. 1

(3) $\mathcal{K}4 \vdash \Box\alpha \supset \Box\beta$; (2), **4**, PL

Theorem 4. *For* $\mathcal{K}4$ *the following holds:*

(a) $\mathcal{K}4 \vdash (\boxdot\alpha \wedge \boxdot(\alpha \supset \beta)) \supset \boxdot\beta$

(b) $\mathcal{K}4 \vdash \boxdot\alpha \supset \boxdot\boxdot\alpha$

(c) *If* $\mathcal{K}4 \vdash \alpha$, *then* $\mathcal{K}4 \vdash \boxdot\alpha$

(d) $\mathcal{K}4 \vdash \boxdot(\alpha \wedge \beta) \equiv (\boxdot\alpha \wedge \boxdot\beta)$

(e) $\mathcal{K}4 \vdash \boxdot(\alpha \supset \beta) \supset (\boxdot\alpha \supset \boxdot\beta)$

(f) $\mathcal{K}4 \vdash \boxdot(\alpha \equiv \beta) \supset (\boxdot\alpha \equiv \boxdot\beta)$.

Proof. (a)

(1) $\vdash(\alpha\wedge(\alpha\supset\beta))\supset\beta$

(2) $\vdash(\Box\alpha\wedge\Box(\alpha\supset\beta))\supset\Box\beta$; (1) Deriv. 1, $\mathbf{K}_1$

(3) $\vdash(\Box\alpha\wedge\Box(\alpha\supset\beta)\wedge\alpha\wedge(\alpha\supset\beta))\supset(\Box\beta\wedge\beta)$; (1), (2) PL

(4) $\vdash((\Box\alpha\wedge\alpha)\wedge(\Box(\alpha\supset\beta)\wedge(\alpha\supset\beta)))\supset(\Box\beta\wedge\beta)$; (3), PL

(5) $\vdash(\boxdot\alpha\wedge\boxdot(\alpha\supset\beta))\supset\boxdot\beta$; (4) Def.

Proof (b)–(f) (exercise).

As can be seen, these properties of $\boxdot$ are similar to those of $\Box$. However, there are properties specific to $\boxdot$, e.g., $\mathcal{K}4\vdash\boxdot\alpha\supset\alpha$ and those from Th. 1.

Second Substitution Theorem. $\mathcal{K}4\vdash\boxdot(\beta\equiv\gamma)\supset(\alpha(\beta)\equiv\alpha(\gamma))$.

Proof (induction on the complexity of α).

α=p. Then $\mathcal{K}4\vdash\boxdot(\beta\equiv\gamma)\supset(\beta\equiv\gamma)$; holds by PL

$\alpha=\neg\delta$. We have the following derivations:

(1) $\mathcal{K}4\vdash\boxdot(\beta\equiv\gamma)\supset(\delta(\beta)\equiv\delta(\gamma))$; by ind. hyp.

(2) $\mathcal{K}4\vdash\delta(\beta)\equiv\delta(\gamma)\supset(\neg\delta(\beta)\equiv\neg\delta(\gamma))$; by PL

(3) $\mathcal{K}4\vdash\boxdot(\beta\equiv\gamma)\supset(\neg\delta(\beta)\equiv\neg\delta(\gamma))$; (1), (2) PL

i.e., $\mathcal{K}4\vdash\boxdot(\beta\equiv\gamma)\supset(\alpha(\beta)\equiv\alpha(\gamma))$.

$\alpha=\delta\supset\varepsilon$; and the theorem holds for δ and ε, i.e.,

(1) $\mathcal{K}4\vdash\boxdot(\beta\equiv\gamma)\supset(\delta(\beta)\equiv\delta(\gamma))$

(2) $\mathcal{K}4\vdash\boxdot(\beta\equiv\gamma)\supset(\varepsilon(\beta)\equiv\varepsilon(\gamma))$, and then

(3) $\mathcal{K}4\vdash\boxdot(\beta\equiv\gamma)\supset((\delta(\beta)\equiv\delta(\gamma))\wedge(\varepsilon(\beta)\equiv\varepsilon(\gamma)))$ and then

(4) $\mathcal{K}4\vdash\boxdot(\beta\equiv\gamma)\supset((\delta(\beta)\supset\varepsilon(\beta))\wedge(\delta(\gamma)\supset\varepsilon(\gamma)))$

from (3) by PL (as above in the proof of First Subst. Th.).

And this means:

(5) $\mathcal{K}4\vdash\boxdot(\beta\equiv\gamma)\supset(\alpha(\beta)\equiv\alpha(\gamma))$

$\alpha=\Box\delta$; and the theorem holds for δ, i.e.,

(1) $\mathcal{K}4\vdash\boxdot(\beta\equiv\gamma)\supset(\delta(\beta)\equiv\delta(\gamma))$; by ind. hyp.

(2) $\mathcal{K}4\vdash\Box\boxdot(\beta\equiv\gamma)\supset\Box(\delta(\beta)\equiv\delta(\gamma))$; (1) Deriv. 1

(3) $\mathcal{K}4\vdash\Box(\delta(\beta)\equiv\delta(\gamma))\supset(\Box\delta(\beta)\equiv\Box\delta(\gamma))$; Deriv. 1, PL

(4) $\mathcal{K}4\vdash\Box\boxdot(\beta\equiv\gamma)\supset(\Box\delta(\beta)\equiv\Box\delta(\gamma))$; (2), (3) PL

(5) $\mathcal{K}4\vdash\boxdot(\beta\equiv\gamma)\supset\Box\boxdot(\beta\equiv\gamma)$; since $\boxdot(\beta\equiv\gamma)\equiv(\Box(\beta\equiv\gamma)\wedge(\beta\equiv\gamma))$

(by Def) and $(\Box(\beta\equiv\gamma)\wedge(\beta\equiv\gamma))\supset\Box(\beta\equiv\gamma)$; by PL, and $\Box(\beta\equiv\gamma)\supset\Box\boxdot(\beta\equiv\gamma)$; by Th. 1(a) and (b).

(6) $\mathcal{K}4 \vdash \boxdot(\beta\equiv\gamma)\supset(\Box\delta(\beta)\equiv\Box\delta(\gamma))$; (4), (5) PL

i.e., $\mathcal{K}4 \vdash \boxdot(\beta\equiv\gamma)\supset(\alpha(\beta)\equiv\alpha(\gamma))$.

From the second substitution theorem a *formalized* version of the first substitution theorem can be derived as corollary.

Corollary. $\mathcal{K}4 \vdash \Box(\beta\equiv\gamma)\supset\Box(\alpha(\beta)\equiv\alpha(\gamma))$.

Proof.

(1) $\mathcal{K}4 \vdash \boxdot(\beta\equiv\gamma)\supset(\alpha(\beta)\equiv\alpha(\gamma))$; Sec. Subst. Th.

(2) $\mathcal{K}4 \vdash \Box\boxdot(\beta\equiv\gamma)\supset\Box(\alpha(\beta)\equiv\alpha(\gamma))$; (1) Deriv. 1

(3) $\mathcal{K}4 \vdash \Box(\beta\equiv\gamma)\supset\Box(\alpha(\beta)\equiv\alpha(\gamma))$; (2), Th. 1, (a) and (b).

Theorems of $\mathcal{T}$

T$_1$ $\vdash \alpha\supset\Diamond\alpha$

(1) $\Box\neg\alpha\supset\neg\alpha$; **T**, Subst.

(2) $\alpha\supset\neg\Box\neg\alpha$; (1) PL

(3) $\alpha\supset\Diamond\alpha$; (2), Def. $\Diamond$

T$_2$ $\vdash \Diamond(\alpha\supset\Box\alpha)$

(1) $\Box\alpha\supset\Diamond\Box\alpha$; by **T**$_1$

(2) $\Diamond(\alpha\supset\Box\alpha)\equiv(\Box\alpha\supset\Diamond\Box\alpha)$; **K**$_6$

(3) $\Diamond(\alpha\supset\Box\alpha)$; (1), (2), PL

T$_3$ $\Box\alpha\supset\Diamond\alpha$ (follows from **T** and **T**$_1$).

Theorems of $\mathcal{S}4$

S4$_1$ $\vdash \Diamond\Diamond\alpha\supset\Diamond\alpha$

(1) $\Box\neg\alpha\supset\Box\Box\neg\alpha$; **4**, Subst.

(2) $\neg\Box\Box\neg\alpha\supset\neg\Box\neg\alpha$; (1) PL

(3) $\Diamond\Diamond\alpha\supset\Diamond\alpha$; (2) Interch.

S4$_2$ $\vdash \Box\alpha\equiv\Box\Box\alpha$

(1) $\Box\Box\alpha\supset\Box\alpha$; **T**, Subst.: $\Box\alpha/p$

(2) $\Box\alpha\supset\Box\Box\alpha$; **4**, Subst.

(3) $\Box\alpha\equiv\Box\Box\alpha$; (1), (2) PL

S4$_3$ $\vdash \Diamond\alpha\equiv\Diamond\Diamond\alpha$

(1) $\Diamond\alpha\supset\Diamond\Diamond\alpha$; **T**$_1$, Subst.

(2) $\Diamond\alpha\equiv\Diamond\Diamond\alpha$; (1) $\mathcal{S}4_1$.

The following formulas are also the theorems of $\mathcal{S}4$ (show that):

S4$_4$ $\vdash \Box\Diamond\alpha\equiv\Box\Diamond\Box\Diamond\alpha$.

The proof does not involve the axiom **T**, and therefore it is also a theorem of $\mathcal{K}4$; for details, comp. Boolos, [1993], 9.

S4$_5$ $\vdash \Diamond\Box\alpha\equiv\Diamond\Box\Diamond\Box\alpha$

S4$_6$ $\vdash \Box(\alpha\supset\beta)\supset\Box(\Box\alpha\supset\Box\beta)$

S4$_7$ $\vdash (\Box\alpha\vee\Box\beta)\equiv\Box(\Box\alpha\vee\Box\beta)$

Theorem of $\mathcal{S}5$

S5$_1$ $\vdash \Diamond\Box\alpha\supset\Box\alpha$

(1) $\Diamond\neg\alpha\supset\Box\Diamond\neg\alpha$; **5**, Subst.
(2) $\neg\Box\Diamond\neg\alpha\supset\neg\Diamond\neg\alpha$; (1) PL
(3) $\Diamond\Box\alpha\supset\Box\alpha$; (2) Interch.

S5$_2$ $\vdash \Diamond\alpha\equiv\Box\Diamond\alpha$

(1) $\Diamond\alpha\supset\Box\Diamond\alpha$; **5**, Subst.
(2) $\Box\Diamond\alpha\supset\Diamond\alpha$; **T**, Subst.
(3) $\Diamond\alpha\equiv\Box\Diamond\alpha$; (1), (2) PL

S5$_3$ $\vdash \Box\alpha\equiv\Diamond\Box\alpha$

(1) $\Box\alpha\supset\Diamond\Box\alpha$; **T**$_1$, Subst.
(2) $\Diamond\neg\alpha\supset\Box\Diamond\neg\alpha$; **5**, Subst.
(3) $\neg\Box\Diamond\neg\alpha\supset\neg\Diamond\neg\alpha$; (2) PL
(4) $\Diamond\Box\alpha\supset\Box\alpha$; (3) Interch.
(5) $\Box\alpha\equiv\Diamond\Box\alpha$; (1), (4) PL

S5$_4$ $\vdash \Box\alpha\supset\Box\Box\alpha$

(1) $\Box\alpha\supset\Diamond\Box\alpha$; **T**$_1$, Subst.
(2) $\Diamond\Box\alpha\equiv\Box\Diamond\Box\alpha$; **S5**$_2$
(3) $\Box\alpha\supset\Box\Diamond\Box\alpha$; (1), (2) PL
(4) $\Box\alpha\supset\Box\Box\alpha$; (3), **S5**$_3$ Repl.

By this theorem it follows that **4**: $\Box p\supset\Box\Box p$ is a theorem of $\mathcal{S}5$, and then $\mathcal{S}5\supseteq\mathcal{S}4$ (since both systems contain $\mathcal{T}$ and the same rules of deduction). As can be argued $\mathcal{S}5\neq\mathcal{S}4$ since the axiom **5** is not $\mathcal{S}4$-valid (show that!).

S5$_5$ $\vdash \Box(\alpha\vee\Box\beta)\equiv(\Box\alpha\vee\Box\beta)$

(1) $\Box(\alpha\vee\Box\beta)\supset(\Box\alpha\vee\Diamond\Box\beta)$; **K**$_{10}$
(2) $\Box(\alpha\vee\Box\beta)\supset(\Box\alpha\vee\Box\beta)$; (1) **S5**$_3$, Repl.
(3) $(\Box\alpha\vee\Box\Box\beta)\supset\Box(\alpha\vee\Box\beta)$; **K**$_2$
(4) $(\Box\alpha\vee\Box\beta)\supset\Box(\alpha\vee\Box\beta)$; (3), **S4**$_2$
(5) $\Box(\alpha\vee\Box\beta)\equiv(\Box\alpha\vee\Box\beta)$; (2), (4) PL

The following formulas of L_{PML} are also theorems of $\mathcal{S}5$ (show that!):

$\mathbf{S5_6}$ $\vdash\Box(\alpha\vee\Diamond\beta)\equiv(\Box\alpha\vee\Diamond\beta)$ (use $\mathbf{S5_5}$)

$\mathbf{S5_7}$ $\vdash\Diamond(\alpha\wedge\Diamond\beta)\equiv(\Diamond\alpha\wedge\Diamond\beta)$

The theorems of $\mathcal{S}5$: $\mathbf{S5_2}$, $\mathbf{S5_3}$, $\mathbf{S4_1}$, $\mathbf{S4_2}$ show that any sequence of modal operators can be reduced in $\mathcal{S}5$ to the last operator of the sequence.

$\mathbf{S5_8}$ $\vdash\Box(\Diamond\alpha\supset\beta)\equiv\Box(\alpha\supset\Box\beta)$

The system $\mathcal{S}5$ and the system $\mathcal{B}$

The system $\mathcal{S}5$ can be axiomatized alternatively as $\mathcal{S}5^*=\mathcal{S}4+\mathbf{B}$, i.e., by adding to $\mathcal{S}4$ of the Brouwer's axiom $\mathbf{B}$: $p\supset\Box\Diamond p$. As we saw above the axiom $\mathbf{4}$: $\Box p\supset\Box\Box p$ is already a theorem of $\mathcal{S}5$. Then to prove that $\mathcal{S}5^*\vdash\alpha$ iff $\mathcal{S}5\vdash\alpha$ is enough to show the following: $\mathcal{S}5^*\vdash\mathbf{5}$: $\Diamond p\supset\Box\Diamond p$ and $\mathcal{S}5\vdash p\supset\Box\Diamond p$, and this happens as follows.

$\mathcal{S}5\vdash\alpha\supset\Box\Diamond\alpha$

(1) $\alpha\supset\Diamond\alpha$; $\mathbf{T_1}$

(2) $\Diamond\alpha\supset\Box\Diamond\alpha$; $\mathbf{5}$, Subst.

(3) $\alpha\supset\Box\Diamond\alpha$; (1), (2) PL

Hence, the axiom $\mathbf{B}$: $p\supset\Box\Diamond p$ is a theorem of $\mathcal{S}5$.

$\mathcal{S}5^*\vdash\Diamond\alpha\supset\Box\Diamond\alpha$

(1) $\Diamond\alpha\supset\Box\Diamond\Diamond\alpha$; $\mathbf{B}$, Subst.: $\Diamond\alpha/p$

(2) $\Diamond\alpha\supset\Box\Diamond\alpha$; (1), $\mathbf{S4_3}$

I.e., $\mathcal{S}5^*$ proves the axiom $\mathbf{5}$: $\Diamond p\supset\Box\Diamond p$.

The following formulas are theorems of $\mathbf{B}$ (prove that):

$\mathbf{B_1}$ $\vdash(\Diamond\Box\alpha\wedge\Diamond\Box\beta)\supset\Box\Diamond(\alpha\wedge\beta)$

$\mathbf{B_2}$ $\vdash\Diamond\Box\alpha\supset\Box\Diamond\alpha$

Theorems of $\mathcal{GL}$

$\mathbf{GL_1}$ $\vdash\Box\alpha\supset\Box\Box\alpha$[7]

(1) $\alpha\supset((\Box\Box\alpha\wedge\Box\alpha)\supset(\Box\alpha\wedge\alpha))$; by PL:

$p\supset((q\wedge r)\supset(r\wedge p))$ + Subst.

(2) $\Box\alpha\supset\Box((\Box\Box\alpha\wedge\Box\alpha)\supset(\Box\alpha\wedge\alpha))$; (1) Deriv. 1

(3) $\Box\alpha\supset\Box(\Box(\Box\alpha\wedge\alpha)\supset(\Box\alpha\wedge\alpha))$; (2), $\mathbf{K_1}$, Repl.

[7] This theorem of $\mathcal{GL}$ was independently proved by D. de Jongh, S. Kripke and G. Sambin (apud. G. Boolos, [1993], 11). Since $\mathbf{4}$: $\Box p\supset\Box\Box p$ is provable in $\mathcal{GL}$ sometimes $\mathcal{GL}$ is known as the following extension of $\mathcal{K}$: $\mathcal{K}$+$\mathbf{4}$+$\mathbf{W}$ (where $\mathcal{K}$+$\mathbf{4}$=$\mathcal{K}4$ (=BML: Basic Modal Logic)); comp. C. Smorynski, [1985], 65, 71.

(4) $\Box(\Box(\Box\alpha\wedge\alpha)\supset(\Box\alpha\wedge\alpha))\supset\Box(\Box\alpha\wedge\alpha)$; **W**, Subst.: $\Box\alpha\wedge\alpha/p$

(5) $\Box\alpha\supset\Box(\Box\alpha\wedge\alpha)$; (3), (4) PL

(6) $\Box\alpha\supset(\Box\Box\alpha\wedge\Box\alpha)$; (5), $\mathbf{K}_1$, Repl.

(7) $\Box\alpha\supset\Box\Box\alpha$; (6), PL

$\mathbf{GL}_2$ $\vdash\Box(\Box\alpha\supset\alpha)\equiv\Box\alpha$

(1) $(\Box\alpha\wedge\alpha)\supset(\Box\alpha\supset\alpha)$; PL

(2) $\Box(\Box\alpha\wedge\alpha)\supset\Box(\Box\alpha\supset\alpha)$; (1) Deriv. 1

(3) $\Box\alpha\supset\Box(\Box\alpha\wedge\alpha)$; comp. formula (5) in the proof of $\mathbf{GL}_1$

(4) $\Box\alpha\supset\Box(\Box\alpha\supset\alpha)$; (2), (3), PL

(5) $\Box(\Box\alpha\supset\alpha)\supset\Box\alpha$; **W**, Subst.

(6) $\Box(\Box\alpha\supset\alpha)\equiv\Box\alpha$; (4), (5), PL

$\mathbf{GL}_3$ $\vdash\Box(\Box\alpha\wedge\alpha)\equiv\Box\alpha$

(1) $\Box(\Box\alpha\wedge\alpha)\equiv(\Box\Box\alpha\wedge\Box\alpha)$; $\mathbf{K}_1$

(2) $(\Box\Box\alpha\wedge\Box\alpha)\supset\Box\alpha$; PL

(3) $\Box(\Box\alpha\wedge\alpha)\supset\Box\alpha$; (1), (2), PL

$((p\equiv q)\supset[(q\supset r)\supset(p\supset r)])$

(4) $\Box\alpha\supset\Box(\Box\alpha\wedge\alpha)$; from (3) of $\mathbf{GL}_2$

(5) $\Box(\Box\alpha\wedge\alpha)\equiv\Box\alpha$; (3), (4), PL[8]

$\mathbf{GL}_4$ $\vdash\Box\bot\equiv\Box\Diamond p$[9]

(1) $\bot\supset\Diamond p$; PL

(2) $\Box\bot\supset\Box\Diamond p$; (1) Deriv. 1

(3) $p\supset\top$; PL

(4) $\Diamond p\supset\Diamond\top$; (3) Deriv. 3

(5) $\Diamond\top\equiv\neg\Box\bot\equiv\Box\bot\supset\bot$; Interch., PL

(6) $\Diamond p\supset(\Box\bot\supset\bot)$; (4), (5), Repl.

(7) $\Box\Diamond p\supset\Box(\Box\bot\supset\bot)$; (6) Deriv. 1

(8) $\Box(\Box\bot\supset\bot)\supset\Box\bot$; **W**, Subst.

(9) $\Box\Diamond p\supset\Box\bot$; (7), (8), PL

---

[8] Note that since by $\mathbf{GL}_1$ the system $\mathcal{GL}$ extends $\mathcal{K}4$, it follows that $\mathbf{GL}_2$ follows from the step (3) of the given proof (i.e., from **Th1** ((a) and (b)) of $\mathcal{K}4$) *via* **W** and PL. Similarly, $\mathbf{GL}_3$ follows directly from **Th1** (a) and (b) of $\mathcal{K}4$. Generally, since $\mathcal{GL}$ extends $\mathcal{K}4$, all the theorems of $\mathcal{K}4$ are also theorems of $\mathcal{GL}$.

[9] As we mentioned at the beginning of this section, the symbols "$\bot$" and "$\top$" denote the logical falsity and the logical truth, respectively. $\mathbf{GL}_4$ is Theorem 21 of Boolos [1993], 12.

(10) $\Box\bot\equiv\Box\Diamond p$; (2), (9), PL

GL$_5$ (a) $\vdash\Box(p\equiv\neg\Box p)\equiv\Box(p\equiv\neg\Box\bot)$

(b) $\vdash\Box(p\equiv\Box p)\equiv\Box(p\equiv\top)$

(c) $\vdash\Box(p\equiv\Box\neg p)\equiv\Box(p\equiv\Box\bot)$

(d) $\vdash\Box(p\equiv\neg\Box\neg p)\equiv\Box(p\equiv\bot)$[10]

Proofs (see G. Boolos [1993], 13-14, Theorem 24).

4.1.2. Semantics of modal logic

While in the preceding section we saw the syntactical aspects of the best-known system of modal logic, in this section we set out the semantics for these systems, and then the way the two sections are correlated by the soundness and completeness theorems.[11]

4.1.2.1. Concepts. Soundness theorem for $\mathcal{K}$

$W=\{w_1,w_2,...\}$ is a nonempty set of possible worlds. R is a dyadic relation defined over W[12] called the *accessibility* relation: when w_1Rw_2 we say that w_2 is *accessible* from w_1 or it is a *possible world relative to* w_1 or, finally, that w_1 *can see* w_2.

If R is a relation defined over W, then:

R is *reflexive* if for all $w\in W$: wRw.

R is *irreflexive* if for no w: wRw.

R is *symmetric* if for all w_1, w_2: if w_1Rw_2, then w_2Rw_1.

R is *transitive* if for all w_1, w_2, w_3: if w_1Rw_2 and w_2Rw_3, then w_1Rw_3.

R is *euclidean* if for all w_1, w_2, w_3: if w_1Rw_2 and w_1Rw_3, then w_2Rw_3 (or w_3Rw_2).

[10] **GL**$_5$ (a)–(d) will play a special role in "decoding" of the so-called self-referential (constant) sentences of L_{PA}; see the next section 4.2.

[11] The section 4.1.2 contains the basic aspects of this semantics, as it is taken today as being the standard one. The elaboration of these items is based in essence on the following sources: G. Boolos [1993], Chs. 4-6, G.E. Hughes and M.J. Cresswell [1996], Chs. 1-3, 6-8, C. Smorynski [1985], Chs. 1-3, M. Fitting [1993].

[12] In the sense that for every pair (w_1, w_2) we can tell whether or not w_1Rw_2.

A *frame* is an ordered pair $\langle W, R \rangle$, where W and R are the items defined above.

We say of a frame $\langle W, R \rangle$ that it has a property P (e.g., reflexivity, transitivity) iff R has this property.

A frame $\langle W, R \rangle$ is finite iff W is finite.

A *valuation* V on W is an assignment of values to the propositional variables in the worlds of W. By wVp we mean that "w verifies p", i.e., p is true in w.

The triple $\langle W, R, V \rangle$ is a *model based on* the frame $\langle W, R \rangle$, and V is a valuation function.

The meaning of the relation $w \models \alpha$ can be stated precisely by the following inductive definition.

Definition 1. *Let* $M = \langle W, R, V \rangle$ *be a model and* $w \in W$. *Then*

(1) $w \models \mathrm{p}$ *iff* wVp.

(2) $w \models \neg\alpha$ *iff* $w \not\models \alpha$.

(3) $w \models \alpha \supset \beta$ *iff either* $w \not\models \alpha$ *or* $w \models \beta$.

(4) $w \models \Box\alpha$ *iff for any* $w_i \in W$ *such that* wRw_i, $w_i \models \alpha$.

Similarly, the truth values of the formulas (not in the primitive notation) $\alpha\wedge\beta$, $\alpha\vee\beta$, $\alpha\equiv\beta$, $\Diamond\alpha$ in a world w can be defined (as consequences of 1-4) (exercise).

Definition 2. *A formula* α *is valid in a model* $M = \langle W, R, V \rangle$ *iff* $w \models \alpha$ *for all* $w \in W$.

Definition 3. *A formula* α *is valid in a frame* $\langle W, R \rangle$ *iff* α *is valid in all models based on* $\langle W, R \rangle$.

The concept of validity in modal logic is relative to the modal systems. Since the modal systems we are interested in are extensions of $\mathcal{K}$, let us firstly consider in detail the validity of $\mathcal{K}$ (or $\mathcal{K}$-validity) and the *soundness theorem* for $\mathcal{K}$.

Definition. *A formula* α *of* L$_{\mathrm{PML}}$ *is* $\mathcal{K}$*-valid iff* α *is valid in any frame* $\langle W, R \rangle$.

The soundness theorem for $\mathcal{K}$. *If* $\mathcal{K} \vdash \alpha$, *then* α *is valid in any frame* $\langle W, R \rangle$.

To prove the soundness of $\mathcal{K}$ boils down to show the following:

(1) All axioms of $\mathcal{K}$ are $\mathcal{K}$-valid.

(2) The rules of deduction (Subst, MP and N) are valid (i.e., they preserve in the conclusion the validity of premise(s)).

(1) *All axioms of $\mathcal{K}$ are $\mathcal{K}$-valid.*

(a) All *valid* formulas of PL are $\mathcal{K}$-valid. This follows from the fact that the validity of a formula α in a world *w* does not require in any way a reference to another world; therefore α will be true in every world and then valid in every model and, finally, α will be valid in every frame $\langle W, R\rangle$.

(b) The axiom **K**: □(p⊃q)⊃(□p⊃□q) is $\mathcal{K}$-valid.

(*Reductio*). Suppose that **K** is not $\mathcal{K}$-valid, i.e., there is a frame $\langle W, R\rangle$ in which **K** is not valid. Then there is a model $\langle W, R, V\rangle$ based on $\langle W, R\rangle$ in which **K** is not valid, and therefore there is a world $w \in W$ in which **K** is false,[13] i.e., (1) $w \nvDash$ **K**. From (1) it follows (2) $w \vDash$ □(p⊃q) and (3) $w \nvDash$ (□p⊃□q). From (3) it follows (4) $w \vDash$ □p and (5) $w \nvDash$ □q. From (5) it follows (by Def. 1 (4)) (6) there is a world $w_i \in W$ such that wRw_i and $w_i \nvDash$ q. From (4) it follows (7) $w_i \vDash$ p (since wRw_i) and then (8) $w_i \nvDash$ (p⊃q) (from (6) and (7)). Now, since wRw_i it follows that $w \nvDash$ □(p⊃q); contradicting (2).

(2) *All deductive rules are $\mathcal{K}$-valid.*

(a) MP is a valid rule of deduction.

Proof. If α and α⊃β are valid in a frame $\langle W, R\rangle$, then both formulas are valid in every model based on $\langle W, R\rangle$ and so they are true in every $w \in W$. Therefore, by Def. 1(3), β is true in any $w \in W$. Whence β is also valid in $\langle W, R\rangle$.

(b) *Subst* is a valid rule of deduction.

Let α(p) be a formula of L_{PML} containing the propositional variable p, and α(β/p) (α(β) for short) be the formula obtained from α(p) by substituting a formula β for p (in all of its occurrences) in α(p). Then the following result holds.

Lemma. *If* α(p) *is valid in the frame* $\langle W, R\rangle$, *then* α(β) *is also valid in* $\langle W, R\rangle$.

Proof. Let $\langle W, R\rangle$ be a frame in which α(p) is valid. Let $M = \langle W, R, V\rangle$ be an arbitrary model based on $\langle W, R\rangle$. Let $w \in W$ be an arbitrary world. Let

[13] We often refer to a *possible* world relative to the world *w* using the symbols *w*', w_i, w_j, w^*.

$M^* = \langle W, R, V^* \rangle$ be a model based on the same frame $\langle W, R \rangle$, where V^* is defined as follows:

wV^*p (in M^*) iff $w \models \beta$ (in M), and

wV^*q (in M^*) iff wVq (for q $\neq$ p).

By induction on the complexity of α(p) the following equivalence can be argued:

(Eq) $w \models \alpha$(p) (in M^*) iff $w \models \alpha(\beta)$ (in M).

We have the following cases:

(1) α(p) = p; then $w \models$ p (in M^*) iff $w \models \beta$ (in M) (by def. of V^*).

(2) α(p) = q (q $\neq$ p); then $w \models$ q (in M^*) iff $w \models$ q (in M) (by def. of V^*).

(3) $\alpha(\mathrm{p}) = \neg\gamma(\mathrm{p})$; and the lemma holds for γ(p), i.e.,
$w \models \gamma$(p) (in M^*) iff $w \models \gamma(\beta)$ (in M). So,
$w \models \neg\gamma$(p) (in M^*) iff $w \models \neg\gamma(\beta)$ (in M) (by PL), i.e.,
$w \models \alpha$(p) (in M^*) iff $w \models \alpha(\beta)$ (in M).

(4) $\alpha(\mathrm{p}) = \gamma \supset \delta$; and the lemma holds for γ and δ, i.e.,
$w \models \gamma$(p) (in M^*) iff $w \models \gamma(\beta)$ (in M), and
$w \models \delta$(p) (in M^*) iff $w \models \delta(\beta)$ (in M); whence by PL:
$w \models \gamma(\mathrm{p}) \supset \delta(\mathrm{p})$ (in M^*) iff $w \models \gamma(\beta) \supset \delta(\beta)$ (in M) i.e.,
$w \models \alpha$(p) (in M^*) iff $w \models \alpha(\beta)$ (in M).

(5) $\alpha(\mathrm{p}) = \Box\gamma(\mathrm{p})$; and the lemma holds for γ(p).

We have: $w \models \Box\gamma$(p) (in M*) iff for every w' such that wRw', $w' \models \gamma$(p) (in M*) iff $w' \models \gamma(\beta)$ (by ind. hyp.) (in M) iff $w \models \Box\gamma(\beta)$ (in M).

Now, since α(p) is valid in $\langle W, R \rangle$ (by hypothesis of lemma) it follows that α(p) is valid in M^*. And therefore, by (Eq) $\alpha(\beta)$ is valid in M. And since in the argument above M and w were arbitrary, it follows that $\alpha(\beta)$ is valid in $\langle W, R \rangle$.

(c) N *is a valid rule of deduction*.

Proof. If α is valid in $\langle W, R \rangle$, then α is valid in every model based on $\langle W, R \rangle$, and then α is true in every $w \in W$. So, for every $w_i \in W$, such that wRw_i: $w_i \models \alpha$, and then $w_i \models \Box\alpha$. Whence $\Box\alpha$ is valid in $\langle W, R \rangle$.

4.1.2.2. Theorems

As we saw above, any formula α provable in $\mathcal{K}$ is $\mathcal{K}$-valid, i.e., valid in any frame $\langle W,R\rangle$. And then $\mathcal{K}$ is a sound system of modal logic. Of a frame $\langle W,R\rangle$ we say that it is *a frame for a modal system* $\mathcal{S}$ if any theorem of $\mathcal{S}$ is valid on that frame. Now, since all systems considered here, other than $\mathcal{K}$, are proper extensions of $\mathcal{K}$, to verify that a frame $\langle W,R\rangle$ is a frame for $\mathcal{S}$ is to only verify that the proper axioms of $\mathcal{S}$ are valid on that frame.

An important logical problem is to establish whether the following co-extensivity holds:

The class $\mathcal{C}$ of all frames for $\mathcal{S}$ = the class $\mathcal{C}^*$ of all frames having the property (properties) *P*, equivalently: $\langle W,R\rangle$ is a frame for $\mathcal{S}$ iff *R* has the property (properties) *P*. And to establish this means to establish that the *proper* axioms of $\mathcal{S}$ are valid in $\langle W,R\rangle$ iff *R* has the property (properties) *P*, i.e., to establish the equivalences of the following form:

$\mathrm{PrAx}_{\mathcal{S}}$ *are valid in* $\langle W,R\rangle$ *iff R has the property* (*properties*) *P*.

Let us in what follows consider these equivalences for any modal system other than $\mathcal{K}$.

Theorem 1. **T**: $\Box p\supset p$ *is valid in* $\langle W,R\rangle$ *iff* $\langle W,R\rangle$ *is reflexive*.

This theorem follows from the proof of the following conditionals:

(a) If **T**: $\Box p\supset p$ is valid in $\langle W,R\rangle$, then $\langle W,R\rangle$ is reflexive.

(b) If $\langle W,R\rangle$ is reflexive, then **T**: $\Box p\supset p$ is valid in $\langle W,R\rangle$.

Proof. (a) Assume hypothesis. Let $\langle W,R,V\rangle$ be a model based on $\langle W,R\rangle$, $w\in W$ an arbitrary world and *V* defined as follows: w_iVp iff wRw_i, i.e.,

(Eq)[14] $w_i\models$ p iff wRw_i.

Now, if for $w_i\in W$ such that wRw_i $w_i\models$ p it follows that $w\models\Box$p. And since $w\models\Box p\supset p$ (by hyp.) it follows that $w\models$ p. Whence, by (Eq), *wRw*, i.e., $\langle W,R\rangle$ is reflexive.

Remark. A variant of proof can also be given by contraposition in the following way. We keep the above model $\langle W,R,V\rangle$ and consider this time that $\langle W,R\rangle$ is not reflexive, i.e., there is just a world $w\in W$ such that not

[14] In what follows we refer to "*wV*p" (or "*V*(p,*w*)=1") only using, *via* Def. 1, the notation $w\models$ p, with no risk of introducing any confusion. The hint the valuation function *V* is defined each time is borrowed from Boolos [1993], 73-74.

wRw. Then from (Eq) and not wRw it follows that $w \models \Box$p (since for any $w_i \neq w$ such that wRw_i: $w_i \models$ p) and $w \not\models$ p, respectively. Therefore $w \not\models \Box$p$\supset$p, i.e., **T**: $\Box$p$\supset$p is not valid in $\langle W, R\rangle$.

Proof. (b) Assume hypothesis. Let $\langle W, R, V\rangle$ be an arbitrary model based on $\langle W, R\rangle$ and $w \in W$ an arbitrary world such that $w \models \Box\alpha$. Then for every $w_i \in W$ such that wRw_i $w_i \models$ p. Since $\langle W, R\rangle$ is reflexive (by hyp.), i.e., wRw holds, it follows that $w \models$ p. And then $w \models \Box$p$\supset$p. Since $\langle W, R, V\rangle$ and w were arbitrary, it follows that **T**: $\Box$p$\supset$p is valid in $\langle W, R\rangle$.

A proof by *reductio* can also be given (exercise).

Theorem 2. **4**: $\Box$p$\supset\Box\Box$p *is valid in* $\langle W, R\rangle$ *iff* $\langle W, R\rangle$ *is transitive*.

As above, we have to prove:

(a) If $\Box$p$\supset\Box\Box$p is valid in $\langle W, R\rangle$, then $\langle W, R\rangle$ is transitive.

(b) If $\langle W, R\rangle$ is transitive, then $\Box$p$\supset\Box\Box$p is valid in $\langle W, R\rangle$.

Proof. (a) Assume hypothesis. Let $\langle W, R, V\rangle$ be a model based on $\langle W, R\rangle$, $w_1, w_2, w_3 \in W$ such that w_1Rw_2 and w_2Rw_3 and V defined as follows:

(Eq) $w_i \models$ p iff w_1Rw_i.

Then $w_1 \models \Box$p (by (Eq)). And since $w_1 \models \Box$p$\supset\Box\Box$p (by hyp.), it follows that $w_1 \models \Box\Box$p. By hypothesis w_1Rw_2, and then $w_2 \models \Box$p and since w_2Rw_3 (by hyp.), we have $w_3 \models$ p. Whence, by (Eq) w_1Rw_3, i.e., $\langle W, R\rangle$ is transitive (since the model $\langle W, R, V\rangle$ and w_1, w_2 and w_3 are arbitrary).

Remark. A proof for (a) can also be given by *reductio*. Consider the model M defined above and that $\langle W, R\rangle$ is not transitive, i.e., there are $w_1, w_2, w_3 \in W$ such that w_1Rw_2, w_2Rw_3 and *not* w_1Rw_3. We also assume the hypothesis of (a), i.e., **4**: $\Box$p$\supset\Box\Box$p is valid in $\langle W, R\rangle$. Then, for this setting a contradiction can be derived in the following way. By (Eq), $w_1 \models \Box$p. But since not w_1Rw_3 and w_2Rw_3 (by hyp.) it follows that $w_3 \not\models$ p (by (Eq)) and then $w_2 \not\models \Box$p, respectively. Again, by w_1Rw_2 and $w_2 \not\models \Box$p, it follows that $w_1 \not\models \Box\Box$p. Hence, finally, $w_1 \not\models \Box$p$\supset\Box\Box$p, contrary to the hypothesis.

Proof. (b) Assume hypothesis, i.e., $\langle W, R\rangle$ is transitive. Let $\langle W, R, V\rangle$ be an arbitrary model based on $\langle W, R\rangle$ and $w_1 \in W$ an arbitrary world such that

$w_1 \models \Box$p. If $w_1 R w_2$ then $w_2 \models$p; and if $w_2 R w_3$, then $w_1 R w_3$ (by transitivity) and therefore $w_3 \models$p and then $w_2 \models \Box$p. It follows that $w_1 \models \Box\Box$p; whence $w_1 \models \Box$p$\supset\Box\Box$p. Therefore, **4**: $\Box$p$\supset\Box\Box$p is valid in $\langle W, R\rangle$.

Remark. A proof of (b) can also be given by *reductio* (exercise).

Theorem 3. 5: $\Diamond$p$\supset\Box\Diamond$p *is valid in* $\langle W, R\rangle$ *iff* $\langle W, R\rangle$ *is euclidean*.

(a) If $\Diamond$p$\supset\Box\Diamond$p is valid in $\langle W, R\rangle$, then $\langle W, R\rangle$ is euclidean.

(b) If $\langle W, R\rangle$ is euclidean, then $\Diamond$p$\supset\Box\Diamond$p is valid in $\langle W, R\rangle$.

Proof. (a) Suppose that $\Diamond$p$\supset\Box\Diamond$p is valid in $\langle W, R\rangle$. Let $\langle W, R, V\rangle$ be a model based on $\langle W, R\rangle$ such that $w_1 R w_3$, $w_1 R w_2$ and V is defined by

(Eq) $w_i \models$p iff $w_i = w_3$.

Since $\Diamond$p$\supset\Box\Diamond$p is valid in $\langle W, R\rangle$ it follows that $w_1 \models \Diamond$p$\supset\Box\Diamond$p. And since $w_3 \models$p (by (Eq)) and $w_1 R w_3$ (by hyp.) it follows that $w_1 \models \Diamond$p. And then $w_1 \models \Box\Diamond$p. And since $w_1 R w_2$ (by hyp.) it follows that $w_2 \models \Diamond$p. Hence there is a world w_i such that $w_2 R w_i$ and $w_i \models$p. By (Eq), $w_i = w_3$, and then $w_2 R w_3$ holds; i.e., $\langle W, R\rangle$ is euclidean.

Proof. (b) (*reductio*). Suppose that $\langle W, R\rangle$ is euclidean and $\Diamond$p$\supset\Box\Diamond$p is not valid in $\langle W, R\rangle$. Then there is a model based on $\langle W, R\rangle$ and $w_1 \in W$ such that

(1) $w_1 \not\models \Diamond$p$\supset\Box\Diamond$p.

(2) $w_1 \models \Diamond$p, and

(3) $w_1 \not\models \Box\Diamond$p.

(4) There is a world w_2 such that $w_1 R w_2$ and $w_2 \models$p; from (2).

(5) There is a world w_3 such that $w_1 R w_3$ and $w_3 \not\models \Diamond$p; from (3).

(6) From $w_1 R w_2$ and $w_1 R w_3$ it follows that $w_3 R w_2$ (since R is euclidean).

(7) $w_2 \not\models p$; from (5) and (6).

But (7) and (4) are contradictory.

Theorem 4. B: p$\supset\Box\Diamond$p *is valid in* $\langle W, R\rangle$ *iff* $\langle W, R\rangle$ *is symmetrical*.

(a) If **B**: p$\supset\Box\Diamond$p is valid in $\langle W, R\rangle$ then $\langle W, R\rangle$ is symmetrical.

(b) If $\langle W, R\rangle$ is symmetrical, then **B**: p$\supset\Box\Diamond$p is valid in $\langle W, R\rangle$.

Proof. (a) Assume that p⊃□◇p is valid in $\langle W, R\rangle$. Let $\langle W, R, V\rangle$ be a model based on $\langle W, R\rangle$, $w_1, w_2 \in W$ such that $w_1 R w_2$ and V defined as follows:

(Eq) $w_i \models$ p iff $w_i = w_1$.

Now, since p⊃□◇p is valid in $\langle W, R\rangle$ it follows that $w_1 \models$ p⊃□◇p. By (Eq) it follows that $w_1 \models$ p. And then $w_1 \models$ □◇p (by MP). Since $w_1 R w_2$ (by hyp.) it follows that $w_2 \models$ ◇p. Hence there is a world w_i such that $w_2 R w_i$ and $w_i \models$ p. It follows that $w_i = w_1$ (by (Eq)). Therefore, $w_2 R w_1$ holds, and then $\langle W, R\rangle$ is symmetrical.

Proof. (b) Suppose $\langle W, R\rangle$ is a symmetrical frame. Let $\langle W, R, V\rangle$ be a model based on $\langle W, R\rangle$ and $w_1 \in W$ such that $w_1 \models$ p, and $w_1 R w_2$. Then since $\langle W, R\rangle$ is symmetrical, it follows that $w_2 R w_1$, and since p is true in w_1 it follows that $w_2 \models$ ◇p for every w_2 such that $w_1 R w_2$. Hence $w_1 \models$ □◇p, and therefore $w_1 \models$ p⊃□◇p. Since $\langle W, R, V\rangle$ is arbitrary and since if in a world p is true then □◇p is also true, it follows that p⊃□◇p is valid in $\langle W, R\rangle$.

Remark. From the theorems 1-4 immediately follows the *soundness theorems* for all the systems considered here (cf. 4.1.1). Since all of these systems are proper extensions of $\mathcal{K}$, and $\mathcal{K}$ was proved to be sound (by 4.1.2.1), the proof of soundness for these modal systems boils down to show for each of them that their proper axioms are valid according to the respective definition of validity. For $\mathcal{T}$, for example, we must prove that any theorem of $\mathcal{T}$ is $\mathcal{T}$-valid; i.e., if $\mathcal{T} \vdash \alpha$ and $\langle W, R\rangle$ is reflexive, then α is $\mathcal{T}$-valid. And this is reducible to show that if $\langle W, R\rangle$ is reflexive, then the proper axiom of $\mathcal{T}$, □p ⊃ p, is valid. But this is proved by Theorem 1(b). Such is the case with all the other modal systems considered above.[15]

4.1.3. Completeness of modal systems

As we saw in the preceding section all modal systems considered are *sound*. This means that all the provable formulas in each of them are valid according to the corresponding concept of validity. This section has the goal

[15] The special case of $\mathcal{GL}$ will be treated in Sect. 4.1.3.3 below.

to show the converse: each of these systems is also complete; i.e., if $\mathcal{S}$ is such a system, then all its valid formulas are provable in $\mathcal{S}$. The two properties, soundness and completeness, together, show that the following co-extensivity holds: $\mathcal{S} \vdash \alpha$ iff α is $\mathcal{S}$-valid (i.e., the set of provable formulas of $\mathcal{S}$ and the set of valid formulas of $\mathcal{S}$ do coincide).

4.1.3.1. Canonical models

An elegant tool for proving the completeness is the use of *canonical models* for these systems. Each such system $\mathcal{S}$ has a canonical model M_{can} with the following remarkable property:[16]

(*Can*) α is valid in M_{can} iff $\mathcal{S} \vdash \alpha$.

The key step of the proof of completeness using this technique is the following: show firstly that the frame $\langle W, R\rangle$ of the canonical model of $\mathcal{S}$ has the properties required by the respective concept of validity (reflexivity, transitivity, symmetry, etc.). Then we reason as follows: since $\mathcal{S}$ is sound, a formula α provable in $\mathcal{S}$ will be $\mathcal{S}$-valid, i.e., α is valid in *any* frame $\langle W, R\rangle$ satisfying the particular properties (refl., trans., etc.), and therefore α is valid in the frame of canonical model; and this means that α is valid in the canonical model for $\mathcal{S}$. Whence, by (*Can*), α is provable in $\mathcal{S}$.

Let us see what is a canonical model.

Maximal consistent sets of formulas

If in the preceding considerations (4.1.2.1) the worlds were simply points in a set of possible worlds, this time we are also interested in their "nature". Then a natural way to see what is a world is to consider it as a consistent and complete set of formulas (propositions) describing it. Let us detail.

Definition 1. *A set* Γ *of modal formulas is said to be* $\mathcal{S}$*-inconsistent* ($\mathcal{S}$*-inc*, *for short*) *iff there is a finite set* $\Gamma_0 = \{\alpha_1,...,\alpha_n\}$ of Γ (i.e., $\Gamma_0 \subseteq \Gamma$) *such that*

$$\mathcal{S} \vdash \neg(\alpha_1 \wedge ... \wedge \alpha_n),$$

otherwise Γ *is* $\mathcal{S}$*-consistent* ($\mathcal{S}$*-con*, *for short*).

Let us abbreviate $\alpha_1 \wedge ... \wedge \alpha_n$ by Conj(Γ_0). Then, evidently, Conj(Γ_0)$\wedge\alpha$ is Conj($\Gamma_0 \cup \alpha$).

[16] A fact proved by Lemma 6 below.

Obvious, if $\Gamma = \{\alpha\}$, then Γ is $\mathcal{S}$-inconsistent iff $\mathcal{S} \vdash \neg\alpha$. And if $\mathcal{S} \nvdash \alpha$, then $\neg\alpha$ is $\mathcal{S}$-con (argue!).

Definition 2. Γ *is said to be maximal (max, for short) iff for every formula* α: $\alpha \in \Gamma$ *or* $\neg\alpha \in \Gamma$.

By Def. 1 and Def. 2, Γ is a maximal $\mathcal{S}$-consistent set (abrev. *max $\mathcal{S}$-con*) iff Γ is maximal and $\mathcal{S}$-consistent.

Lemma 1. *If* Γ *is an $\mathcal{S}$-con set of formulas and* α *any formula, then* $\Gamma \cup \{\alpha\}$ *is $\mathcal{S}$-con or* $\Gamma \cup \{\neg\alpha\}$ *is $\mathcal{S}$-con.*

Proof (*reductio***)** Assume hypothesis and that both $\Gamma \cup \{\alpha\}$ and $\Gamma \cup \{\neg\alpha\}$ are $\mathcal{S}$-inc. Then there are finite sets $\Gamma_0^1 \subseteq \Gamma$ and $\Gamma_0^2 \subseteq \Gamma$ such that

(1) $\vdash \neg(\mathrm{Conj}(\Gamma_0^1) \wedge \alpha)$; by Def.1

(2) $\vdash \neg(\mathrm{Conj}(\Gamma_0^2) \wedge \neg\alpha)$; by Def.1

(3) $\vdash \mathrm{Conj}(\Gamma_0^1) \supset \neg\alpha$; (1) by PL

(4) $\vdash \mathrm{Conj}(\Gamma_0^2) \supset \alpha$; (2) by PL, and then

(5) $\vdash (\mathrm{Conj}(\Gamma_0^1) \wedge \mathrm{Conj}(\Gamma_0^2)) \supset (\alpha \wedge \neg\alpha)$; by PL

(6) $\vdash \neg(\mathrm{Conj}(\Gamma_0^1) \wedge \mathrm{Conj}(\Gamma_0^2))$; (5) by PL

I.e., $\vdash \neg\mathrm{Conj}(\Gamma_0^1 \cup \Gamma_0^2)$. But $\Gamma_0^1 \cup \Gamma_0^2 \subseteq \Gamma$, and therefore Γ is $\mathcal{S}$-inc.

Lemma 2. *Let* Γ *be any max $\mathcal{S}$-con set. Then the following holds:*

(1) *For any formula* α *strictly one member of* $\{\alpha, \neg\alpha\}$ *is in* Γ.

(2) $\alpha \supset \beta \in \Gamma$ *iff* $\alpha \notin \Gamma$ *or* $\beta \in \Gamma$.

(3) *If* $\mathcal{S} \vdash \alpha$, *then* $\alpha \in \Gamma$.

(4) *If* $\alpha \in \Gamma$ *and* $\mathcal{S} \vdash \alpha \supset \beta$, *then* $\beta \in \Gamma$.

Proof. (1) follows immediately from $\mathcal{S}$-consistency and maximality.

(2) $\alpha \supset \beta \in \Gamma$ iff $\neg\alpha \vee \beta \in \Gamma$ iff $\neg\alpha \in \Gamma$ or $\beta \in \Gamma$ (argue this!) iff $\alpha \notin \Gamma$ (by $\mathcal{S}$-con) or $\beta \in \Gamma$.

(3) $\mathcal{S} \vdash \alpha$ iff $\mathcal{S} \vdash \neg(\neg\alpha)$ iff $\neg\alpha$ is $\mathcal{S}$-inc iff $\neg\alpha \notin \Gamma$ iff $\alpha \in \Gamma$.

(4) (Exercise).

Lemma 3. *Let* Δ *be any $\mathcal{S}$-con set. Then there exist a max $\mathcal{S}$-con set* Γ *such that* $\Delta \subseteq \Gamma$.

Proof. Let us consider an enumeration of all modal formulas: $\gamma_1, \gamma_2, \gamma_3, \ldots$ Define a sequence $\Gamma_0, \Gamma_1, \Gamma_2, \ldots$ of sets of formulas as follows:

$$\Gamma_0 = \Delta$$

$$\Gamma_{n+1} = \begin{cases} \Gamma_n \cup \{\gamma_{n+1}\} & \text{if it is } \mathcal{S}\text{-con} \\ \Gamma_n \cup \{\neg\gamma_{n+1}\} & \text{otherwise.} \end{cases}$$

Let $\Gamma = \cup\Gamma_n$ $(n = 0,1,2,\ldots)$.

We have the following:

(1) Every Γ_n is $\mathcal{S}$-con, since $\Gamma_0 = \Delta$ is consistent (by hyp.). And if Γ_n is consistent, then either $\Gamma_n \cup \{\gamma_{n+1}\}$ is consistent and then $\Gamma_{n+1} = \Gamma_n \cup \{\gamma_{n+1}\}$ (by def.), or $\Gamma_n \cup \{\gamma_{n+1}\}$ is $\mathcal{S}$-inc and then $\Gamma_{n+1} = \Gamma_n \cup \{\neg\gamma_{n+1}\}$ is $\mathcal{S}$-con (by Lemma 1).

(2) Γ *is $\mathcal{S}$-con*. Otherwise, there would be an $\mathcal{S}$-inc finite subset of Γ. But every subset of Γ is some Γ_n and then such Γ_n would be $\mathcal{S}$-inc (contradicting (1)).

(3) Γ *is maximal*. Since if for some formula γ_{k+1}, $\gamma_{k+1} \notin \Gamma$, then $\gamma_{k+1} \notin \Gamma_{k+1}$, and this means that $\Gamma_k \cup \{\gamma_{k+1}\}$ is $\mathcal{S}$-inc, and therefore $\Gamma_k \cup \{\neg\gamma_{k+1}\} = \Gamma_{k+1}$ is $\mathcal{S}$-con, in which case $\neg\gamma_{k+1} \in \Gamma$.

Let us observe that the above lemmas are not specifically modal; i.e., they hold for any system $\mathcal{S}$ containing PL (comp. Ch. 1, Sect. 3.3.2). Instead, for the following lemma, $\mathcal{S}$ will be any normal modal system. First of all, a definition:

$\Gamma_\Box = \{\alpha | \Box\alpha \in \Gamma\}$ ($w_\Box$ will have the corresponding meaning).

Lemma 4. *Let* Γ *be any $\mathcal{S}$-con set such that* $\neg\Box\alpha \in \Gamma$. *Then* $\Gamma_\Box \cup \{\neg\alpha\}$ *is $\mathcal{S}$-con.*

Proof (*reductio*). Assume hypothesis and that the set $\Gamma_\Box \cup \{\neg\alpha\}$ is $\mathcal{S}$-inc. Then there is a finite $\Gamma_0 \subseteq \Gamma_\Box$ such that:

(1) $\vdash \neg(\mathrm{Conj}(\Gamma_0) \wedge \neg\alpha)$; by Def.1

(2) $\vdash \mathrm{Conj}(\Gamma_0) \supset \alpha$; (1) PL

(3) $\vdash \Box\,\mathrm{Conj}(\Gamma_0) \supset \Box\alpha$; (2) Deriv.1

But if $\Gamma_0 = \{\alpha_1, \ldots, \alpha_n\}$, then $\Box\,\mathrm{Conj}(\Gamma_0) \equiv (\Box\alpha_1 \wedge \ldots \wedge \Box\alpha_n)$ (i.e., $\mathrm{Conj}(\Box\alpha_i)$) (i = 1,...,$n$) by **K1**. And then

(4) $\vdash\neg(\mathrm{Conj}(\Box\alpha_i) \wedge \neg\Box\alpha)$.

Since all formulas $\Box\alpha_i$, $\neg\Box\alpha$ are in Γ, it follows that Γ is $\mathcal{S}$-inc (contra hyp.).

Canonical models

To define a canonical model means to define the respective items: *W, R, V*:

W = the set of all maximal $\mathcal{S}$-con sets

R: $\Gamma R\Delta$ iff $\Gamma_\Box \subseteq \Delta$

(i.e., w_1Rw_2 iff for every formula α: if $\Box\alpha \in w_1$, then $\alpha \in w_2$)

The condition $\Gamma_\Box \subseteq \Delta$ in the definition of R is the minimal condition imposed on Δ to be a possible world relative to Γ (i.e., to be accessible from Γ).

V: $w \models \mathrm{p}$ iff $\mathrm{p} \in w$, i.e., a propositional variable p is true in a world w iff p is a member of w.

Now, the following lemmas are fundamental concerning canonical models.

Lemma 5. *Let* $M_{can} = \langle W, R, V\rangle$. *Then for every formula* α *and every* $w \in W$:

$w \models \alpha$ *iff* $\alpha \in w$.

Proof (induction on the complexity of α).

Basis. $\alpha = \mathrm{p}$; then lemma holds by def. of V for M_{can}.

Induction. $\alpha = \neg\beta$; and the lemma holds for β. We have:

$w \models \neg\beta$ iff $w \not\models \beta$ iff $\beta \notin w$ (by hyp.) iff $\neg\beta \in w$ (by max).

$\alpha = \beta \supset \gamma$; and the lemma holds for β and γ. Then $w \models \beta \supset \gamma$ iff $w \not\models \beta$ or $w \models \gamma$ iff $\beta \notin w$ or $\gamma \in w$ (by hyp.) iff $\beta \supset \gamma \in w$ (by Lemma 2(2) above).

$\alpha = \Box\beta$; and Lemma holds for β.

(a) If $\Box\beta \in w$, then for every $w_i \in W$ such that wRw_i: $\beta \in w_i$ (by def. of R for M_{can}). Hence $w_i \models \beta$ (by ind. hyp.), and therefore $w \models \Box\beta$.

(b) Conversely, we must prove: If $w \models \Box\beta$, then $\Box\beta \in w$; equivalently (by contraposition): If $\Box\beta \notin w$, then $w \not\models \Box\beta$. Let us suppose that $\Box\beta \notin w$. Then $\neg\Box\beta \in w$ (by max), and therefore the set $w_\Box \cup \{\neg\beta\}$ is $\mathcal{S}$-con (by Lemma 4, above). Hence there is a max $\mathcal{S}$-con set w_i such that wRw_i (since

$w_\Box \subseteq w_i$ and $\neg\beta \in w_i$). Whence $\beta \notin w_i$ (by con) and then $w_i \not\models \beta$. Finally, since wRw_i, it follows that $w \not\models \Box\beta$.

Lemma 6. *For any formula* α *the following holds:*

(*Can*) α *is valid in* M_{can} *for* $\mathcal{S}$ *iff* $\mathcal{S} \vdash \alpha$.

Proof. (a) If $\mathcal{S} \vdash \alpha$, then $\alpha \in \Gamma$ for every max $\mathcal{S}$-con Γ (by Lemma 2(3), above). And then $w \models \alpha$, for every $w \in W$ in M_{can} (by Lemma 5), i.e., α is valid in M_{can}.

(b) For the converse we argue for the corresponding conditional obtained by contraposition. If $\mathcal{S} \not\vdash \alpha$, then $\neg\alpha$ is $\mathcal{S}$-consistent. And then there is a maximal $\mathcal{S}$-con set $w \in W$ such that $\neg\alpha \in w$; whence $\alpha \notin w$ (by con). Therefore, $w \not\models \alpha$ (by Lemma 5 above) and then α is not valid in M_{can}.

4.1.3.2. Completeness (*via* canonical models)

Using (*Can*) of Lemma 6, the completeness[17] of the systems $\mathcal{K}$, $\mathcal{T}$, $\mathcal{K}4$, $\mathcal{S}4$, $\mathcal{S}5$ and $\mathcal{B}$ can easily be proved.

As we saw in 4.1.2.2, for these systems the following holds: *the class of frames for* $\mathcal{S}$ (where $\mathcal{S}$ is any system from this list) does coincide with *the class of all frames having the property* (*properties*) *P* required by the corresponding definition of $\mathcal{S}$-validity (reflexivity, symmetry, etc.). So in order to use (*Can*) for proving completeness of a modal system, all we have to prove is that *the canonical model for* $\mathcal{S}$ *is based on a frame for* $\mathcal{S}$, and then we reason as follows: since a formula α is $\mathcal{S}$-valid, then α is valid in every frame for $\mathcal{S}$ and therefore α is also valid in the frame for canonical model of $\mathcal{S}$. Whence α is valid in the canonical model M_{can}, and therefore, by (*Can*), α is a theorem of $\mathcal{S}$.

For the system $\mathcal{K}$ the completeness is an immediate result, since a formula α is $\mathcal{K}$-valid iff α is valid in every frame $\langle W, R\rangle$, and then α is valid in any model based on $\langle W, R\rangle$, and then valid in M_{can}. Hence, by (*Can*), α is provable in $\mathcal{K}$, i.e., $\mathcal{K} \vdash \alpha$.

Completeness Theorem for $\mathcal{T}$. *If* α *is* $\mathcal{T}$*-valid, then* $\mathcal{T} \vdash \alpha$.

[17] Remember that a system $\mathcal{S}$ is complete if all $\mathcal{S}$-valid formulas are theorems of $\mathcal{S}$.

Proof. According to the definition of $\mathcal{T}$-validity and the definition of R in M_{can}, all we have to prove is that R is *reflexive*, i.e., for any α and any $w \in W$: If $\Box\alpha \in w$, then $\alpha \in w$. Now, by axiom **T**: $\Box p \supset p$ of $\mathcal{T}$ and Subst, it follows that $\mathcal{T} \vdash \Box\alpha \supset \alpha$. And then $\Box\alpha \supset \alpha \in w$ (for any $w \in W$) (by 4.1.3.1, Lemma 2(3)). If $\Box\alpha \in w$, then $\alpha \in w$ (since otherwise $\neg\alpha \in w$; but $\{\Box\alpha, \Box\alpha \supset \alpha, \neg\alpha\}$ would be $\mathcal{S}$-inc).

Completeness Theorem for $\mathcal{K}4$. *If* α *is* $\mathcal{K}4$*-valid, then* $\mathcal{K}4 \vdash \alpha$.

Proof. We must prove that R of the canonical model is transitive; i.e., for any $w_1, w_2, w_3 \in W$: if $\Box\alpha \in w_1$, then $\alpha \in w_3$. Since $\mathcal{K}4 \vdash \Box p \supset \Box\Box p$, by Subst it follows that $\mathcal{K}4 \vdash \Box\alpha \supset \Box\Box\alpha$, and then $\Box\alpha \supset \Box\Box\alpha \in w$ (for any $w \in W$). Now, since $\Box\alpha \in w_1$ and $\Box\alpha \supset \Box\Box\alpha \in w_1$ it follows that $\Box\Box\alpha \in w_1$ (by Lemma 2(4)), and since $\Box\Box\alpha \in w_1$ and $w_1 R w_2$, it follows that $\Box\alpha \in w_2$ (by def. of R in M_{can}). And, finally, since $w_2 R w_3$, it follows that $\alpha \in w_3$ and this means that R is transitive.

Completeness Theorem for $\mathcal{S}4$. *If* α *is* $\mathcal{S}4$*-valid, then* $\mathcal{S}4 \vdash \alpha$.

Proof (follows from the proofs for $\mathcal{T}$ and $\mathcal{K}4$).

Completeness Theorem for $\mathcal{B}$. *If* α *is* $\mathcal{B}$*-valid, then* $\mathcal{B} \vdash \alpha$.

Proof. We must prove that R of M_{can} is reflexive and symmetrical. But reflexivity is required by the fact that $\mathcal{B}$ contains $\mathcal{T}$. So what remains to be proved is that R of M_{can} is symmetrical, i.e., if $w_1 R w_2$, then $w_2 R w_1$. In terms of M_{can} this means: if $\Box\alpha \in w_2$, then $\alpha \in w_1$, equivalently: if $\alpha \notin w_1$, then $\Box\alpha \notin w_2$.

Suppose that $\alpha \notin w_1$. Since $\mathcal{B} \vdash \neg\alpha \supset \Box\Diamond\neg\alpha$ (from axiom **B** and Subst: $\neg\alpha$ for p) it follows that $\neg\alpha \supset \Box\Diamond\neg\alpha \in w$, for any $w \in W$. Now, since $\alpha \notin w_1$, it follows that $\neg\alpha \in w_1$ (by max). And then $\Box\Diamond\neg\alpha \in w_1$ (by Lemma 2(4)). And since $w_1 R w_2$ (by def. of R in M_{can}) it follows that $\Diamond\neg\alpha \in w_2$, equivalently $\neg\Box\alpha \in w_2$, and then $\Box\alpha \notin w_2$ (by con).

Completeness Theorem for $\mathcal{S}5$. *If* α *is* $\mathcal{S}5$*-valid, then* $\mathcal{S}5 \vdash \alpha$.

The system $\mathcal{S}5$ properly extends $\mathcal{T}$, $\mathcal{S}4$ and $\mathcal{B}$, i.e., $\mathcal{S}5 \vdash$ **T**: $\Box p \supset p$ (immediate, by the construction of $\mathcal{S}5$), $\mathcal{S}5 \vdash$ **4**: $\Box p \supset \Box\Box p$ (comp. **S5**$_4$ of 4.1.1.). $\mathcal{S}5 \vdash$ **B**: $p \supset \Box\Diamond p$ (by the theorem $p \supset \Diamond p$ of $\mathcal{T}$ and **5**). So $\mathcal{S}5$-validity must contain all the three properties for validity (refl., symm. and trans.). Hence the completeness of $\mathcal{S}5$ follows from the respective proofs for $\mathcal{T}$, $\mathcal{K}4$ and $\mathcal{B}$.

4.1.3.3. $\mathcal{GL}$ is non-canonical

As we saw, by a *normal* modal system $\mathcal{S}$ we understand any extension (proper or not) of the modal system $\mathcal{K}$. When $\mathcal{S}$ is both sound and complete with respect to a class $\mathcal{C}$ of frames, then $\mathcal{S}$ is *characterized* by this class $\mathcal{C}$. And a frame $\langle W, R\rangle$ is a *frame for* $\mathcal{S}$ if every theorem of $\mathcal{S}$ is valid in $\langle W, R\rangle$. If there is a world $w \in W$ such that $w \not\models \alpha$, then we say that α is *falsified* in the frame $\langle W, R\rangle$.

By the property M_{can} of canonical model for a normal system $\mathcal{S}$ we conclude that the respective $\mathcal{S}$ is characterized by its canonical model. But this does not imply that any such system is characterized *by the frame* of its canonical model, since this *frame* may not be a frame for $\mathcal{S}$.

A system $\mathcal{S}$ is called *canonical* iff the frame $\langle W, R\rangle$ of its canonical model is a frame for $\mathcal{S}$.

The proof that $\mathcal{GL}$ is non-canonical runs as follows:[18]

1. To show that if for a modal system $\mathcal{S}$ the set $S = \{\neg\Box\alpha \mid \not\vdash \alpha\}$ is S-consistent, then the frame of the canonical model of $\mathcal{S}$ contains a world w such that wRw.

2. To show that if $\mathcal{GL}$ is such a system, then the axiom **W**: $\Box(\Box p \supset p) \supset \Box p$ is not valid in such a frame.

Argument for (2). Suppose that w^* is the world in M_{can} and therefore in the frame $\langle W, R\rangle$ of M_{can} such that w^*Rw^*. Suppose that $w \models p$ for all $w \in W$ with $w \neq w^*$ and $w^* \not\models p$. By the following simple argument it follows that

$w^* \not\models$ **W**: $\Box(\Box p \supset p) \supset \Box p$.

(1) $w^* \not\models \Box p$; since $w^* \not\models p$ and w^*Rw^*.

(2) $w^* \models \Box p \supset p$; (1) by PL

(3) $w \models \Box p \supset p$; since p is true in every world different from w^*.

(4) $w \models \Box p \supset p$; for every w (including w^*); by (2) and (3).

(5) $w^* \models \Box(\Box p \supset p)$; (4).

(6) $w^* \not\models \Box(\Box p \supset p) \supset \Box p$; (1), (5), PL.

Therefore, if the frame $\langle W, R\rangle$ of the canonical model for $\mathcal{GL}$ does contain a world w^* such that w^*Rw^*, then in such a frame the axiom **W** of $\mathcal{GL}$ is *falsifiable*. And this means that the frame $\langle W, R\rangle$ of the canonical

[18] This is only an outline. For the proof in all of its details, comp. G.E. Hughes and M.J.Creswell [1996], 138-141.

model of $\mathcal{GL}$ is not a frame for $\mathcal{GL}$. Hence $\mathcal{GL}$ is a *non-canonical system*.

It also follows the following fact: If **W** is valid in a frame $\langle W,R\rangle$, then $\langle W,R\rangle$ is *irreflexive*. The validity of **W** in a frame $\langle W,R\rangle$ also requires the *transitivity* of $\langle W,R\rangle$ and that $\langle W,R\rangle$ does not contain an infinite sequence *Seq* of the form $w_1Rw_2R\ldots$ (i.e., a sequence in which every term has a successor). Let us consider these aspects.

Theorem 1. *If* **W** *is valid in* $\langle W,R\rangle$, *then* $\langle W,R\rangle$ *is transitive.*

Equivalently, by contraposition: *If* $\langle W,R\rangle$ *is not transitive, then* **W** *is not valid in* $\langle W,R\rangle$.

Proof.[19] Assume hypothesis, i.e., for w_1, w_2 and w_3 from *W*: w_1Rw_2 and w_2Rw_3, but not w_1Rw_3. Define $\langle W,R,V\rangle$ a model based on $\langle W,R\rangle$, where *V* is defined as follows: (Eq) $w\models$ p iff $w\neq w_2$ and $w\neq w_3$. Then we have:

(1) $w_1\not\models\Box$p; since w_1Rw_2 and $w_2\not\models$ p.

(2) Let $w_i\in W$ be any world such that w_1Rw_i. w_i cannot be w_3 since by definition not w_1Rw_3. So will remain the following two cases:

(a) $w_i = w_2$. Since w_2Rw_3 and $w_3\not\models$ p it follows that $w_2\not\models\Box$p and therefore $w_2\models\Box$p$\supset$p.

(b) In any other world $w_i\in W$ (different from w_2 and w_3) we have $w_i\models\Box$p$\supset$p (since in such worlds $w_i\models$ p).

(3) $w_1\models\Box(\Box$p$\supset$p); from (a) and (b) of (2).

(4) $w_1\not\models\Box(\Box$p$\supset$p)$\supset\Box$p; (1), (3), PL.

Theorem 2. **W**: $\Box(\Box$p$\supset$p)$\supset\Box$p *is valid in* $\langle W,R\rangle$ *iff* $\langle W,R\rangle$ *does not contain an infinite Seq.*

(a) If **W**: $\Box(\Box$p$\supset$p)$\supset\Box$p is valid in $\langle W,R\rangle$, then $\langle W,R\rangle$ does not contain an infinite *Seq*.

(b) If $\langle W,R\rangle$ does not contain an infinite *Seq*, then **W**: $\Box(\Box$p$\supset$p)$\supset\Box$p is valid in $\langle W,R\rangle$.

Proof. (a) (*Reductio*). Assume hypothesis of (a) and that $\langle W,R\rangle$ does contain an infinite sequence *Seq*: $w_1Rw_2R\ldots$. Let $\langle W,R,V\rangle$ be a model

[19] Cf. G.E. Hughes and M.J. Cresswell [1996], 178. For more interesting things about the non-canonical system $\mathcal{GL}$, comp. G. Boolos [1993], Ch. 7 and C. Smorynski [1985], Ch. 2, Sect. 2.

based on $\langle W, R\rangle$, where V is defined as follows: (Eq) $w \models$ p iff $w \notin Seq$. So we have the following derivations:

(1) $w_i \not\models \Box$p, for every $w_i \in Seq$, since for any w_i there is w_{i+1} such that $w_i R w_{i+1}$ (by hyp.) and $w_{i+1} \not\models$ p (by def. of V).

(2) Let w be any world such that $w_i R w$. We have the following two cases:

(a) $w \notin Seq$; then $w \models$ p (by def. of V), and then $w_i \models \Box$p$\supset$p (by PL).

(b) $w \in Seq$; then $w \not\models$ p and therefore $w_i \not\models \Box$p. It follows that $w_i \models \Box$p$\supset$p (by PL).

(3) $w_i \models \Box(\Box$p$\supset$p); from (a) and (b).

(4) $w_i \not\models \Box(\Box$p$\supset$p)$\supset\Box$p; (1) and (3).

This means that there is a model $\langle W, R, V\rangle$ based on the frame $\langle W, R\rangle$ in which **W** is not valid. Whence, **W** is not valid in $\langle W, R\rangle$; contra hypothesis.

Proof. (b) (*Reductio*). Assume hypothesis of (b) and that **W**: $\Box(\Box$p$\supset$p)$\supset\Box$p is not valid in $\langle W, R\rangle$. This means that there is a world $w_1 \in W$ in which **W** is false. So, we have the following derivations:

(1) $w_1 \not\models \Box(\Box$p$\supset$p)$\supset\Box$p; by hyp.

(2) $w_1 \models \Box(\Box$p$\supset$p); (1), PL

(3) $w_1 \not\models \Box$p; (1), PL

(4) There is a world $w_2 \in W$ such that $w_1 R w_2$ and $w_2 \not\models$ p; from (3).

(5) $w_2 \models \Box$p$\supset$p; from (2) and that $w_1 R w_2$.

(6) $w_2 \not\models \Box$p, from (5) and (4) by PL

(7) There is a world $w_3 \in W$ such that $w_2 R w_3$ and $w_3 \not\models$ p; from (6).

(8) $w_3 \models \Box$p$\supset$p; from (2) and that $w_1 R w_2$ and $w_2 R w_3$ plus transitivity.

(9) $w_3 \not\models \Box$p; (7) and (8) by PL

(10) There is a world $w_4 \in W$ such that $w_3 R w_4$ and $w_4 \not\models$ p; (9), and so on.

And then we got an infinite *Seq* $w_1 R w_2 R ...$ in $\langle W, R\rangle$, contradicting the hypothesis of (b).

4.1.3.4. Finite models

As we saw in the preceding section the canonical models, by property *Can*, are a useful tool for proving the completeness theorems for modal systems. However, this technique cannot be applied to the system $\mathcal{GL}$, since $\mathcal{GL}$ is non-canonical, viz. the frame of its canonical model is not a frame for $\mathcal{GL}$. Nevertheless, the completeness of $\mathcal{GL}$ (and of the other systems considered here) can be proved using *finite models*.

The finite models are "weaker" forms of the canonical models, in the following sense: in order to show that an $\mathcal{S}$-valid formula α is a theorem of $\mathcal{S}$ we do not consider the infinite list of all modal formulas (as in the construction of canonical models) but *only* the formula α; more exactly, only the *finite* set of all subformulas of α, since the truth-value of α in a world $w \in W$ depends only on the truth-values of its subformulas in that world.

As we know, to prove the completeness of a system $\mathcal{S}$ means to prove the following conditional:

Compl. If α is $\mathcal{S}$-valid, then $\mathcal{S} \vdash \alpha$, equivalently (by contraposition):

Compl. If $\mathcal{S} \nvdash \alpha$, then α is not $\mathcal{S}$-valid.

And in order to show that α is not $\mathcal{S}$-valid is enough to construct a finite model $M^* = \langle W^*, R^*, V^* \rangle$, based on a frame $\langle W^*, R^* \rangle$ having the properties required by $\mathcal{S}$-validity (refl., trans., etc.), in which α is falsifiable, i.e., there is a $w \in W$ such that $w \nvDash \alpha$.

Let us detail.[20]

By a subformula of a formula α we understand any part of α, including α itself, which satisfies the clauses of the definition of formula of L_{PML} (comp. 4.1.1).

Example. Let α = □(¬p∧¬□(q∨r)). Then the subformulas of α are: p, ¬p, ¬□(q∨r), □(q∨r), q∨r, q, r, ¬p∧¬□(q∨r), □(¬p∧¬□(q∨r)).

The construction of the finite model proceeds as follows. We begin by considering the set of all subformulas of a formula α, i.e., $Set_\alpha = \{\beta \mid \beta$ is a subformula of $\alpha\}$. Now, let $Set^*_\alpha = \mathrm{Set}_\alpha \cup \{\neg\beta \mid \beta \in \mathrm{Set}_\alpha\}$. Therefore, Set^*_α is a *finite* set containing all subformulas of α and their negations.

[20] The considerations of Sect. 4.1.3.4 are partially based on Hughes and Gresswell [1996], Ch. 8.

To construct the finite model $M^* = \langle W^*, R^*, V^* \rangle$ means to define the corresponding items.[21]

W^* is the set of all maximal $\mathcal{S}$-consistent sets of formulas constructed from the members of Set^*_α (they will be called α-max-$\mathcal{S}$-con sets[22]). For such α-max-$\mathcal{S}$-con set the lemmas 1 and 2 of 4.1.3.1 evidently holds (for any formulas α, β of Set^*_α). Lemma 3 (4.1.3.1) also holds, and let us prove it for the form required here.

Lemma 3*. *Let $\Delta \subseteq Set^*_\alpha$ be any $\mathcal{S}$-con set. Then there exist an α-max-$\mathcal{S}$-con set Γ such that $\Delta \subseteq \Gamma$.*

Proof.[23] Let $\gamma_1,...,\gamma_n$ be an enumeration of the formulas of Set^*_α. We form Γ as follows.

$\Gamma_0 = \Delta$, and for $0 \leq k < n$

$$\Gamma_{k+1} = \begin{cases} \Gamma_k \cup \{\gamma_{k+1}\} \text{ if this is consistent} \\ \Gamma_k \cup \{\neg\gamma_{k+1}\}; \text{ otherwise.} \end{cases}$$

So defined, Γ is $\mathcal{S}$-con, since $\Gamma_0 = \Delta$ is $\mathcal{S}$-con (by hyp.) and for any *k* if Γ_k is $\mathcal{S}$-con, then either $\Gamma_k \cup \{\gamma_{k+1}\}$ is $\mathcal{S}$-con, and then $\Gamma_{k+1} = \Gamma_k \cup \{\gamma_{k+1}\}$ (by Def.), or $\Gamma_k \cup \{\gamma_{k+1}\}$ is $\mathcal{S}$-inc and then $\Gamma_{k+1} = \Gamma_k \cup \{\neg\gamma_{k+1}\}$ (by Lemma 1, 4.1.3.1). So $\Gamma_n = \Gamma$ is $\mathcal{S}$-con and max.

In brief, the finite model $M^* = \langle W^*, R^*, V^* \rangle$ is defined as follows:

W^*: the set of all α-max-$\mathcal{S}$-con sets

R^* (depends on $\mathcal{S}$; for $\mathcal{K}$ and $\mathcal{T}$ it is that given for canonical models)

V^*: $w \models$ p iff p$\in w$, if $\text{p} \in Set_\alpha$; and arbitrary if $\text{p} \notin Set_\alpha$.

Lemma 4 of 4.1.3.1 also holds. Now, since the definition of R^* (accessibility relation) for the finite model $M^* = \langle W^*, R^*, V^* \rangle$ for the systems $\mathcal{K}$ and $\mathcal{T}$ is that given for the canonical models (comp. 4.1.3.1), Lemma 5 also holds (as we shall see), but only for $\mathcal{K}$ and $\mathcal{T}$. But since the definition of R^* for the other systems will be different from that given for

[21] It is similar to the construction of the canonical model, but relativized to Set^*_α.

[22] In the sense of Def. 1 and Def. 2 of 4.1.3.1.

[23] Similar to the proof of Lemma 3 Sect. 4.1.3.1 (but this time relativized to Set^*_α).

canonical models, the proof of the clause regarding the *modal*[24] formulas β (of the form $\Box\gamma$) must be given for any specific system.

Let us detail. First of all, the proof of Lemma 5 (we call it Lemma 5*) for the finite models.

Lemma 5*. *For any* $\beta \in Set_\alpha$ *and any* $w \in W^*$: $w \models \beta$ *iff* $\beta \in w$.

Proof (induction on the complexity of β).[25]

Basis. $\beta = p$. The lemma holds by def. of V^*.

Induction. $\beta = \neg\gamma$. Then we have the following derivations: $w \models \beta$ iff $w \models \neg\gamma$ iff $\not\models \gamma$ (by 4.1.2.1, Def 1(2)) iff $\gamma \notin w$ (by ind. hyp.) iff $\neg\gamma \in w$ (by α-max) iff $\beta \in w$.

$\beta = \gamma \supset \delta$. Then the following holds: $w \models \beta$ if $w \models \gamma \supset \delta$ iff $w \not\models \gamma$ or $w \models \delta$ (by 4.1.2.1 Def 1(3)) iff $\gamma \notin w$ or $\delta \in w$ (by ind. hyp.) iff $\gamma \supset \delta \in w$ (by 4.1.3.1, Lemma 2(2)) iff $\beta \in w$.

$\beta = \Box\gamma$. What must be proved are the following conditionals:

(a) If $\beta \in w$, then $w \models \beta$, and

(b) If $w \models \beta$, then $\beta \in w$.

(a) Suppose $\beta \in w$, i.e., $\Box\gamma \in w$. Then $\gamma \in w'$ for every $w' \in W^*$ such that wRw' (by Def. of R^*). Whence $w' \models \gamma$ and therefore $w \models \Box\gamma$, i.e., $w \models \beta$.

(b) is equivalent (by contraposition) to: if $\beta \notin w$, then $w \not\models \beta$. Suppose therefore that $\beta \notin w$, i.e., $\Box\gamma \notin w$. Then $\neg\Box\gamma \in w$ (by max), and therefore the set $w_\Box \cup \{\neg\gamma\}$ is $\mathcal{S}$-consistent (by Lemma 4 (Sect. 4.1.3.1), where $w_\Box = \{\delta \mid \Box\delta \in w\}$). Since all formulas $\Box\delta$ are formulas of Set_α (by hyp.), all formulas $\delta \in Set_\alpha$. So in the set $w_\Box \cup \{\neg\gamma\}$ only the formulas $\neg\gamma$ may be or not in Set_α. But anyway $\neg\gamma \in Set_\alpha^*$. Now, since $w_\Box \cup \{\neg\gamma\}$ is $\mathcal{S}$-con, it follows, by Lemma 3*, that there is an α-max-$\mathcal{S}$-con set *w'* such that $w_\Box \cup \{\neg\gamma\} \subseteq w'$. And since for all δ, if $\Box\delta \in w$, then $\delta \in w'$, it follows that wR^*w' and $\neg\gamma \in w'$. Then $\gamma \notin w'$ (by $\mathcal{S}$-con). And since $\Box\gamma \in Set_\alpha$, it

[24] In the proof by induction, given above for Lemma 5, the *Basis* and *Induction* for formulas $\neg\gamma$ and $\gamma\supset\delta$ remain unchanged.

[25] As can be expected, the proof of this lemma mimics the proof of Lemma 5 from Sect. 4.1.3.1.

follows that $\gamma \in Set_\alpha$. And since $\gamma \notin w'$ and Lemma 5* holds for γ, it follows that $w' \not\models \gamma$ (by ind. hyp.); and since wRw' it follows that $w \not\models \Box\gamma$, i.e., $w \not\models \beta$.

Now, to sum up, this apparatus allows us to prove the completeness of modal systems, in the following way. We must prove the completeness in its equivalent form (by contraposition). Let *Compl** be the *argument* used in proving completeness theorem for $\mathcal{S}$.

Compl*. If $\mathcal{S} \nvdash \alpha$, then α is not $\mathcal{S}$-valid.

Suppose that $\mathcal{S} \nvdash \alpha$, then $\{\neg\alpha\}$ is $\mathcal{S}$-consistent. So, by Lemma 3*, there is an α-max-$\mathcal{S}$-con set $w \in W^*$ such that $\neg\alpha \in w$. It follows that $\alpha \notin w$. Whence, by Lemma 5*, $w \not\models \alpha$. And then α is not valid in the finite model $M^* = \langle W^*, R^*, V^* \rangle$ and therefore α is not valid in the finite frame $\langle W^*, R^* \rangle$; i.e., α is not $\mathcal{S}$-valid.

As can be observed, if we also consider the soundness of $\mathcal{S}$, then these results give the co-extensivity of the following items: "α is $\mathcal{S}$-valid", "α is valid in all finite $\mathcal{S}$-frames" and "α is provable in $\mathcal{S}$".

4.1.3.5. Completeness of modal systems (*via* finite models)

As we mentioned above, unlike the canonical models, in the case of *finite* models the relation R^* is specific to each system $\mathcal{S}$. And this does imply that Lemma 5* must be proved every time for modal formulas (in the inductive step). So in using *Compl** for proving the completeness of $\mathcal{K}$-$\mathcal{GL}$ only the following two things must be shown:

(1*) R^* has the property required by the respective definition of $\mathcal{S}$-validity (i.e., $\langle W, R \rangle$ is a frame for $\mathcal{S}$).

(2*) Lemma 5* (for modal formulas) holds for $\mathcal{S}$.[26]

Completeness of $\mathcal{K}$

The completeness of $\mathcal{K}$ easy follows from Compl*, since if $\mathcal{K} \nvdash \alpha$, then α is not valid in $M^* = \langle W^*, R^*, V^* \rangle$, and therefore α is not valid in $\langle W^*, R^* \rangle$. So, there is a frame $\langle W, R \rangle$ in which α is not valid, and then α is

[26] With some changes, this strategy of proving the completeness of modal systems is that from G. Boolos [1993], Ch. 5.

not $\mathcal{K}$-valid (according to the definition of $\mathcal{K}$-validity).

For all the other systems we must show *in every case*, that the finite frame $\langle W^*, R^*\rangle$ has the properties required by the respective definition of $\mathcal{S}$-validity. Briefly, what must be shown is that R^* has those properties.

Completeness of $\mathcal{T}$

As for $\mathcal{K}$, the definition of R^* for $\mathcal{T}$ is that of the canonical models, i.e., wR^*w' iff $w_\Box \subseteq w'$ (comp. 4.1.3.1). The Lemma 5^* is that proven above, and so (2^*) holds. So the proof of its completeness is reducible to the proof that R^* is *reflexive*; viz. to show that for any formula $\Box\beta$ and any $w \in W^*$: if $\Box\beta \in w$, then $\beta \in w$. But this is clear, since otherwise $\Box\beta \in w$ and $\beta \notin w$, and then $\neg\beta \in w$. But in this case w would be $\mathcal{S}$-inc, since $\mathcal{S} \vdash \Box\beta \supset \beta$ (using the axiom **T**), viz. $\mathcal{S} \vdash \neg(\Box\beta \wedge \neg\beta)$.

Observe that if α has no modal operators, then the Set_α has no formula of the form $\Box\beta$, so the set $w_\Box$ would be empty and automatically $w_\Box \subseteq w$. Hence (1^*) also holds.

Completeness of $\mathcal{K}4$

As we know, $\mathcal{K}4 = \mathcal{K}$+**4**: $\Box p \supset \Box\Box p$. This time the definition of R^* must be changed, since the worlds $w \in W^*$ do contain only formulas of Set_α^* (i.e., subformulas of α or their negations), and if $\Box\beta \in w$, this does not imply, by the axiom **4**, that $\Box\Box\beta \in w$, since α may not contain $\Box\Box\beta$ as a subformula. And then to show that R^* is transitive requires the change of definition of R^* (for finite models). And this is the following:

wR^*w' iff for all $\Box\beta$: if $\Box\beta \in w$, then $\Box\beta$, $\beta \in w'$.

So defined, R^* is evidently transitive, and then (1^*) holds.

Now it must be shown that (2^*) also holds, viz. for any $\Box\gamma \in Sat_\alpha$ and any $w \in W^*$:

$w \models \Box\gamma$ iff $\Box\gamma \in w$.

Equivalently, we must prove the following two conditionals:

(a) If $\Box\beta \in w$, then $w \models \Box\gamma$.

(b) If $w \models \Box\gamma$, then $\Box\gamma \in w$.

(a) Suppose $\Box\gamma \in w$ and wRw'. Then $\Box\gamma,\ \gamma \in w'$ (by def. of R^* for $\mathcal{K}4$).

Then $w' \models \gamma$ (by ind. hyp.); whence $w \models \Box\gamma$.

(b) This conditional is equivalent (by contraposition) to

If $\Box\gamma \notin w$, then $w \not\models \Box\gamma$.

Suppose that $\Box\gamma \notin w$, but $\Box\gamma \in Set_\alpha$. Then $\neg\Box\gamma \in w$ (since w is α-max-$\mathcal{K}4$-con). To derive (b) it must be shown that the following set is $\mathcal{K}4$-consistent:

$\Delta = \{\delta|\Box\delta \in w\} \cup \{\Box\delta|\Box\delta \in w\} \cup \{\neg\gamma\}$

Δ is $\mathcal{K}4$-consistent.

Reductio. Suppose that Δ is $\mathcal{K}4$-inconsistent, i.e.,

(1) $\mathcal{K}4 \vdash \neg(\delta_1 \wedge \ldots \wedge \delta_n \wedge \Box\delta_1 \wedge \ldots \wedge \Box\delta_n \wedge \neg\gamma)$, where $\Box\delta_i$ $(1 \leq i \leq n)$ are all formulas of this form in w.

Let $D = \delta_1 \wedge \ldots \wedge \delta_n$ and since $\Box\delta_1 \wedge \ldots \wedge \Box\delta_n \equiv \Box(\delta_1 \wedge \ldots \wedge \delta_n)$ (by 4.1.1, $\mathbf{K_1}$) let $\Box D = \Box(\delta_1 \wedge \ldots \wedge \delta_n)$. Hence

$\mathcal{K}4 \vdash (D \wedge \Box D) \supset \gamma$; PL

(2) $\mathcal{K}4 \vdash \Box(D \wedge \Box D) \supset \Box\gamma$; (1) Deriv. 1

(3) $\mathcal{K}4 \vdash (\Box D \wedge \Box\Box D) \supset \Box\gamma$; (3), 4.1.1, $\mathbf{K_1}$

(4) $\mathcal{K}4 \vdash \Box D \supset \Box\Box D$; **4**, Subst.

(5) $\mathcal{K}4 \vdash \Box D \supset \Box\gamma$; (3), (4), PL

($\vdash ((p \wedge q) \supset r) \supset ((p \supset q) \supset (p \supset r))$, Subst, MP)

In extenso,

(6) $\mathcal{K}4 \vdash \Box(\delta_1 \wedge \ldots \wedge \delta_n) \supset \Box\gamma$, and then

(7) $\mathrm{K4} \vdash (\Box\delta_1 \wedge \ldots \wedge \Box\delta_n) \supset \Box\gamma$; (6), 4.1.1, $\mathbf{K_1}$

(8) $\mathrm{K4} \vdash \neg(\Box\delta_1 \wedge \ldots \wedge \Box\delta_n \wedge \neg\Box\gamma)$; (7), PL.

Now, since all formulas $\Box\delta_1, \ldots, \Box\delta_n$, $\neg\Box\gamma$ are in w, it follow that w is $\mathcal{K}4$-inc. Then if w is $\mathcal{K}4$-con, Δ is $\mathcal{K}4$-con too. And then, by Lemma 3^*, there is an α-max-$\mathcal{S}$-con set w' such that (*) wR^*w', since if $\Box\delta \in w$, then $\Box\delta,\ \delta \in w'$ (by construction of Δ and the fact that $\Delta \subseteq w'$), and (**) $\neg\gamma \in w'$, and then $\gamma \notin w'$ (by $\mathcal{K}4$-con). And since Lemma 5^* holds for γ (by ind. hyp.), it follows that $w' \not\models \gamma$, and by (*) $w \not\models \Box\gamma$. Therefore, (2^*) also holds.

Using $Compl^*$ the completeness of $\mathcal{K}4$ follows.

Completeness of $\mathcal{S}4$

Since $\mathcal{S}4 = \mathcal{K}$+**T**: $\Box p \supset p$+**4**: $\Box p \supset \Box\Box p$ and by keeping the definition of R^* given above for $\mathcal{K}4$ (R^* being evidently also reflexive), it follows, by the above proofs for $\mathcal{T}$ and $\mathcal{K}4$, that $\mathcal{S}4$ is complete.

Completeness of $\mathcal{B}$

As we know, $\mathcal{B} = \mathcal{K}$+**T**: $\Box p \supset p$+**B**: $p \supset \Box\Diamond p$.

R^* is defined in the following way:

wR^*w' iff for all $\Box\beta$: If $\Box\beta \in w$, then $\beta \in w'$ and if $\Box\beta \in w'$, then $\beta \in w$.

So defined, the symmetry of R^* is evident and since $\mathcal{B}$ contains **T**: $\Box p \supset p$, also the reflexivity holds. And then (1^*) holds.

What remains to be proved is the Lemma 5^* for modal formulas, i.e., for any $\Box\gamma \in Sat_\alpha$ and any $w \in W^*$:

$w \models \Box\gamma$ iff $\Box\gamma \in w$,

or equivalently the two conditionals.

(a) If $\Box\gamma \in w$, then $w \models \Box\gamma$.

(b) If $w \models \Box\gamma$, then $\Box\gamma \in w$.

(a) Suppose $\Box\gamma \in w$ and wR^*w'. Then $\gamma \in w'$ (by def. of R^* for $\mathcal{B}$). Then $w' \models \gamma$ (by ind. hyp.). Since w' is arbitrary and wRw', $w \models \Box\gamma$.

(b) As above, this conditional is equivalent (by contraposition) to:

If $\Box\gamma \notin w$, then $w \not\models \Box\gamma$.

Suppose that $\Box\gamma \notin w$, but $\Box\gamma \in Set_\alpha$. Then $\neg\Box\gamma \in w$ (by max). To derive (b) it must be shown that the following set is $\mathcal{B}$-consistent.

$\Delta = \{\delta | \Box\delta \in w\} \cup \{\neg\Box\varepsilon | \Box\varepsilon \in Sat_\alpha$ and $\neg\varepsilon \in w\} \cup \{\neg\gamma\}$.

Δ is $\mathcal{B}$-consistent.

Reductio. Suppose that Δ is $\mathcal{B}$-inconsistent, i.e.,

(1) $\mathcal{B} \vdash \neg(\delta_1 \wedge \ldots \wedge \delta_n \wedge \neg\Box\varepsilon_1 \wedge \ldots \wedge \neg\Box\varepsilon_m \wedge \neg\gamma)$, where $\Box\delta_i$ are all formulas of the form $\Box\delta$ in w and $\neg\Box\varepsilon_i$ are all formulas of the form $\neg\Box\varepsilon$ such that $\Box\varepsilon_i \in Sat_\alpha$ and $\neg\varepsilon \in w$.

(2) $\mathcal{B} \vdash (\delta_1 \wedge \ldots \wedge \delta_n \wedge \neg\Box\varepsilon_1 \wedge \ldots \wedge \neg\Box\varepsilon_m) \supset \gamma$; (1), PL

(3) $\mathcal{B} \vdash \Box(\delta_1 \wedge \ldots \wedge \delta_n \wedge \neg\Box\varepsilon_1 \wedge \ldots \wedge \neg\Box\varepsilon_m) \supset \Box\gamma$; (2), Deriv. 1

(4) $\mathcal{B} \vdash (\Box\delta_1 \wedge \ldots \wedge \Box\delta_n \wedge \Box\neg\Box\varepsilon_1 \wedge \ldots \wedge \Box\neg\Box\varepsilon_m) \supset \Box\gamma$; (3), $\mathbf{K_1}$

Let $p = \Box\delta_1 \wedge \ldots \wedge \Box\delta_n$, $q = \Box\neg\Box\varepsilon_1 \wedge \ldots \wedge \Box\neg\Box\varepsilon_m$, $r = \Box\gamma$.

(5) $\mathcal{B} \vdash \neg\varepsilon_i \supset \Box\Diamond\neg\varepsilon_i$ $(1 \le i \le m)$, by **B**, equivalently

$\mathcal{B} \vdash \neg\varepsilon_i \supset \Box\neg\Box\varepsilon_i$, and then

(6) $\mathcal{B} \vdash (\neg\varepsilon_1\wedge\ldots\wedge\neg\varepsilon_m)\supset(\Box\neg\Box\varepsilon_1\wedge\ldots\wedge\Box\neg\Box\varepsilon_m)$, (5), PL

Let $s = \neg\varepsilon_1 \wedge \ldots \wedge \neg\varepsilon_m$.

(7) $(\Box\delta_1\wedge\ldots\wedge\Box\delta_n\wedge\neg\varepsilon_1\wedge\ldots\wedge\varepsilon_m)\supset\Box\gamma$; (4), (6) by $\vdash((p\wedge q)\supset r)\supset[(s\supset q)\supset((p\wedge s)\supset r)]$, Subst. and MP (two times)

(8) $\mathcal{B} \vdash \neg(\Box\delta_1\wedge\ldots\wedge\Box\delta_n\wedge\neg\varepsilon_1\wedge\ldots\wedge\neg\varepsilon_m\wedge\neg\Box\gamma)$; (7), PL

Since all formulas $\Box\delta_i$, $\neg\varepsilon_i$ and $\neg\Box\gamma$ are in w, it follows, by (8), that w is $\mathcal{B}$-incon.

Now, since Δ is $\mathcal{B}$-con, by Lemma 3* there is a w' such that (*) wRw' and (**) $\neg\gamma \in w'$. (a) holds, since if $\Box\delta \in w$, then $\delta \in w'$. And if $\Box\varepsilon \in w'$, then $\varepsilon \in w$, since if $\varepsilon \notin w$, then $\neg\varepsilon \in w$ and then by the definition of Δ, $\neg\Box\varepsilon \in \Delta$, and then $\neg\Box\varepsilon \in w'$, making w' $\mathcal{B}$-inc. Now, from (*) and (**) it follows that $\gamma \notin w'$, and then $w' \not\models \gamma$ (by ind. hyp.), and since wRw', it follows that $w \not\models \Box\gamma$. So (2*) also holds.

Using *Compl**, we obtain the completeness of $\mathcal{B}$.

Completeness of $\mathcal{S}5$

$\mathcal{S}5 = \mathcal{K}$+**T**: $\Box p\supset p$+**5**: $\Diamond p\supset\Box\Diamond p$.

R^* is defined in the following way:

wRw' iff for all $\Box\beta$: $\Box\beta \in w$ iff $\Box\beta \in w'$.

As can be easily verified, all properties required by the definition of $\mathcal{S}5$-validity hold: reflexivity, symmetry and transitivity. So (1*) holds.

To prove (2*) means to show that for any $\Box\gamma$ and any $w \in W^*$:

$w \models \Box\gamma$ iff $\Box\gamma \in w$,

or, equivalently, the two conditionals:

(a) If $\Box\gamma \in w$, then $w \models \Box\gamma$, and

(b) the converse of (a).

(a) Suppose that $\Box\gamma \in w$ and wRw'. Then by def. of R^*, $\Box\gamma \in w'$. But by the axiom **T**: $\Box p\supset p$ it follows that $\gamma \in w'$, and then $w' \models \gamma$ (by ind. hyp.). Since w' is arbitrary and wRw': $w \models \Box\gamma$.

(b) is equivalent to: If $\Box\gamma \notin w$, then $w \not\models \Box\gamma$. Suppose that $\Box\gamma \notin w$ and that

$\Box\gamma\in Set_\alpha$. Then $\neg\Box\gamma\in w$ (by max). As in the preceding proofs, (b) can be derived from the argument showing that the following set is $\mathcal{S}5$-consistent:

$\Delta=\{\Box\delta|\Box\delta\in w\}\cup\{\neg\Box\varepsilon|\neg\Box\varepsilon\in w\}\cup\{\neg\gamma\}$.

Reductio. Suppose Δ is $\mathcal{S}5$-inconsistent, i.e.,

(1) $\mathcal{S}5\vdash\neg(\Box\delta_1\wedge...\wedge\Box\delta_n\wedge\neg\Box\varepsilon_1\wedge...\wedge\neg\Box\varepsilon_m\wedge\neg\gamma)$, where $\Box\delta_i$ are all formulas of the form $\Box\delta$ in w and $\neg\Box\varepsilon_i$ are all formulas of the form $\neg\Box\varepsilon$ in w, and derive

(2) $\mathcal{S}5\vdash\neg(\Box\delta_1\wedge...\wedge\Box\delta_n\wedge\neg\Box\varepsilon_1\wedge...\wedge\neg\Box\varepsilon_m\wedge\neg\Box\gamma)$.

Then since all formulas $\Box\delta_i$, $\neg\Box\varepsilon_i$ and $\neg\Box\gamma$ are in w, by (2) it follows that w is $\mathcal{S}5$-inconsistent. Hence Δ is $\mathcal{S}5$-con. Now we reason as above (in the proofs for $\mathcal{K}4$ or $\mathcal{B}$) and conclude that $w\not\models\Box\gamma$. So (2^*) also holds. Whence, finally, by *Compl** the completeness of $\mathcal{S}5$ follows (write down the full proof!).

Completeness of $\mathcal{GL}$

$\mathcal{GL}=\mathcal{K}+\mathbf{W}$: $\Box(\Box p\supset p)\supset\Box p$.

R^* is defined as follows:

wR^*w' iff (a) for all $\Box\beta\in w$: $\Box\beta$, $\beta\in w'$, and

(b) there is a formula $\Box\varepsilon\in w'$: $\neg\Box\varepsilon\in w$.

R^* is *transitive*. Suppose wR^*w' and $w'R^*w''$. So, if $\Box\delta\in w$, then $\Box\delta\in w'$, and therefore $\Box\delta$, $\delta\in w''$ (by (a)). Hence wRw'' holds. And since wRw' there is a formula $\Box\varepsilon\in w'$ and $\neg\Box\varepsilon\in w$ (by (b)). And since $w'Rw''$, $\Box\varepsilon\in w''$ (by (a)). Hence there is a formula $\Box\varepsilon$ such that $\Box\varepsilon\in w''$ and $\neg\Box\varepsilon\in w$, and therefore wRw'' holds.

R^* is *irreflexive*. If R^* were reflexive, i.e., wR^*w, then there would be a formula $\Box\varepsilon\in w$ and $\neg\Box\varepsilon\in w$, contrary to the $\mathcal{GL}$-consistency of w. Since $\langle W^*,R^*\rangle$ is also finite (comp. 4.1.3.3, *Theorem* 2), it follows that $\langle W^*,R^*\rangle$ is a frame for $\mathcal{GL}$; and therefore (1^*) holds. It must be shown that (2^*) also holds, viz. for any formula $\Box\gamma\in Set_\alpha$ and any $w\in W^*$:

$w\models\Box\gamma$ iff $\Box\gamma\in w$.

Equivalently, the following conditionals must be proved:

(a) If $\Box\gamma\in w$, then $w\models\Box\gamma$.

(b) If $w\models\Box\gamma$, then $\Box\gamma\in w$.

(a) Suppose $\Box\gamma \in w$ and wRw'. Then $\gamma \in w'$ and then $\gamma \in Set_\alpha$. Therefore $w' \models \gamma$ (by ind. hyp.) and then $w \models \Box\gamma$.

(b) This conditional is equivalent to

If $\Box\gamma \notin w$, then $w \not\models \Box\gamma$.

Suppose therefore that $\Box\gamma \notin w$ but $\Box\gamma \in Set_\alpha$. Then $\neg\Box\gamma \in w$. To derive (b) is enough to show that the following set is $\mathcal{GL}$-consistent:

$\Delta = \{\delta | \Box\delta \in w\} \cup \{\Box\delta | \Box\delta \in w\} \cup \{\Box\gamma, \neg\gamma\}$

Δ *is* $\mathcal{GL}$*-consistent*.

Reductio.

(1) $\mathcal{GL} \vdash \neg(\delta_1 \wedge ... \wedge \delta_n \wedge \Box\delta_1 \wedge ... \wedge \Box\delta_n \wedge \Box\gamma \wedge \neg\gamma)$; hyp.

(2) $\mathcal{GL} \vdash \neg(\delta_1 \wedge ... \wedge \delta_n \wedge \Box\delta_1 \wedge ... \wedge \Box\delta_n) \vee \neg(\Box\gamma \wedge \neg\gamma)$; (1), PL

(3) $\mathcal{GL} \vdash (\delta_1 \wedge ... \wedge \delta_n \wedge \Box\delta_1 \wedge ... \wedge \Box\delta_n) \supset (\Box\gamma \supset \gamma)$; (2), PL

(4) $\mathcal{GL} \vdash (\Box\delta_1 \wedge ... \wedge \Box\delta_n \wedge \Box\Box\delta_1 \wedge ... \wedge \Box\Box\delta_n) \supset \Box(\Box\gamma \supset \gamma)$;
(3) Deriv. 1, $\mathbf{K_1}$, 4.1.1

(5) $\mathcal{GL} \vdash \Box(\Box\gamma \supset \gamma) \supset \Box\gamma$; by **W**

(6) $\mathcal{GL} \vdash (\Box\delta_1 \wedge ... \wedge \Box\delta_n \wedge \Box\Box\delta_1 \wedge ... \wedge \Box\Box\delta_n) \supset \Box\gamma$; (4), (5), PL

(7) $\mathcal{GL} \vdash \Box\delta_i \supset \Box\Box\delta_i$; $\mathbf{GL_1}$, 4.1.1

(8) $\mathcal{GL} \vdash (\Box\delta_1 \wedge ... \wedge \Box\delta_n) \supset (\Box\Box\delta_1 \wedge ... \wedge \Box\Box\delta_n)$; (7), PL

(9) $\mathcal{GL} \vdash (\Box\delta_1 \wedge ... \wedge \Box\delta_n) \supset \Box\gamma$; (6), (8) by PL:
$(\vdash ((p \wedge q) \supset r) \supset ((p \supset q) \supset (p \supset r))$, Subst., MP)

(10) $\mathcal{GL} \vdash \neg(\Box\delta_1 \wedge ... \wedge \Box\delta_n \wedge \neg\Box\gamma)$; (9), PL

But since all formulas $\Box\delta_1, ..., \Box\delta_n$, $\neg\Box\gamma$ are in w, it follows, by (10), that w is $\mathcal{GL}$-inconsistent. Hence Δ is $\mathcal{GL}$-consistent. Then there is an α-max-$\mathcal{S}$-con set w' such that $\Delta \subseteq w'$, wR^*w' and $\neg\gamma \in w'$. Then $\gamma \notin w'$, but $\gamma \in Set_\alpha$ (since $\Box\gamma \in Set_\alpha$). Hence $w' \not\models \gamma$ (by ind. hyp.), and so $w \not\models \Box\gamma$. So (2^*) also holds.

Whence by *Compl** it follows the completeness of $\mathcal{GL}$.

4.2. Modal logic of provability

In the preceding section (4.1) we presented some of the well-known systems of propositional modal logic, their syntax and semantics. In the context we have in view, the system $\mathcal{GL}$ is that primarily interest us, since it is usually called *the modal logic of provability*, viz. the modal logic as it is applied to the investigation of the notion of *provability* of PA^{ax}.

The language of $\mathcal{GL}$ is that specified in 4.1.1 (and *Remark*); i.e., the primitive symbols are p, q, r,..., $\perp$, $\supset$, $\Box$, (,), and then the notion "formula" will be strictly defined by the following clauses:

1. every propositional variable is a formula.
2. $\perp$ is a formula.
3. if α is a formula, then $\neg\alpha$ and $\Box\alpha$ are formulas.
4. if α and β are formulas, then $\alpha\supset\beta$ is a formula.

4.2.1. Provability predicate for PA^{ax}. Fixed points. Löb's Theorem

As we saw (comp. Ch. 3, 4.1), $Pf(y,x)$: "*y is the Gödel number of a proof of the formula with Gödel number x*" is a primitive recursive relation. So, it is formally expressible in PA^{ax} (cf. Ch. 3, 3.4) by a formula $\Pi(y,x)$. Let $\exists y\Pi(y,x)$ be the Σ_1-formula whose intuitive meaning is "the formula with Gödel number *x* is provable", viz. it is the Σ_1-relation expressed by the Σ_1-formula $\exists y\Pi(y,x)$. Let Bew(x)[27] be the formula $\exists y\Pi(y,x)$, usually called *the provability predicate for* PA^{ax}.

Notation. As we know, if *n* is the Gödel number of a formula α, then $\bar{n}$ is the *numeral* for *n*. In what follows the expression $\ulcorner\alpha\urcorner$ will be used with the same meaning: the numeral corresponding to the Gödel number of α. "Bew$\ulcorner\alpha\urcorner$)" and "Bew[α]" will mean "α is provable", according as α is a sentence (closed formula of L_{PA}) or α has free variables. This notation is borrowed from Boolos [1993], Ch. 2. Finally, since "$\perp$" is a primitive symbol of the language of $\mathcal{GL}$, in the considerations of this section the symbol "$\perp$" will be taken as primitive logical symbol of L_{PA}, and then $\perp$, Bew($\ulcorner\perp\urcorner$), $\neg$Bew($\ulcorner\perp\urcorner$) will be formulas of L_{PA}.

Even if we read both expressions "Bew($\ulcorner\alpha\urcorner$)" and "$\vdash\alpha$" as "α is provable", there is a notable difference between them: Bew($\ulcorner\alpha\urcorner$) is a *formula* of L_{PA} (i.e., it belongs to the object-language), while "$\vdash\alpha$" is an

[27] As we mentioned, "Bew" is for the Germ. "beweisbar" (provable).

expression of the metalanguage.

The predicate Bew(x) is characterized by the following (provable) properties:[28]

D_1. If $\vdash\alpha$, then $\vdash \mathrm{Bew}(\ulcorner\alpha\urcorner)$.

D_2. $\vdash \mathrm{Bew}(\ulcorner\alpha\supset\beta\urcorner)\supset\mathrm{Bew}(\ulcorner\alpha\urcorner)\supset\mathrm{Bew}(\ulcorner\beta\urcorner)$.

D_3. $\vdash \mathrm{Bew}(\ulcorner\alpha\urcorner)\supset\vdash \mathrm{Bew}(\ulcorner\mathrm{Bew}(\ulcorner\alpha\urcorner)\urcorner)$.

As we saw above (comp. Ch. 3, 4.2.2.1) by Diagonal Lemma (DL) if $\beta(x)$ (with x free) is an arbitrary formula of L_{PA}, then there is a sentence G such that $PA^{ax}\vdash G\equiv\beta(\bar{g})$, where g is the Gödel number of G.

G is called the *fixed point* for $\beta(x)$, and this provable equivalence may be taken to assert that G has the property expressed by the formula $\beta(x)$. Now, since Bew(x), i.e., $\exists y\Pi(y,x)$, and $\neg$Bew(x), i.e., $\neg\exists y\Pi(y,x)$, are formulas of L_{PA} with x free, by DL for these formulas there are the sentences G (fixed points of the corresponding formulas) such that

(a) $PA^{ax}\vdash G\equiv\mathrm{Bew}(\ulcorner G\urcorner)$

(b) $PA^{ax}\vdash G\equiv\neg\mathrm{Bew}(\ulcorner G\urcorner)$.

The sentence G in (b) is a sentence *equivalent* to a sentence asserting that G is not provable. Then by Gödel's Theorem (comp. Ch. 3, 4.2.2.2) if PA^{ax} is consistent, then G is not provable in PA^{ax}.

What happens with the sentence G in (a), is it provable or not? This question was raised by L. Henkin[29] and the answer was given by M. Löb[30]. Actually, what Löb proved is a stronger fact: viz., *any sentence implied by its own provability is provable*, subject of the well-known Löb's Theorem.

Löb's Theorem. *If* $PA^{ax}\vdash\mathrm{Bew}(\ulcorner\alpha\urcorner)\supset\alpha$, *then* $PA^{ax}\vdash\alpha$.

Proof.[31] Let (1) $PA^{+}=PA^{ax}\cup\{\neg\alpha\}$, an axiomatizable extension of PA^{ax}.

---

[28] These properties are called *Hilbert-Bernays-Löb derivability conditions*. Formulated initially (in a clumsy way) by Hilbert and Bernays [1939], §5: 1c) and 2c), they were simplified by Löb [1955] (in the form given here). The label "derivability conditions" comes from the fact that these properties are sufficient conditions on Bew(x) and a formal system S for deriving in S the Gödelian second incompleteness theorem. For the proof of D_1-D_3, comp. G. Boolos [1993], Ch. 2 and C. Smorynski [1984], 446-7 (Theorem 1.1), and [1985], Ch. 0 (Lemma 5.13).

[29] Cf. L. Henkin [1952].

[30] Cf. M.H. Löb [1954].

[31] This proof is a variant of Buss proof; comp. S. Buss [1998], 122. For other proofs, comp. C. Smorynski [1977], Sect. 4.1 (in J. Barwise (ed.) [1977], G. Boolos [1993], Ch. 3 (an elegant proof using the idea of fixed point (by DL) of the formula Bew(x)$\supset$S and the derivability conditions D_1-D_3; there the author also mentioned a variant proof due to Kreisel

Then (2) PA^+ is consistent iff $PA^{ax} \nvdash \alpha$ (argue that!). Let $Con(PA^+)$ be the formula expressing the consistency of PA^+. Evidently, the following holds (3) $Con(PA^+)$ iff $\neg Bew(\ulcorner\alpha\urcorner)$ in PA^{ax}. Suppose that $PA^{ax} \vdash Bew(\ulcorner\alpha\urcorner) \supset \alpha$, equivalently (4) $PA^{ax} \vdash \neg\alpha \supset \neg Bew(\ulcorner\alpha\urcorner)$. It follows that (5) $PA^{ax} \vdash \neg\alpha \supset Con(PA^+)$ (from (3) and (4)). And this does imply $PA^{ax}, \neg\alpha \vdash Con(PA^+)$, where $PA^{ax}, \neg\alpha$ means $PA^{ax} \cup \{\neg\alpha\}$. And this means (6) $PA^+ \vdash Con(PA^+)$, contrary to second incompleteness theorem. Whence PA^+ is inconsistent, and therefore (7) $PA^{ax} \vdash \alpha$ (by (2)).

From this result easy follows that any fixed point of Bew(x) is provable, since $PA^{ax} \vdash \alpha \equiv Bew(\ulcorner\alpha\urcorner)$ (by DL), and then $PA^{ax} \vdash Bew(\ulcorner\alpha\urcorner) \supset \alpha$ (by PL). Whence, by Löb's Theorem $PA^{ax} \vdash \alpha$.

Let us observe that the converse of Löb's Theorem also holds: If $PA^{ax} \vdash \alpha$, then $PA^{ax} \vdash Bew(\ulcorner\alpha\urcorner) \supset \alpha$. Since, as we know (by Ch. 2, Sect. 3.1, Ax1) $PA^{ax} \vdash \alpha \supset (Bew(\ulcorner\alpha\urcorner) \supset \alpha)$ and if $PA^{ax} \vdash \alpha$, then $PA^{ax} \vdash Bew(\ulcorner\alpha\urcorner) \supset \alpha$ (by MP). And then, as a general result the following holds:

(*Eq*) $PA^{ax} \vdash Bew(\ulcorner\alpha\urcorner) \supset \alpha$ iff $PA^{ax} \vdash \alpha$.

4.2.2. $\mathcal{GL}$ and PA^{ax}

For the subject we want to analyse here the following questions are central: Wherein the correspondence between $\mathcal{GL}$ and PA^{ax} consists? Why the *modal* system $\mathcal{GL}$ is properly taken as the logic of provability for the formal system of *arithmetic* PA^{ax}? In order to answer these questions, let us begin with what is meant by a *translation* of the formulas of $L_{\mathcal{GL}}$ in the corresponding formulas of L_{PA}.[32] First of all, a *realization* (*)[33] is a function from the set of propositional variables of modal logic to the sentences of L_{PA}. Then the translation α^* (always a formula of L_{PA}) of a modal formula is defined inductively as follows:

$p^* = S$; where p is a propositional variable of the modal language

$\perp^* = \perp$

and Takeuti [1974]). For an interesting proof of Löb's Theorem, *via* DL, and a proof of its *formalized* form (*via* Kreisel and Takenti's [1974] proof) comp. C. Smorynski [1984], 451-3. For some other proof, comp. H. Friedman and M. Sheard [1987].

[32] The considerations of this section are essentially based on G. Boolos [1993], Ch. 3, C. Smorynski [1985], Ch. 1.

[33] Or *interpretation*, the name given by Boolos [1993], 51, C. Smorynski [1985], 64.

$(\alpha\supset\beta)^* = (\alpha^*\supset\beta^*)$; where $\alpha\supset\beta$ is a modal formula

$(\Box\alpha)^* = \mathrm{Bew}[\alpha^*]$, or $\mathrm{Bew}(\ulcorner\alpha^*\urcorner)$ if α is closed.

Given this translation, the correspondence between the modal language of $\mathcal{GL}$ and the arithmetic language L_{PA} can be easily seen. If, for example, * is an arbitrary realization and $\alpha = \neg\Box\bot\supset\neg\Box\neg\Box\bot$, then $\alpha^* = (\neg\Box\bot\supset\neg\Box\neg\Box\bot)^*$ is the formula of L_{PA}:

$\neg\mathrm{Bew}(\ulcorner\bot\urcorner)\supset\neg\mathrm{Bew}(\ulcorner\neg\mathrm{Bew}(\ulcorner\bot\urcorner)\urcorner)$, whose meaning is:

"If PA^{ax} is consistent, then the formula expressing its consistency (i.e., $\neg\mathrm{Bew}(\ulcorner\bot\urcorner)$) is not provable in PA^{ax} " (and this, as we know, is just the second incompleteness theorem (comp. Ch. 3, 4.2.1.2). If $\alpha = \Box\neg p$, then $\alpha^* = \mathrm{Bew}(\ulcorner\neg S\urcorner)$, i.e., S is disprovable (refutable). If $\alpha = \neg\Box p\wedge\neg\Box\neg p$, then $\alpha^* = \neg\mathrm{Bew}(\ulcorner S\urcorner)\wedge\neg\mathrm{Bew}(\ulcorner\neg S\urcorner)$, i.e., S is undecidable, etc.

What makes $\mathcal{GL}$ be the logic of provability is the following relation between $\mathcal{GL}$ and PA^{ax} (under the interpretation of $\Box$ as Bew(x)).

Theorem 1. $\mathcal{GL}\vdash\alpha$ *if and only if for every realization* *, $PA^{ax}\vdash\alpha^*$.

This means that every theorem of $\mathcal{GL}$ is a theorem of PA^{ax} for every realization (the *arithmetical soundness* of $\mathcal{GL}$) and its converse, every formula of $\mathcal{GL}$ that is provable in PA^{ax} for every realization is also provable in $\mathcal{GL}$ (*arithmetical completeness* of $\mathcal{GL}$[34]).

By Theorem 1, $\mathcal{GL}\vdash\Box(\Box\bot\supset\bot)\supset\Box\bot$ (comp. Sect. 4.1.1, Theorem **GL**$_2$, or by Axiom **W**), equivalent $\mathcal{GL}\vdash\neg\Box\bot\supset\neg\Box\neg\Box\bot$, if and only if $PA^{ax}\vdash(\neg\Box\bot\supset\neg\Box\neg\Box\bot)^*$, i.e., $PA^{ax}\vdash\neg\mathrm{Bew}(\ulcorner\bot\urcorner)\supset\neg\mathrm{Bew}(\ulcorner\neg\mathrm{Bew}(\ulcorner\bot\urcorner)\urcorner)$, i.e., the second incompleteness theorem.

Another example illustrating the content of Theorem 1 is given by the following argument:

(1) $\mathcal{GL}\vdash\neg\Box\bot\supset\neg\Box\neg\Box\bot$; from **W** by PL

(2) $\mathcal{GL}\vdash\Box\bot\supset\Box\Box\bot$; by 4.1.1, Th. **GL**$_1$

(3) $\mathcal{GL}\vdash\neg\Box\Box\bot\supset\neg\Box\bot$; (2), PL

(4) $\mathcal{GL}\vdash\neg\Box\Box\bot\supset\neg\Box\neg\Box\bot$; (1), (3), PL

(5) $\mathcal{GL}\vdash\neg\Box\Box\bot\supset\neg\Box\neg\neg\Box\bot$; Repl.

(6) $\mathcal{GL}\vdash\neg\Box\Box\bot\supset(\neg\Box\neg\Box\bot\wedge\neg\Box\neg\neg\Box\bot)$; (4), (5), PL

[34] A remarkable result due to R. Solovay [1976]. For a proof of both parts of this theorem (*arithmetical* soundness and *arithmetical* completeness of $\mathcal{GL}$), comp. also G. Boolos [1993], 59 and Ch. 9, respectively.

(7) $\mathrm{PA}^{ax} \vdash (\neg\Box\Box\bot \supset (\neg\Box\neg\Box\bot \wedge \neg\Box\neg\neg\Box\bot))^{*}$; (6) by Theorem 1. What (7) says is the following thing: *if the inconsistency of* PA^{ax} *is not provable, then neither its consistency, nor its inconsistency is provable in* PA^{ax}. I.e., the formula of L_{PA} $\neg\mathrm{Bew}(\ulcorner\bot\urcorner)$, expressing the consistency of PA^{ax}, is undecidable in PA^{ax}.

Remark. As can be observed from some of these examples, $\mathcal{GL}$ gives an interesting analysis of the so-called *constant sentences* of L_{PA}, corresponding to *letterless formulas* of $L_{\mathcal{GL}}$.

If α is a letterless sentence, then the truth value of α^* is the same for all realizations *. And, evidently, for every constant sentence S of L_{PA} there is a letterless sentence β such that for all realizations * the following holds: $S = \beta^*$. And for these constant sentences there is an effective method which allows us to establish their truth. The proof of this fact is based on a normal form theorem, according to which for any letterless formula β there is a letterless formula γ in normal form (i.e., γ is a truth-functional combination of formulas of the form $\Box^i\bot$ $(i = 0,1,\ldots,k)$) such that $\mathcal{GL} \vdash \beta \equiv \gamma$.[35]

Comments. As we saw (comp. 4.1.1), the system $\mathcal{GL}$ is a proper extension of $\mathcal{K}$, its proper axiom being **W**: $\Box(\Box p \supset p) \supset \Box p$. The rules of $\mathcal{GL}$ are those of any *normal* modal systems: Subst, MP and N. But $\mathcal{GL}$ is *not* an extension (proper or not) of $\mathcal{T}$ and then it is different from all modal systems containing the axiom **T**: $\Box p \supset p$ (e.g., $\mathcal{S}4$, $\mathcal{S}5$, $\mathcal{B}$). And this is a relevant fact, since if $\Box p \supset p$ were provable in $\mathcal{GL}$, then $\Box\bot \supset \bot$ would be provable (by Subst), and then $\Box(\Box\bot \supset \bot)$ would be provable as well (by N). Whence, finally, $\Box\bot$ would be a theorem of $\mathcal{GL}$ (by **W** and MP). Hence $\mathcal{GL}$ would be inconsistent. Therefore, if $\mathcal{GL}$ is consistent, then not any formula of the form $\Box\alpha \supset \alpha$ is a theorem of $\mathcal{GL}$.

Since $\mathcal{GL} \nvdash \Box p \supset p$, it follows that $\mathcal{GL} \nvdash \Box\bot \supset \bot$ (by Subst), equivalently $\mathcal{GL} \nvdash \neg\Box\bot$, equivalently $\mathrm{GL} \nvdash \Diamond\top$. But $\mathcal{GL} \vdash \top$. This means that the set of its theorems *is not closed under possibility* (though as a *normal* modal system it is closed under necessity (the rule N)).[36]

Now, the following things hold for PA^{ax}. By soundness of PA^{ax} if

[35] The proof of decidability of constant sentences uses the idea of finding the letterless formula β corresponding to β^* (such that $S = \beta^*$), construct the normal form γ of β and apply PL; for details, cf. G. Boolos [1979], 62, G. Boolos [1993], Ch. 7.

[36] Remember that if R is a rule of deduction and S is a set of formulas, S is closed under R if it contains all formulas deducible by R from the members of S.

$PA^{ax} \vdash S$, then S is true in M (the standard model of L_{PA}), therefore $Bew(\ulcorner S \urcorner)$ is a true sentence. But $Bew(\ulcorner S \urcorner)$ is just the formalization of "$\vdash S$". Then for every sentence S of L_{PA}, $Bew(\ulcorner S \urcorner) \supset S$ is a true sentence of L_{PA} (and this holds for any sentence S of L_{PA}). Therefore, if α is a modal formula, then $(\Box\alpha \supset \alpha)^*$ is a true sentence of L_{PA}, for any realization *.

Moreover, any sentence derivable by applications of MP is true. Since all these formulas are true for every realization, the modal system that embeds all these cases is a new one, called $\mathcal{GLS}$ (Gödel-Löb-Solovay). So, the axioms of $\mathcal{GLS}$ are all theorems of $\mathcal{GL}$ and all sentences of the form $\Box\alpha \supset \alpha$, and MP as its sole rule of inference.[37] For this system the theorem above also holds. And then the rule N of $\mathcal{GL}$ is not a derivable rule in $\mathcal{GLS}$, too. Otherwise, since $\Box\bot \supset \bot$ is an axiom of $\mathcal{GLS}$, $\Box(\Box\bot \supset \bot)$ would be a theorem of $\mathcal{GLS}$. Therefore $PA^{ax} \vdash \Box(\Box\bot \supset \bot)^*$, i.e., $PA^{ax} \vdash Bew(\ulcorner Bew(\ulcorner\bot\urcorner) \supset \bot\urcorner)$. And then $Bew(\ulcorner Bew(\ulcorner\bot\urcorner) \supset \bot\urcorner)$ is true (by soundness of PA^{ax}). Whence $PA^{ax} \vdash Bew(\ulcorner\bot\urcorner) \supset \bot$, i.e., $PA^{ax} \vdash \neg Bew(\ulcorner\bot\urcorner)$, viz. $PA^{ax} \vdash Con_{PA^{ax}}$[38] (contrary to second incompleteness theorem).

Unlike $\mathcal{GL}$, the modal system $\mathcal{GLS}$ *is closed under possibility*, since if $\mathcal{GLS} \vdash \alpha$ and since all formulas of the form $\Box\alpha \supset \alpha$ are axions of $\mathcal{GLS}$ (for any α), it follows that $\mathcal{GLS} \vdash \Box\neg\alpha \supset \neg\alpha$, equivalently, $\mathcal{GLS} \vdash \alpha \supset \neg\Box\neg\alpha$ (by PL), i.e., $\mathcal{GLS} \vdash \Diamond\alpha$ (by Interch.). And then all formulas of the form $\top$, $\Diamond\top$, $\Diamond\Diamond\top$, etc. are theorems of $\mathcal{GLS}$.

Now, let us ask the following question: how the derivability conditions D_1-D_3 might be like if we "read" them *modally*? The answer can easily be given by the following simple analysis. Let α and β be modal formulas, let α^*, β^* be the formulas of L_{PA} under an arbitrary realization *. Then the derivability conditions D_1-D_3 can be expressed in the following way:

D_1^*. If $PA^{ax} \vdash \alpha^*$, then $PA^{ax} \vdash (\Box\alpha)^*$.

D_2^*. $PA^{ax} \vdash (\Box(\alpha \supset \beta) \supset (\Box\alpha \supset \Box\beta))^*$.

D_3^*. $PA^{ax} \vdash (\Box\alpha \supset \Box\Box\alpha)^*$.

[37] And then $\mathcal{GLS}$ is not a *normal* system of modal logic.

[38] Equivalently, *via* first incompleteness theorem, PA^{ax} *proves* its inconsistency (detail!).

And then the modal counterpart of D_1^* - D_3^* are, evidently, the following expressions:

N. If $\vdash\alpha$, then $\vdash\Box\alpha$.

K. $\vdash\Box(\alpha\supset\beta)\supset(\Box\alpha\supset\Box\beta)$.

4. $\vdash\Box\alpha\supset\Box\Box\alpha$,

where **N** is the necessity rule, **K** is the modal axiom of $\mathcal{K}$ and **4** is the proper axiom of $\mathcal{K}4$. Given that Subst and MP are valid rules of deduction for every normal system of modal logic, it follows that what we have here is just the modal system $\mathcal{K}4$. And this fact leads to the following result.

Theorem 2. *If* $\mathcal{K}4\vdash\alpha$, *then for every realization* *, $PA^{ax}\vdash\alpha^*$.

Remark. Evidently, $\mathcal{K}4$ and $\mathcal{GL}$ are not identical. But if to $\mathcal{K}4$ is added the so-called *Löb's Rule* (if $\vdash(\Box\alpha\supset\alpha)$, then $\vdash\alpha$), then, as can be expected, the resulting system and $\mathcal{GL}$ are equivalent (i.e., the two systems prove the same theorems).

As we know, $\Box p\supset p$ is not a theorem of $\mathcal{GL}$. $(\Box p\supset p)^*$ is the formula of L_{PA} $Bew(\ulcorner S\urcorner)\supset S$. Such a formula is called *reflection principle* for PA^{ax}. Actually, a reflection principle is a *formal* assertion stating the soundness of a formal system. For a given α, $Bew(\ulcorner\alpha\urcorner)\supset\alpha$ is called the reflection principle of α. By (*Eq*) (comp. Sect. 4.2.1 above) not every reflection is provable. If α is the Gödel sentence G, then by (*Eq*) $PA^{ax}\nvdash Bew(\ulcorner G\urcorner)\supset G$. On the other hand, this result can be obtained also directly from $\vdash(p\equiv\neg q)\supset(q\supset p)\equiv p)$, plus Subst G/p, $Bew(\ulcorner G\urcorner)$/q, $PA^{ax}\vdash G\equiv\neg Bew(\ulcorner G\urcorner)$ and MP (detail!).[39]

As we know (comp. Ch. 3, Sect. 4.2.1.2(2)), a formal theory *T* is consistent iff there is no formula α such that $T\vdash\alpha$ and $T\vdash\neg\alpha$.[40] *T* is ω-consistent iff for no formula $\alpha(x)$ the following simultaneous hold: (a) for every *n*, $T\vdash\neg\alpha(\bar{n})$ and (b) $T\vdash\exists x\alpha(x)$. Evidently, if $T\nvdash\exists x\alpha(x)$, this does imply that *T* does not prove any formula, and then *T* is consistent. But the converse is not true, i.e., the conditional if *T* is consistent, then *T* is ω-consistent does not hold.

---

[39] For excellent comments on *reflection principles*, comp. C. Smorynski [1977] and G. Kreisel and A. Levy [1968].

[40] Equivalently, *T* does not prove any formula, no contradiction is provable in *T*, $T\nvdash\bot$, $T\nvdash 0=\bar{1}$. The closed formula expressing the consistency of PA^{ax} is $\neg$ $Bew(\ulcorner\bot\urcorner)$.

Argument. Let $PA^+ = PA^{ax} \cup \{\neg G\}$ (where G is the Gödelian sentence for PA^{ax}), equivalently, $PA^+ = PA^{ax} \cup Bew(\ulcorner \perp \urcorner)$ (since $PA^{ax} \vdash G \equiv Con(PA^{ax}) \equiv \neg Bew(\ulcorner \perp \urcorner)$, (by (EQ) (below, p.310)). Since $PA^{ax} \nvdash G$ (if PA^{ax} is consistent), then PA^+ is also consistent (by Ch. 2, Sect. 3.5.3, Lemma 2). Hence the intuitive sentence $Pf(n, \perp)$ is false *for every n*, and then $PA^{ax} \vdash \neg\Pi(\bar{n}, \ulcorner \perp \urcorner)$ for every n (by formal expressibility of $Pf(y,x)$ by $\Pi(y,x)$ (cf. Ch. 3, 3.4, final Remark) and then $PA^+ \vdash \neg\Pi(\bar{n}, \ulcorner \perp \urcorner)$ (since $PA^{ax} \subseteq PA^+$). But $PA^+ \vdash Bew(\ulcorner \perp \urcorner)$, viz. $PA^+ \vdash \exists y \Pi(y, \ulcorner \perp \urcorner)$, and this means that PA^+ is ω-inconsistent.

The 1-*consistency* of a theory T is the special case of ω-consistency, when α(x) from the definition of ω-consistency is a *decidable* formula of L_{PA}. Evidently, the following holds: ω-consistency does imply 1-consistency and 1-consistency does imply consistency.

And then, by the argument above, from the consistency of PA^{ax} does not follow that PA^{ax} is 1-consistent too.

Theorem 3. *Let* α *be a closed formula of* L_{PA} *such that* $PA^{ax} \nvdash \alpha$. *Then if* PA^{ax} *is* 1-*consistent, then* $PA^{ax} \nvdash Bew(\ulcorner\alpha\urcorner)$.

Proof (*reductio*). Suppose that PA^{ax} is 1-consistent, $PA^{ax} \nvdash \alpha$ and $PA^{ax} \vdash Bew(\ulcorner\alpha\urcorner)$. Then since α is not provable in PA^{ax}, it follows that for every n, $Pf(n,a)$ is false (where a is the Gödel number of α). It follows that $PA^{ax} \vdash \neg\Pi(\bar{n}, \ulcorner\alpha\urcorner)$ for every n (since $\Pi(y,x)$ formally expresses $Pf(y,x)$ in PA^{ax}). Now, since $PA^{ax} \vdash Bew(\ulcorner\alpha\urcorner)$, equivalently $PA^{ax} \vdash \exists y \Pi(y, \ulcorner\alpha\urcorner)$ it follows that PA^{ax} is 1-inconsistent (contrary to the hypothesis).

Corollary. *If* PA^{ax} *is* 1-*consistent, then* PA^{ax} *has not the theorems of the following form:* $\perp$, $Bew(\ulcorner \perp \urcorner)$, $Bew(\ulcorner Bew(\ulcorner \perp \urcorner)\urcorner)$, *etc.*

Proof. (By Theorem 3, for $\alpha = \perp$). An independent argument for this corollary can also be given in the following terms. If $PA^{ax} \vdash \perp$, then PA^{ax} is inconsistent and therefore it is 1-inconsistent (contra hyp.). Or, since $PA^{ax} \vdash \neg G \equiv Bew(\ulcorner \perp \urcorner)$, then if $Bew(\ulcorner \perp \urcorner)$ were provable in PA^{ax}, then, by Ch. 3, 4.2.1.2 (Gödel's first incompleteness theorem, part (2)), PA^{ax} would be 1-inconsistent.

Now, from Theorem 1 and Corollary it follows that the modal system $\mathcal{GL}$ has not the theorems of the form $\perp$, $\Box\perp$, $\Box\Box\perp$, etc.

As we saw (comp. Ch. 3, 4.2.1.2, Gödel's second incompleteness theorem), the conditional "if PA^{ax} is consistent, then the consistency of PA^{ax}

is not provable in PA^{ax} " is *formalizable* and, moreover, *provable* in PA^{ax}. Similarly, the following conditional is also formalizable and provable in PA^{ax}: "if PA^{ax} is 1-consistent, then the inconsistency of PA^{ax} is not provable in PA^{ax} ". The inconsistency of PA^{ax} is just the formula $Bew(\ulcorner\bot\urcorner)$. And by definition, PA^{ax} is 1-consistent if the following items does not hold simultaneously: $PA^{ax}\vdash\exists x\alpha(x)$ and for every n: $PA^{ax}\vdash\neg\alpha(\overline{n})$, i.e.,

Not($\vdash\exists x\alpha(x)$ and for every n: $\vdash\neg\alpha(\overline{n})$,

equivalently

If $\vdash\exists x\alpha(x)$, then Not for every n: $\vdash\neg\alpha(\overline{n})$.

This conditional expressing 1-consistency is formalizable and, moreover, the *whole* conditional mentioned above is formalizable and provable in PA^{ax} [41], i.e., $PA^{ax}\vdash 1Con\supset\neg Bew(\ulcorner\bot\urcorner)$.So is the case with formulas of the form $1Con\supset\neg Bew(\ulcorner Bew(\ulcorner\bot\urcorner)\urcorner)$, $1Con\supset\neg Bew(\ulcorner Bew(\ulcorner Bew(\ulcorner\bot\urcorner)\urcorner)\urcorner)$.

Using this result some other argument for the fact that consistency of PA^{ax} does not imply 1-consistency of PA^{ax} can be given. For, as we just mentioned, $PA^{ax}\vdash 1Con\supset\neg Bew(\ulcorner Bew(\ulcorner\bot\urcorner)\urcorner)$, and then if the implication $\neg Bew(\ulcorner\bot\urcorner)\supset 1Con$ were provable in PA^{ax}, then, by PL, the formula $\neg Bew(\ulcorner\bot\urcorner)\supset\neg Bew(\ulcorner Bew(\ulcorner\bot\urcorner)\urcorner)$ would be provable as well. Whence, by contraposition, $PA^{ax}\vdash Bew(\ulcorner Bew(\ulcorner\bot\urcorner)\urcorner)\supset Bew(\ulcorner\bot\urcorner)$; and therefore $PA^{ax}\vdash Bew(\ulcorner\bot\urcorner)$ (by Löb's Theorem). This means that PA^{ax} would be 1-inconsistent (since $PA^{ax}\vdash\neg G\equiv Bew(\ulcorner\bot\urcorner)$, whence $PA^{ax}\vdash\neg G$, and then, by first incompleteness theorem (part (2)) PA^{ax} is 1-inconsistent).

Let us illustrate once more the content of Theorem 1. As we saw (comp. 4.1.1)

GL5 (a) $\vdash\Box(p\equiv\neg\Box p)\equiv\Box(p\equiv\neg\Box\bot)$.

Then, by Theorem 1,

$PA^{ax}\vdash((\Box(p\equiv\neg\Box p)\equiv\Box(p\equiv\neg\Box\bot))^*$, i.e.,

$PA^{ax}\vdash Bew(\ulcorner S\equiv\neg Bew(\ulcorner S\urcorner)\urcorner)\equiv Bew(\ulcorner S\equiv\neg Bew(\ulcorner\bot\urcorner)\urcorner)$.

Then the following assertion *is provable* in PA^{ax}: "a sentence S of L_{PA} is equivalent in PA^{ax} to its own unprovability iff S is equivalent to the consistency of PA^{ax} ".

Informally, from this expression provable in PA^{ax} the following biconditional can be derived:

---

[41] For details, comp. C. Smorynski [1977], 4.2 and G. Kreisel and A. Lévy [1968].

(EQ) $PA^{ax} \vdash S \equiv \neg Bew(\ulcorner S \urcorner)$ iff $PA^{ax} \vdash S \equiv \neg Bew(\ulcorner \perp \urcorner)$.[42]

This "iff" can also be proved independently. Since the left result does hold by DL (i.e., S is the fixed point of the formula $\neg Bew(x)$), all that remains to be proved is the right result. And this means that we have to show the following

(a) $PA^{ax} \vdash \neg Bew(\ulcorner \perp \urcorner) \supset S$, and

(b) $PA^{ax} \vdash S \supset \neg Bew(\ulcorner \perp \urcorner)$.

Proof. (a) We have the following derivations:

(1) $PA^{ax} \vdash S \equiv \neg Bew(\ulcorner S \urcorner)$; by DL

(2) $PA^{ax} \vdash \neg S \equiv Bew(\ulcorner S \urcorner)$; (1), PL

(3) $PA^{ax} \vdash Bew(\ulcorner \neg S \equiv Bew(\ulcorner S \urcorner) \urcorner)$; (2) by D_1 (of Sect. 4.2.1)

(4) $PA^{ax} \vdash Bew(\ulcorner \neg S \urcorner) \equiv Bew(\ulcorner Bew(\ulcorner S \urcorner) \urcorner)$; (3)

(The derivation of (4) from (3) can be carried out *via* the modal system $\mathcal{GL}$, since (3) has the form $\Box(\alpha \equiv \beta)$, and then, by Deriv. 2 (of 4.1.1), we obtain $\Box\alpha \equiv \Box\beta$, whose arithmetical counterpart is (4)).

(5) $PA^{ax} \vdash Bew(\ulcorner S \urcorner) \supset Bew(\ulcorner Bew(\ulcorner S \urcorner) \urcorner)$; by Sect. 4.2.1, D_3

(6) $PA^{ax} \vdash Bew(\ulcorner S \urcorner) \supset Bew(\ulcorner \neg S \urcorner)$; (5), (2), PL

(7) $PA^{ax} \vdash S \supset (\neg S \supset \perp)$; by PL

(8) $PA^{ax} \vdash \neg S \supset Bew(\ulcorner \perp \urcorner)$.

In order to derive (8) from (7) we proceed as follows: apply D_1 to the formula from (7) and then D_2 (twice), obtaining $PA^{ax} \vdash Bew(\ulcorner S \urcorner) \supset Bew(\ulcorner \neg S \urcorner) \supset Bew(\ulcorner \perp \urcorner))$, which together with D_3, *via* PL, gives $PA^{ax} \vdash Bew(\ulcorner S \urcorner) \supset Bew(\ulcorner \perp \urcorner)$. Whence, by (1) and PL, $PA^{ax} \vdash \neg S \supset Bew(\ulcorner \perp \urcorner)$, (i.e., (8)) and therefore $PA^{ax} \vdash \neg Bew(\ulcorner \perp \urcorner) \supset S$, equivalently $PA^{ax} \vdash Con(PA^{ax}) \supset S$.

Remark. Since PA^{ax} (if consistent) does not prove S, i.e., $PA^{ax} \nvdash S$,[43] by one application of *modus tollens*, we derive $PA^{ax} \nvdash Con(PA^{ax})$; and this is just another proof of the Gödel's second incompleteness theorem.

(b) We have the following derivations:

(1) $PA^{ax} \vdash \perp \supset S$; PL

[42] Remember, $\neg Bew(\ulcorner \perp \urcorner)$ is the formula expressing in L_{PA} the *consistency* of PA^{ax}; equivalent notation $Con(PA^{ax})$.

[43] See below, for S = G, the proof of the first part of Gödel's first incompleteness theorem.

(2) $PA^{ax} \vdash Bew(\ulcorner \bot \supset S \urcorner)$; (1), D_1

(3) $PA^{ax} \vdash Bew(\ulcorner \bot \urcorner) \supset Bew(\ulcorner S \urcorner)$; (2), D_2

(4) $PA^{ax} \vdash Bew(\ulcorner \bot \urcorner) \supset \neg S$; (3), by (1) from Proof (a)

(5) $PA^{ax} \vdash S \supset \neg Bew(\ulcorner \bot \urcorner)$; (4), PL, i.e.,

$PA^{ax} \vdash S \supset Con(PA^{ax})$.

From the proofs (a) and (b) it follows $PA^{ax} \vdash S \equiv \neg Bew(\ulcorner \bot \urcorner)$, i.e., $PA^{ax} \vdash S \equiv Con(PA^{ax})$, and therefore it follows (EQ). And what (EQ) shows is that a sentence asserting its own unprovability is equivalent to the assertion of consistency of PA^{ax}, and the provability of their equivalence shows the uniqueness of such a sentence.

Now, let us take only one half from **GL$_5$** (a) and make the following derivation:

(1) $\mathcal{GL} \vdash \Box(p \equiv \neg \Box p) \supset \Box(p \equiv \neg \Box \bot)$; from **GL$_5$** (a)

(2) $\mathcal{GL} \vdash \Box(p \equiv \neg \Box \bot) \supset (\Box p \equiv \Box \neg \Box \bot)$; by **K**, **K$_1$**

(3) $\mathcal{GL} \vdash \Box \bot \equiv \Box \Diamond p$; by 4.1.1, **GL$_4$**

(4) $\mathcal{GL} \vdash \Box \bot \equiv \Box \Diamond \top$; (3), Subst.

(5) $\mathcal{GL} \vdash \Box \bot \equiv \Box \neg \Box \bot$; (4), Interch.

(6) $\mathcal{GL} \vdash \Box(p \equiv \neg \Box \bot) \supset (\Box p \equiv \Box \bot)$; (2), (5), PL

(7) $\mathcal{GL} \vdash \Box(p \equiv \neg \Box p) \supset (\Box p \equiv \Box \bot)$; (1), (6), PL

Therefore

$PA^{ax} \vdash (\Box(p \equiv \neg \Box p) \supset (\Box p \equiv \Box \bot))^*$, i.e.,

$PA^{ax} \vdash Bew(\ulcorner S \equiv \neg Bew(\ulcorner S \urcorner) \urcorner) \supset (Bew(\ulcorner S \urcorner) \equiv Bew(\ulcorner \bot \urcorner))$.

And this means that PA^{ax} proves the following assertion: "if S is equivalent to its own unprovability, then S is provable in PA^{ax} iff PA^{ax} is inconsistent".

Similarly, we can find the assertions corresponding to **GL$_5$** (b)-(d), provable in PA^{ax} (exercise).

Now the following example illustrates the use of modal logic ($\mathcal{GL}$) in proving the undecidability of the Gödelian sentence G in PA^{ax}.[44]

---

[44] Actually, $\mathcal{GL}$ is used only in the proof of the second part of first incompleteness theorem (that of non-provability in PA^{ax} of $\neg$G). Evidently, the use of modal logic can be avoided if, for example, from (3) in the second proof we derive directly $PA^{ax} \vdash G$ (of course, under assumption of Σ_1-soundness of PA^{ax}, equivalent, of 1-consistency of PA^{ax} (by Ch. 3, Sect. 4.2.5, Fact 7 Lemma)).

As we know, the undecidable sentence G is a fixed point of $\neg$Bew(x), and then

(1) $PA^{ax} \vdash G \equiv \neg Bew(\ulcorner G \urcorner)$; by DL

(2) $PA^{ax} \vdash G$; hyp.

(3) $PA^{ax} \vdash Bew(\ulcorner G \urcorner)$; (2), D_1

(4) $PA^{ax} \vdash \neg Bew(\ulcorner G \urcorner)$; (1), (2), PL

(5) $PA^{ax} \vdash Bew(\ulcorner G \urcorner) \wedge \neg Bew(\ulcorner G \urcorner)$; (3), (4), PL

(6) $PA^{ax} \vdash \perp$; i.e., PA^{ax} is incon.

Therefore, (1) if PA^{ax} is consistent, then $PA^{ax} \nvdash G$.

On the other hand, we have the following derivation:

(1) $PA^{ax} \vdash G \equiv \neg Bew(\ulcorner G \urcorner)$; by DL

(2) $PA^{ax} \vdash \neg G$; hyp.

(3) $PA^{ax} \vdash Bew(\ulcorner G \urcorner)$; (1), (2), PL

(4) $PA^{ax} \vdash Bew(\ulcorner \neg G \urcorner)$; (2), D_1

(5) $PA^{ax} \vdash Bew(\ulcorner G \urcorner) \wedge Bew(\ulcorner \neg G \urcorner)$; (3), (4), PL

(6) $\mathcal{GL} \vdash (\Box p \wedge \Box \neg p) \supset \Box \perp$; by 4.1.1, $\mathbf{K_1}$, Subst. $\neg p/q$

(7) $PA^{ax} \vdash (Bew(\ulcorner G \urcorner) \wedge Bew(\ulcorner \neg G \urcorner)) \supset Bew(\ulcorner \perp \urcorner)$; (6), Th1

(8) $PA^{ax} \vdash Bew(\ulcorner \perp \urcorner)$; (5), (7), MP.

And this means that PA^{ax} is 1-inconsistent (by Corollary above). Therefore, if PA^{ax} is 1-consistent, then $PA^{ax} \nvdash Bew(\ulcorner \perp \urcorner)$, whence, $PA^{ax} \nvdash \neg G$.

Remark. As we see by this example, the *undecidability* of G does not follow from the (simple) consistency of PA^{ax}. Actually, as we know (comp. Ch. 3, 4.2.1.2), only the unprovability of $\neg$G claims the stronger assumption of 1-consistency of PA^{ax} and this fact can also be argued in the following way. Suppose that under the assumption of consistency the sentence $\neg$G were unprovable. Formally, this means that (1) $PA^{ax} \vdash \neg Bew(\ulcorner \perp \urcorner) \supset \neg Bew(\ulcorner \neg G \urcorner)$. But (2) $PA^{ax} \vdash G \equiv \neg Bew(\ulcorner \perp \urcorner)$ (G being equivalent to the consistency of PA^{ax}). And then, from (1) and (2) it follows (3) $PA^{ax} \vdash \neg Bew(\ulcorner \perp \urcorner) \supset \neg Bew(\ulcorner Bew(\ulcorner \perp \urcorner) \urcorner)$. Whence, by contraposition, $PA^{ax} \vdash Bew(\ulcorner Bew(\ulcorner \perp \urcorner) \urcorner) \supset Bew(\ulcorner \perp \urcorner)$, and then, by Löb's Theorem, $PA^{ax} \vdash Bew(\ulcorner \perp \urcorner)$. But, evidently, $PA^{ax} \nvdash Bew(\ulcorner \perp \urcorner)$. Therefore, the undecidability of G in PA^{ax} can be proved only on the stronger assumption of 1-consistency. And this fact holds for any sentence S of L_{PA} equivalent to a sentence asserting its own unprovability in PA^{ax}.[45]

A similar result can be stated in the following terms: if α is a letterless formula of $L_{\mathcal{GL}}$, then the undecidability of its arithmetical counterpart, α^*, does not follow only from the consistency of PA^{ax}. This means, *via* Theorem 1, that $\mathcal{GL} \nvdash \neg\Box\bot \supset (\neg\Box\alpha \wedge \neg\Box\neg\alpha)$,[46] and therefore $PA^{ax} \nvdash (\neg\Box\bot \supset (\neg\Box\alpha \wedge \neg\Box\neg\alpha))^*$.

4.2.3. Fixed point theorem

Preliminary

Besides the *Theorem 1* (above, Sect. 4.2.2 (*arithmetical soundness and arithmetical completeness of $\mathcal{GL}$*), another result concerning $\mathcal{GL}$ is really remarkable: *the fixed point theorem*, due to Dick de Jongh and Giovanni Sambin.[47]

First of all, remember that the *strong box* $\boxdot$ is defined as follows: $\boxdot\alpha =_{df} \Box\alpha \wedge \alpha$ (comp. 4.1.1). And a formula α is *modalized* (or boxed) *in* p if every occurrence of p in α is in the scope of an occurrence of modal operator $\Box$.

Fixed point theorem (FPT). *Let* α *be a formula modalized in* p. *Then there is a formula* β *containing only the propositional variables of* α *different from* p *and such that*

$\mathcal{GL} \vdash \boxdot(p \equiv \alpha) \equiv \boxdot(p \equiv \beta)$.

The formula β is called a *fixed point* of α.

From this theorem another form of it can be obtained in the following way:

(1) $\mathcal{GL} \vdash \Box\boxdot(p \equiv \alpha) \equiv \Box\boxdot(p \equiv \beta)$; from FPT, using Deriv. 2 (4.1.1).

(2) $\mathcal{GL} \vdash \Box(p \equiv \alpha) \equiv \Box(p \equiv \beta)$; (1), Theorem 1(a) and (b) (4.1.1).

Let us observe that if $\alpha(p)$ is a formula containing only the propositional variable p, then its fixed point will be a letterless formula. And

---

[45] And then no fixed point of $\neg$Bew(x) is provable in PA^{ax}, if PA^{ax} is consistent; and if PA^{ax} is 1-consistent, then no fixed point of $\neg$Bew(x) is disprovable in PA^{ax}.

[46] For details, comp. G. Boolos [1993], 97-98.

[47] This topic is of higher type of complexity. In order to make it as much as possible understandable and to not exceed the intended level of complexity for this book, the results are presented in outline or using comments of some distinguished authors, e.g. G. Boolos and C. Smorynski.

then by **GL₅** (a)-(d) (4.1.1) for $\alpha(p)$: $\neg\Box p$, $\Box p$, $\Box\neg p$ and $\neg\Box\neg p$, respectively, the fixed points of these formulas are, $\neg\Box\bot$, $\top$, $\Box\bot$ and $\bot$, respectively.

The proof of the Fixed Point Theorem is based on the following results:

1. **Theorem.** *Let* β *be a formula of* L_{MPL} *not containing the variable* p. *Then if* $\mathcal{GL}\vdash\boxdot(p\equiv\alpha)\supset(p\equiv\beta)$, *then* $\mathcal{GL}\vdash\boxdot(p\equiv\beta)\supset(p\equiv\alpha)$.[48]

2. $\mathcal{GL}\vdash\boxdot(p\equiv\alpha)\supset(p\equiv\beta)$.

The argument is the following. Firstly, from 1 and 2, using MP, it follows that $\mathcal{GL}\vdash\boxdot(p\equiv\beta)\supset(p\equiv\alpha)$, from which, using Theorem 2 of 4.1.1, it follows that $\mathcal{GL}\vdash\boxdot(p\equiv\beta)\supset\boxdot(p\equiv\alpha)$. Similarly, from 2 it follows that $\mathcal{GL}\vdash\boxdot(p\equiv\alpha)\supset\boxdot(p\equiv\beta)$. Whence, by PL, it follows that $\mathcal{GL}\vdash\boxdot(p\equiv\alpha)\equiv\boxdot(p\equiv\beta)$, i.e., just the Fixed Point Theorem.

So, given the proof of 1, all that remains to be proved is 2. But the proof of 2, in turn, requests the following results: the Craig Interpolation Lemma, the Beth Definability Theorem and the Theorem (a result due to Bernardi). In fact, Craig's result is needed only for the proof of Beth's Theorem, and the Fixed Point Theorem follows directly from Beth's and Bernardi's results. Let us detail.

4.2.3.1. Craig Interpolation Lemma for $\mathcal{GL}$

As we saw (comp. Ch. 1, 2.9, Remark 2), by Interpolation Theorem for PL the following result holds: if $\models\alpha\supset\beta$, then $\alpha\supset\beta$ has an interpolant, i.e., there is a formula γ all of whose variables occur in both α and β such that $\models\alpha\supset\gamma$ and $\models\gamma\supset\beta$. By soundness and completeness of PL^{ax}, this result evidently holds in its syntactic form. This also holds for the modal system $\mathcal{GL}$.

Craig Interpolation Lemma for GL. *If* $\mathcal{GL}\vdash\alpha\supset\beta$, *then* $\alpha\supset\beta$ *has an interpolant* γ, *i.e.,* $\mathcal{GL}\vdash\alpha\supset\gamma$ *and* $\mathcal{GL}\vdash\gamma\supset\beta$ (*with all variables of* γ *occurring in both* α *and* β).

Proof[49] (outline). Let S^0 be the set of formulas all of whose variables occur

[48] For a semantic proof of this result (due to Goldfarb), comp. G. Boolos [1993], 107-108.

[49] The proof of this theorem is due to G. Boolos (comp. G. Boolos [1979], Ch. 14) and C. Smorynski (an elegant exposition of Smorynski's proof is given in G. Boolos [1993], 118-

in α, let S^1 be the set of formulas all of whose variables occur in β. Let $S^{\cap} = S^0 \cap S^1$. Let $Set = \{\delta$ is a subformula of α or of $\beta\}$, let $S^* = \{\neg\delta \mid \delta \in Set\}$. For a set $S \subseteq Set \cup Set^*$ let $S_0 = S \cap S^0$ and $S_1 = S \cap S^1$. Then, evidently, $S = S_0 \cup S_1$. A formula γ *separates* a set $S \subseteq Set \cup Set^*$ iff $\gamma \in S^{\cap}$, $\mathcal{GL} \vdash Conj(S_0) \supset \gamma$ and $\mathcal{GL} \vdash Conj(S_1) \supset \neg\gamma$. S is called *inseparable* iff there exists no formula γ separating S. As can be seen, if S is consistent, then S is inseparable, and if S is inseparable, then both sets S_0 and S_1 are consistent.

As we saw, the system $\mathcal{GL}$ is noncanonical (comp. 4.1.3.3). So for proving its completeness we cannot use the technique of canonical models. Instead, for this proof we successfully used the technique of finite models (comp. 4.1.3.4 and 4.1.3.5). Similar to those consideration let us define the following finite model $M = \langle W,R,V \rangle$:

W = the set of all maximal and inseparable sets (i.e., a set $w \in W$ is inseparable and for any formula $\delta \in Set$: $\delta \in w$ or $\neg\delta \in w$.

R: wRw' iff for any $\Box\varepsilon$: if $\Box\varepsilon \in w$, then $\Box\varepsilon, \varepsilon \in w'$, and there is a formula $\Box\varepsilon \in w'$ such that $\Box\varepsilon \notin w$.

V: $p \in w$ iff $w \models p$.

This definition does guarantee that M is finite, transitive and irreflexive.

The proof of Craig Interpolation Lemma for $\mathcal{GL}$ is based on the following results:

Rez1. Let S be an inseparable set and $\delta \in Set$. Then either $S \cup \{\delta\}$ is inseparable or $S \cup \{\neg\delta\}$ is inseparable.[50]

By *Rez1* every inseparable set S can be extended to a *maximal* and inseparable set w; i.e., for every $\delta \in Set$, either $\delta \in w$ or $\neg\delta \in w$.

Rez2. Let $M = \langle W,R,V \rangle$. Then for every $\delta \in Set$ and $w \in W$:

$w \models \delta$ iff $\delta \in w$.[51]

Finally, the proof of interpolation lemma runs as follows:

Suppose for *reductio* that $\mathcal{GL} \vdash \alpha \supset \beta$ and that the implication $\alpha \supset \beta$ has no interpolant. I.e., there is no formula γ all of whose variables occur in both α and β such that $\mathcal{GL} \vdash \alpha \supset \gamma$ and $\mathcal{GL} \vdash \gamma \supset \beta$. Then the set $\{\alpha, \neg\beta\}$ is inseparable

---

121). Based on these results we only sketch such a proof.

[50] This is *Lemma 9* in Boolos version of Smorynski's original result (cf. G. Boolos [1993], 119).This is the analogue of Lemma 1(Sect.4.1.3.1).

[51] This is *Lemma 10* of Boolos version (cf. G. Boolos [1993], 120). It is the analogue of 4.1.3.1, Lemma 5.

(otherwise there were a formula γ such that $\mathcal{GL}\vdash\alpha\supset\gamma$ and $\mathcal{GL}\vdash\neg\beta\supset\neg\gamma$, equivalently $\mathcal{GL}\vdash\gamma\supset\beta$, and then γ would be an interpolant for $\alpha\supset\beta$ (contra our supposition). Now, since $\{\alpha,\neg\beta\}$ is inseparable, it follows, by *Rez*1, that $\{\alpha,\neg\beta\}\subseteq w$, where w is an inseparable and maximal set from W. Then, by *Rez*2, $w\models\alpha$, $w\models\neg\beta$ and then $w\not\models\beta$. Whence $w\not\models\alpha\supset\beta$. Therefore, by soundness of $\mathcal{GL}$, $\mathcal{GL}\nvdash\alpha\supset\beta$, contradicting the hypothesis.

4.2.3.2. Beth Definability Theorem for $\mathcal{GL}$

The Beth Definability Theorem for $\mathcal{GL}$ can be easily derived from its form for FOL (comp. Ch. 2, 2.6), by a "propositional" reading of this form and by considering its syntactic proof in FOL^{ax}. Both forms of this theorem (for FOL^{ax} and for $\mathcal{GL}$) are essentially based on the Craig Interpolation Lemma.

Beth Definability Theorem. *Let* p *and* q *be two distinct propositional variables, let* δ *and* δ^* *be two formulas of* L_{PML} *different from each other only by the fact that the occurrences of* p *in* δ *are exactly the occurrence of* q *in* δ^*. *Suppose that* $\mathcal{GL}\vdash(\delta\wedge\delta^*)\supset(p\equiv q)$. *Then there is a formula* β *of* L_{PML} *whose all propositional variables are contained in* δ *and distinct from* p. *Then* $GL\vdash\delta\supset(p\equiv\beta)$.

Proof.[52] We have the following derivations:

(1) $\mathcal{GL}\vdash(\delta\wedge\delta^*)\supset(p\equiv q)$; by hyp.

(2) $\mathcal{GL}\vdash((\delta\wedge\delta^*)\supset(p\equiv q))\supset((\delta\wedge p)\supset(\delta^*\supset q))$; by Rule_p (cf. Ch. 2, 3.2.1)

(3) $\mathcal{GL}\vdash(\delta\wedge p)\supset(\delta^*\supset q)$; (1), (2), MP.

Let us firstly observe that in the formula $\delta\wedge p$ does not occur q, and in $\delta^*\supset q$ does not occur p. So, the variables common to both formulas are exactly the variables of δ distinct from p.

Now, since the implication from (3) is provable in $\mathcal{GL}$, by Craig's Interpolation Lemma for $\mathcal{GL}$ it follows that it has an interpolant, i.e., there is a formula γ of L_{PML} whose all variables are contained in δ and distinct from p (and then all variables of γ are contained in both $\delta\wedge p$ and $\delta^*\supset q$) such that

(4) $\mathcal{GL}\vdash(\delta\wedge p)\supset\gamma$

[52] This proof is mimicking that given in Ch. 2, 2.6.

(5) $\mathcal{GL}\vdash\gamma\supset(\delta^*\supset q)$, and then
(6) $\mathcal{GL}\vdash\delta^*\supset(\gamma\supset q)$; (5) by Rule$_p$, MP
(7) $\mathcal{GL}\vdash\delta\supset(p\supset\gamma)$; (4) by Rule$_p$, MP
(8) $\mathcal{GL}\vdash\delta\supset(\gamma\supset p)$; (6), p/q
(9) $\mathcal{GL}\vdash\delta\supset(p\equiv\gamma)$; (7), (8), Rule$_p$, MP.

In order to give all ingredients needed for the proof of Fixed Point Theorem another result is also requested: a proof of the *uniqueness* of fixed points. And this is given by the theorem below, but first time let us take a definition.

Definition. *Let* $B(q_1,...,q_n)$ *be a formula containing* $q_1,...,q_n$ *distinct in pairs and not containing* p. *Let* $\alpha(p) = B(\Box\gamma_1(p),...,\Box\gamma_n(p))$, *where* $\Box\gamma_i(p)$ *are all subformulas of this form of* $\alpha(p)$. *This representation of* $\alpha(p)$ *is called* ***decomposition*** *of* $\alpha(p)$ (*abrev.* ***Decomp***), *where* $\Box\gamma_i(p)$ *are its components and each contains* p.

Evidently, if $\alpha(p)$ is modalized in p, it is decomposable.

4.2.3.3. Theorem (Bernardi)

Let α *be a formula of* L_{MPL} *modalized in* p *and not containing* q, *let* α^* *be a formula of* L_{MPL} *which differs from* α *only by the fact that the occurrences of* q *in* α^* *are exactly the occurrences of* p *in* α. *Then*

$\mathcal{GL}\vdash(\boxdot(p\equiv\alpha)\wedge\boxdot(q\equiv\alpha^*))\supset(p\equiv q)$.

Proof.[53] The idea of the proof is the following: to show firstly that

(*) $\mathcal{GL}\vdash\Box(p\equiv q)\supset(p\equiv q)$,

and then to insert the formula from (*) as the consequent in an implication whose antecedent is just the antecedent of the formula to be proved in $\mathcal{GL}$, in order to derive (*via* the axiom **W**) the result we are looking for. Let us detail.

Since α is modalized in p, let us consider its decomposition $\alpha(p) = B(\Box\gamma_1(p),...,\Box\gamma_n(p))$, where $B(q_1,...,q_n)$ does not contain p. Then we have the following derivations:

(1) $\mathcal{GL}\vdash\Box(p\equiv q)\supset(\Box(\gamma_i(p)\equiv\gamma_i(q))$; by 4.1.1, Sec. Subst. Th, Coroll.

---

[53] This is Smorynski's proof of Bernardi's *syntactic* proof of the uniqueness of fixed points; cf. Smorynski [1985], 76-77. For a *semantic* proof of this theorem, comp. Boolos [1993], 122, Lemma 11.

(2) $\mathcal{GL}\vdash\Box(\gamma_i(p)\equiv\gamma_i(q))\supset(\Box\gamma_i(p)\equiv\Box\gamma_i(q))$; by 4.1.1, **K**, PL

(3) $\mathcal{GL}\vdash\Box(p\equiv q)\supset(\Box\gamma_i(p)\equiv\Box\gamma_i(q))$; (1), (2), PL

(4) $\mathcal{GL}\vdash\Box(p\equiv q)\supset\Box(\Box\gamma_i(p)\equiv\Box\gamma_i(q))$; (3), by 4.1.1, Th3

(5) $\mathcal{GL}\vdash\Box(p\equiv q)\supset\boxdot(\Box\gamma_i(p)\equiv\Box\gamma_i(q))$; (3), (4), PL

(6) $\mathcal{GL}\vdash\boxdot(\Box\gamma_i(p)\equiv\Box\gamma_i(q))\supset(B(\Box\gamma_i(p))\equiv B(\Box(\gamma_i(q)))$;
by 4.1.1 Sec. Subst. Th.

(7) $\mathcal{GL}\vdash\Box(p\equiv q)\supset(B(\Box\gamma_i(p))\equiv B(\Box(\gamma_i(q)))$; (5), (6), PL

(8) $\mathcal{GL}\vdash\Box(p\equiv q)\supset(\alpha(p)\equiv\alpha(q))$; (7) by Decomp.

And then (*) $\mathcal{GL}\vdash\Box(p\equiv q)\supset(p\equiv q)$.

On the other hand, we have the following derivations:

(1) $\mathcal{GL}\vdash(\boxdot(p\equiv\alpha(p))\wedge\boxdot(q\equiv\alpha(q)))\supset(\Box(p\equiv q)\supset(p\equiv q))$; using (*) and PL

(2) $\mathcal{GL}\vdash\Box(\Box(p\equiv q)\supset(p\equiv q))$; by (*) and the rule N

(3) $\mathcal{GL}\vdash\Box(\Box(p\equiv q)\supset(p\equiv q))\supset\Box(p\equiv q)$; by **W**, Subst.

(4) $\mathcal{GL}\vdash\Box(p\equiv q)$; (2), (3), MP

(5) $\mathcal{GL}\vdash(\boxdot(p\equiv\alpha(p))\wedge\boxdot(q\equiv\alpha(q)))\supset(p\equiv q)$; (1), (4), PL

Now, *via* Theorem (Bernardi) and Beth Definability Theorem, the Fixed Point Theorem results in the following way. Let p and q be two distinct variables, let α be a formula modalized in p and not containing q and α^* be a formula differing from α only by the fact that the occurrences of q in α^* are exactly the occurrences of p in α (this is just the hypothesis of Bernardi's result). Then by Theorem (Bernardi) the following holds

$$\mathcal{GL}\vdash(\boxdot(p\equiv\alpha)\wedge\boxdot(q\equiv\alpha^*))\supset(p\equiv q).$$

Now, if we take $\delta=\boxdot(p\equiv\alpha)$ and $\delta^*=\boxdot(q\equiv\alpha^*)$, then the hypothesis of Beth's Definability Theorem (i.e., the construction of δ and δ^* and that $\mathcal{GL}\vdash(\delta\wedge\delta^*)\supset(p\equiv q)$) is satisfied. Whence, by this theorem, it follows that there is a formula β whose all variables are contained in δ (i.e., in $\boxdot(p\equiv\alpha)$, and then in α) and are distinct from p such that $\mathcal{GL}\vdash\boxdot(p\equiv\alpha)\supset(p\equiv\beta)$. For the rest of the argument, see the end of Preliminary (p.314).

Comments

1. From the above form of the theorem concerning the uniqueness of fixed points, another form can be derived.

Corollary. *Let* α *and* α^* *be formulas as in the Theorem* (*Bernardi*). *Then*

$$\mathcal{GL}\vdash(\Box(p\equiv\alpha)\wedge\Box(q\equiv\alpha^*))\supset\Box(p\equiv q)$$

Proof.

(1) $\mathcal{GL}\vdash(\boxdot(p\equiv\alpha)\wedge\boxdot(q\equiv\alpha^*))\supset(p\equiv q)$; Th. (Bernardi)

(2) $\mathcal{GL}\vdash\boxdot((p\equiv\alpha)\wedge(q\equiv\alpha^*))\supset(p\equiv q)$; (1), 4.1.1, Th. 4 (d)

(3) $\mathcal{GL}\vdash\Box((p\equiv\alpha)\wedge(q\equiv\alpha^*))\supset\Box(p\equiv q)$; (2), 4.1.1, Th. 2 (a)

(4) $\mathcal{GL}\vdash(\Box(p\equiv\alpha)\wedge\Box(q\equiv\alpha^*))\supset\Box(p\equiv q)$; (3), 4.1.1, $\mathbf{K_1}$

By 4.2.2, Theorem 1, this does imply that

(a) $PA^{ax}\vdash[(\Box(p\equiv\alpha)\wedge\Box(q\equiv\alpha^*))\supset\Box(p\equiv q)]^*$, where * is an arbitrary realization.[54]

Let us take the following translation of the modal formula in the square brackets of (a): $p^* = S_1$, $q^* = S_2$, $(\alpha)^* = \neg Bew(x)$. Then

(b) $PA^{ax}\vdash(Bew(\ulcorner S_1\equiv\neg Bew(\ulcorner S_1\urcorner)\urcorner)\wedge Bew(\ulcorner S_2\equiv\neg Bew(\ulcorner S_2\urcorner)\urcorner))$
$\supset Bew(\ulcorner S_1\equiv S_2\urcorner)$.

As can be seen, S_1 and S_2 are two fixed points of the formula $\neg Bew(x)$. Informally, what (b) states is the following thing: if $PA^{ax}\vdash S_1\equiv\neg Bew(\ulcorner S_1\urcorner)$ and $PA^{ax}\vdash S_2\equiv\neg Bew(\ulcorner S_2\urcorner)$, then $PA^{ax}\vdash S_1\equiv S_2$; i.e., any two fixed points of the formula $\neg Bew(x)$ are (provably) equivalent in PA^{ax}. And then since the Gödel's sentence G is a fixed point of $\neg Bew(x)$ and by the first incompleteness theorem $PA^{ax}\nvdash G$ (if PA^{ax} is consistent), it follows that no fixed point of $\neg Bew(x)$ is provable in PA^{ax}. And, as we know, since $PA^{ax}\vdash Con(PA)\equiv G$ it follows that (under the same assumption of consistency of PA^{ax}) $PA^{ax}\nvdash Con(PA)$ (and this is the second Gödel's incompleteness theorem).

2. The proof of the full Fixed Point Theorem, given above, is only one of the different types of proofs for this theorem. In his book[55] G. Boolos gives two another proofs, one for the special case in which the modal formula $\alpha(p)$, modalized in p, contains only one variable p (due to Bernardi and Smorynski), a case relevant for explaining a lot of questions concerning the meaning of self-referential arithmetical sentences,[56] and a proof of the general fixed point theorem (due to G. Sambin and Lisa Reidhaar-Olson).[57] And a second proof of the full fixed point theorem (due to Z. Gleit) intends

[54] To avoid any confusion, the star * in the subformula $\Box(q\equiv\alpha^*)$ has the meaning given in Bernardi's Theorem. Only the second star means "realization".

[55] Cf. G. Boolos [1993], Ch. 8.

[56] As we saw in Sect. 4.2.2. Comp. also G. Boolos [1979], Ch. 9: Calculating truth-values of fixed points.

[57] Comp. also G. Boolos and Jeffrey [1991], Ch. 27.

to show that if $\alpha(p)$ is an arbitrary formula modalized in p, then its fixed point has a modal degree $\leq n$ (where n is the number of subformulas of $\alpha(p)$ of the form $\Box\delta$ (in its decomposition)).

Finally, we mention still another strategy to prove the full theorem of fixed point. It is due to C. Smorynski.[58] Let us outline it.

First of all, the formulation of this theorem in Smorynski's paper.

Fixed point theorem.[59] *Let* $\alpha(p)$ *be a formula modalized in* p. *Then there is a formula* β *containing all the variables of* $\alpha(p)$ *excepting* p *such that*

1. $\mathcal{GL}\vdash \boxdot(p\equiv\alpha(p))\supset(p\equiv\beta)$
2. $\mathcal{GL}\vdash \beta\equiv\alpha(\beta)$.

Since there is a theorem on the uniqueness of fixed points (as we saw above by Bernardi's proof), the proof of fixed point theorem may be restricted to the proof of 2. (Since from 2, by N and the definition of $\boxdot$, it follows that $\mathcal{GL}\vdash \boxdot(\beta\equiv\alpha(\beta))$. Whence, by uniqueness of fixed points, *via* PL, it follows 1). And the proof of 2 requests the result of a lemma and its corollary. So, let us present these results: *Lemma*, its *Corollary* and, finally, the proof of 2.

Lemma. (D. de Jongh, C. Smorynski).[60] $\mathcal{GL}\vdash \Box\gamma(\top)\equiv\Box\gamma(\Box\gamma(\top))$ (*where* $\top$ *denotes logical truth; comp. 4.1.1*).

Proof. (a) $GL\vdash \Box\gamma(\top)\supset\Box\gamma(\Box\gamma(\top))$

(1) $\mathcal{GL}\vdash \Box\gamma(\top)\supset(\top\equiv\Box\gamma(\top))$; by PL

(2) $\mathcal{GL}\vdash \Box\gamma(\top)\supset\Box(\top\equiv\Box\gamma(\top))$; (1) by 4.1.1, Theorem 3

(3) $\mathcal{GL}\vdash \Box\gamma(\top)\supset\boxdot(\top\equiv\Box\gamma(\top))$; (1), (2), PL

(4) $\mathcal{GL}\vdash \boxdot(\top\equiv\Box\gamma(\top))\supset(\Box\gamma(\top)\equiv\Box\gamma(\Box\gamma(\top)))$; by 4.1.1, Sec. Subst. Th.

(5) $\mathcal{GL}\vdash \Box\gamma(\top)\supset(\Box\gamma(\top)\supset\Box\gamma(\Box\gamma(\top))$; (3), (4), PL

(6) $\mathcal{GL}\vdash \Box\gamma(\top)\supset(\Box\gamma(\Box\gamma(\top))$; (5), PL

(b) $\mathcal{GL}\vdash \Box\gamma(\Box\gamma(\top))\supset\Box\gamma(\top)$

---

[58] Cf. C. Smorynski [1985], Ch. 1, Sect. 3.

[59] Called by the author "Explicit definability theorem" (Th. 3.5, 79). Actually, the explicit definability in $\mathcal{GL}$ of fixed points is the problem of *existence* of fixed points in $\mathcal{GL}$.

[60] Cf. C. Smorynski [1985], 78, Lemma 3.2, A semantic version of this lemma is given by Lemma 3.18 of [1985], 125.

(1) $\mathcal{GL}\vdash\Box\gamma(\top)\supset\boxdot(\top\equiv\Box\gamma(\top))$; is (3) of the preceding proof

(2) $\mathcal{GL}\vdash\boxdot(\top\equiv\Box\gamma(\top))\supset(\gamma(\top)\equiv\gamma(\Box\gamma(\top))$; by 4.1.1, Sec. Subst. Th.

(3) $\mathcal{GL}\vdash\Box\gamma(\top)\supset(\gamma(\Box\gamma(\top))\equiv\gamma(\top))$; (1), (2), PL (and the symmetry of "$\equiv$")

(4) $\mathcal{GL}\vdash\Box\gamma(\top)\supset(\gamma(\Box\gamma(\top))\supset\gamma(\top))$; (3), PL

(5) $\mathcal{GL}\vdash\gamma(\Box\gamma(\top))\supset(\Box\gamma(\top)\supset\gamma(\top))$; (4), PL

(6) $\mathcal{GL}\vdash\Box\gamma(\Box\gamma(\top))\supset\Box(\Box\gamma(\top)\supset\gamma(\top))$; (5), 4.1.1, Deriv.1

(7) $\mathcal{GL}\vdash\Box(\Box\gamma(\top)\supset\gamma(\top))\supset\Box\gamma(\top))$; by Ax. **W**, Subst.

(8) $\mathcal{GL}\vdash\Box\gamma(\Box\gamma(\top)\supset\Box\gamma(\top))$; (6), (7), PL

The Lemma follows from (a) and (b) by PL.

Corollary. *Let* $\alpha(p) = B(\Box\gamma(p))$. *Then* $\mathcal{GL}\vdash\alpha B(\top)\equiv\alpha(\alpha B(\top))$.

Proof.

(1) $\mathcal{GL}\vdash\Box\gamma B(\top)\equiv\Box\gamma B(\Box\gamma B(\top))$; by Lemma

(2) $\mathcal{GL}\vdash\Box(\Box\gamma B(\top)\equiv\Box\gamma B(\Box\gamma B(\top)))$; (1), N

(3) $\mathcal{GL}\vdash\boxdot(\Box\gamma B(\top))\equiv\Box\gamma B(\Box\gamma B(\top)))$; (1), (2), by PL and Def. of $\boxdot$

(4) $\mathcal{GL}\vdash B(\Box\gamma B(\top))\equiv B(\Box\gamma B(\Box\gamma B(\top)))$; (3) by 4.1.1, Sec. Subst. Th.

(5) $\mathcal{GL}\vdash\alpha B(\top)\equiv\alpha(\alpha B(\top))$; (4) by construction of $\alpha(p)$.

Now, using the corollary of the above lemma, the Smorynski's proof of 2: $\mathcal{GL}\vdash\beta\equiv\alpha(\beta)$ runs as follows. The proof is by induction on n: the number of components in a decomposition of $\alpha(p)$, modalized in p. For the induction step let $\alpha(p) = B(\Box\gamma_1(p),...,\Box\gamma_n(p))$. Let us take the n-component of $\alpha(p)$ and replace all of its occurrences in $\alpha(p)$ by $\Box\gamma_n(q)$, where q does not occur in $\alpha(p)$. Let the resulting formula be α^*, i.e.,

$$\alpha^*(p,q) = B(\Box\gamma_1(p),...,\Box\gamma_{n-1}(p),\Box\gamma_n(q)).$$

And then $\alpha^*(p)$ has only $n-1$ components, for which, by induction hypothesis, there exists a fixed point $\beta^* = \beta^*(q)$, such that

(1) $\mathcal{GL}\vdash\beta^*(q)\equiv B(\Box\gamma_1(\beta^*),...,\Box\gamma_{n-1}(\beta^*),\Box\gamma_n(q))$

Let us consider that all occurrences of q in $\beta^*(q)$ are in occurrences of the n-component $\Box\gamma_n(q)$. So, by Corollary a formula β can be found such that

(2) $\mathcal{GL}\vdash\beta\equiv\beta^*(\beta)$, i.e.,

(3) $\mathcal{GL}\vdash\beta\equiv B(\Box\gamma_1(\beta^*),...,\Box\gamma_{n-1}(\beta^*),\Box\gamma_n(\beta))$; (1) Subst. β/q

(4) $\mathcal{GL} \vdash \beta \equiv B(\Box\gamma_1(\beta),\ldots,\Box\gamma_{n-1}(\beta),\Box\gamma_n(\beta))$; (3) Subst. $\beta/\beta^*(\beta)$

(5) $\mathcal{GL} \vdash \beta \equiv \alpha(\beta)$; (4) by construction of $\alpha(p)$.

This corollary allows us the calculation of fixed points. Let us take some examples.

(a) For Gödel-type sentences: $\alpha(p) = \neg\Box p$. Here $B(q) = \neg q$ and $\gamma(p) = p$. The fixed point of $\alpha(p)$ is $\alpha(B(\top))$, i.e., $\neg\Box\neg\top$, and then, equivalently, $\neg\Box\bot$.

(b) For Henkin-type sentences: $\alpha(p) = \Box p$. Then $B(q) = q$, $\gamma(p) = p$. And then $\beta(\top) = \top$, so $\alpha(B(T)) = \Box\top$, equivalently $\top$.

(c) For Jeroslow-type sentences: $\alpha(p) = \Box\neg p$. The fixed point of $\alpha(p)$ is $\beta = \Box\bot$ (detail!).

(d) For Rogers-type sentences: $\alpha(p) = \neg\Box\neg p$. The fixed point of $\alpha(p)$ is $\beta = \bot$ (detail!).

(e) For Löb-type sentences: $\alpha(p,q) = \Box p \supset q$. $\beta = \Box q \supset q$ (detail).

(f) For Kreisel-type sentences: $\alpha(p,q) = \Box(p \supset q)$. The fixed point of $\alpha(p,q)$ is $\beta = \Box q$.[61]

The investigations on this topic, *modal logic of provability*, are very extensive. Besides the names just mentioned we also mentioned some other important names, e.g. A. Avron [1984], F. Montagna [1984], R. Solovay [1976], S. Valentini [1983], S.N. Artemov and L.D. Beklemishev [2005], S.R. Buss [1998], G. Japaridze, Dick de Jongh [1998], D.M. Gabbay and L. Maksimova [2005].

[61] For calculation of the fixed points, comp. G. Boolos [1993] Ch. 8, [1979] Ch. 9 and C. Smorynski [1985] Ch. 1.

Bibliography

Ackermann, W. [1954] *Solvable Cases of the Decision Problem*, Amsterdam.

Addison, J.W., L. Henkin, A. Tarski (eds.) [1965] *The Theory of Models*, North-Holland.

Anderson, A.R. and N. Belnap [1959] "A simple treatment of truth functions", *Journal of Symbolic Logic*, 24, 301-302.

Artemov, S.N., Lev D. Beklemishev [2005] "Provability logic", in vol. Gabbay, D.M. and F. Guenthner (eds.) [2005], 189-360 (with an extensive bibliography on this topic).

Avron, A. [1984] "On modal systems having arithmetical interpretations", *Journal of Symbolic Logic*, 49(3), 935-942.

Balaguer, M. [1998] *Platonism and Anti-Platonism in Mathematics*, New York, Oxford University Press.

Barwise, J. (ed.) [1977] *Handbook of Mathematical Logic*, North-Holland.

Benacerraf, P. [1967] "God, the Devil and Gödel", *The Monist*, 51, 9-32.

Bernardi, C. [1975] "The fixed point theorem for diagonalizable algebras", *Studia Logica*, 34, 239-251.

[1976] "The uniqueness of the fixed-point in every diagonalizable algebra", *Studia Logica*, 35, 335-343.

Berto, F. [2009] "The Gödel Paradox and Wittgenstein's Reasons", *Philosophia Mathematica* (III), 17, 208-219.

Beth, E. [1951] "A Topological Proof of the Theorem of Löwenheim-Skolem-Gödel", *Indag. Math.*, 13, 436-444.

[1953] "Some consequences of the theorem of Löwenheim-Skolem-Gödel-Malcev", *Indag. Math.*, 15, 66-71.

[1959] *The Foundations of Mathematics*, Amsterdam.

Boolos, G.S. [1979] *The Unprovability of Consistency: An Essay in Modal Logic*, Cambridge University Press.

[1993] *The Logic of Provability*, Cambridge University Press.

Boolos, G.S., J.P. Burgess, R.C. Jeffrey [2002] *Computability and Logic*, 4th ed., Cambridge University Press.

Boolos, G.S. and R. Jeffrey [1991] *Computability and Logic*, 3rd ed., Cambridge University Press.

Boolos, G.S. and G. Sambin [1991] "Provability: the emergence of a mathematical modality", *Studia Logica*, 50, 1-23.
Brouwer, L.E.J., E.W. Beth, A. Heyting (eds.) [1971] *Studies in Logic and the Foundations of Mathematics*, North-Holland Publishing Company, Amsterdam.
Buss, S.R. (ed.) [1998] *Handbook of Proof Theory*, Elsevier Science B.V.
Carnap, R. [1934] *Logische Syntax der Sprache*, Springer-Verlag, Wien GmbH.
Chabert, J.L. [1999] *A History of Algorithms*, Berlin, Springer.
Chang, C. and R.C. Lee [1973] *Symbolic Logic and Mechanical Theorem Proving*, New York, Academic Press.
Church, A. [1936(a)] "A note on the Entscheidungsproblem", *Journal of Symbolic Logic*, vol. 1, 40-41; with a Correction, 101-102.
[1936(b)] "An unsolvable problem of elementary number theory", *Americal Journal of Mathematics*, 58, 345-363.
[1944] *Introduction to Mathematical Logic*, Princeton.
[1956] *Introduction to Mathematical Logic, I*, Princeton.
Copeland, B.J. (ed.) [2004] *The Essential Turing*, Oxford, Clarendon Press.
Copeland, B.J., C.J. Posy, O. Shagrir (eds.) [2013], *Computability*, The MIT Press, Cambridge, Massachusetts.
Couturat, B. [1903] *Opuscules et fragments inédites de Leibniz*, Paris.
Craig, W. [1957] "Linear reasoning. A new form of the Herbrand-Gentzen theorem", *Journal of Symbolic Logic*, 22, 250-268.
Davis, M. [1958] *Computability and Unsolvability*, McGraw-Hill, New York.
Davis, M. (ed.) [2004] *The Undecidable*, Dover Publ., New York.
Dean, W. [2014] "Arithmetical Reflection and the Provability of Soundness", *Philosophia Mathematica*, 24, 31-64.
Dekker, J. [1955] "Productive sets", *Trans. Amer. Math. Soc.,* Vol. 78, 129-149.
Ebbinghaus, H.D.; J. Flum and W. Thomas [1994] *Mathematical Logic*, Springer-Verlag, New York.
Égré, P. [2005] "The Knower Paradox in the Light of Provability Interpretations of Modal Logic", *Journal of Logic, Language and Information*, 14, 13-48
Ehrenfeucht, A. and S. Feferman [1960] "Representability of recursively enumerable sets in formal theories", *Archiv für Mathematische Logik und Grundlagenforschung*, 5, 37-41.

Feferman, S. [1960] "Arithmetization of metamathematics in a general setting", *Fundamenta Mathematicae*, 49, 35-92.
[1962] "Transfinite recursive progressions of axiomatic theories", *Journal of Symbolic Logic*, 27, 259-316.
[1984] "Kurt Gödel: Conviction and caution", *Philosophia Naturalis*, 21, 546-562; also in Feferman [1998], 150-164.
[1991] "Reflecting on Incompleteness", *Journal of Symbolic Logic*, Vol. 56, 1, 1-49.
[1998] *In the Light of Reason*, New York, Oxford University Press.
Fitting, M. [1990] *First-Order Logic and Automated Theorem Proving*, Springer-Verlag.
[1993] "Basic Modal Logic", in vol. Gabbay, Hogger, Robinson (eds.) [1993], 368-449.
Fitting, M. and R.L. Mendelson [1998] *First-Order Modal Logic*, Kluwer Acad. Publ.
Floyd, J., A. Bokulich (eds.) [2017] *Philosophical Explorations of the Legacy of Alan Turing*, Springer.
Franzén, T. [2005] *Gödel's Theorem: An Incomplete Guide to its Use and Abuse*, Wellesley, MA: A.K. Peters.
Frege, G. [1879] *Begriffsschrift*, Louis Nebert, Halle.
[1980] *Gottlob Freges Briefwechsel*, Felix Meiner Verlag, Hamburg.
Friedman, H. and M. Sheard [1987], "An Axiomatic Approach to Self-Referential Truth", *Annals of Pure and Applied Logic*, 33, 1-21.
Gabbay, D.M. and F. Guenthner (eds.) [2005] *Handbook of Philosophical Logic*, 2nd ed., vol. 13, Springer.
Gabbay, D.M., C.J. Hogger, J.A. Robinson (eds.) [1993] *Handbook of Logic in Artificial Intelligence and Logic Programming*, vol. 1, Oxford University Press.
Gabbay, D. and L. Maksimova [2005] *Interpolation and Definability* (*Modal and Intuitionistic Logics*), Clarendon Press, Oxford.
Gallier, J.H. [2003] *Logic for Computer Science. Foundations of Automatic Theorem Proving*, University of Pensylvania.
Gentzen, G. [1934] "Untersuchungen über das logische Schließen", *Mathematische Zeitschrift*, 39, 176-210, 405-431.
Gleit, Z. and W. Goldfarb [1990] "Characters and fixed points in provability logic", *Notre Dame Journal of Formal Logic*, 31, 26-36.

Goldblatt, R. [1978] "Arithmetical necessity, provability and intuitionistic logic", *Theoria*, 44, 38-46.

Gödel, K. [1930] "Die Vollständigkeit der Axiome des logischen Funcktionenkalküls", *Monatshefte für Mathematik und Physik*, 37, 349-360; bilingual (Germ.-Engl.) in K. Gödel [1986], 102-123.

[1931] "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I", *Monatshefte für Mathematik und Physik,* 38, 173-198; a bilingual (Germ.-Engl.) in K. Gödel [1986], 144-195.

[1934] "On undecidable propositions of formal mathematical systems", in K. Gödel [1986], 346-371.

[1986] *Collected Works*, vol. I, Oxford University Press, 1986.

Grelling, K., L. Nelson [1908], "Bemerkungen zu den Paradoxien von Russell and Burali-Forti", *Abhandlungen der Fries'schen Schule II*, Göttingen, 301-334.

Hájek, P., Pudlák [1993] *Metamathematics of First-Order Arithmetic*, Springer, Berlin.

Hamilton, A.G. [1978] *Logic for Mathematicians*, Cambridge University Press.

Hasenjaeger, G. [1953] "Eine Bemerkung zu Henkin's Beweis für die Vollständigkeit des Prädikatenkalküls der ersten Stufe", *Journal of the Symbolic Logic*, 18, 42-48.

Hasenjaeger, G., and H. Scholz [1961], *Grundzüge der mathematischen Logik*, Springer-Verlag, Berlin-Göttingen-Heidelberg.

Henkin, L. [1949] "The completeness of the first-order functional calculus", *Journal of Symbolic Logic*, 14, 159-166.

Herbrand, J. [1930] "Recherches sur la théorie de la demonstration", *Travaux de la Soc. des Sci. et des Lettres de Varsovie*, III, vol. 33, 33-160.

Hermes, H. [1961] *Aufzählbarkeit, Entscheidbarkeit*, *Berechenbarkeit*, Springer, Berlin; Engl. transl. with the title *Enumerability, Decidability, Computability*, Springer, Berlin, 1965.

[1963] *Einführung in die mathematische Logik*, B.G. Teubner Verlagsgesellschaft, Stuttgart.

Hilbert, D. and W. Ackermann [1928] *Grundzüge der theoretischen Logik*, Springer, Berlin; Engl. transl. of 2nd ed. [1938], with the title *Principles of Mathematical Logic*, Chelsea Publishing Co., New York, 1950.
[1946] *Grundzüge der theoretischen Logik*, zweite verbesserte Auflage, Dover Publications, New York.
[1972] *Grundzüge der theoretischen Logik*, 6. Aufl., Springer-Verlag, Berlin.
Hilbert, D. and P. Bernays [1934], [1939] *Grundlagen der Mathematik*, vol. I (1934), vol. II (1939), Berlin, Verlag Von Julius Springer.
Hoffmann, D.W. [2013] *Die Gödel'schen Unvollständigkeitssätze*, Springer Spektrum.
Hofstadter, D.R. [1979] *Gödel, Escher, Bach: An Eternal Golden Braid*, Basic Books, Inc., Publishers.
Hughes, G.E. and M.J. Cresswell [1996] *A New Introduction to Modal Logic*, Routledge, London and New York.
Hunter, G. [1973] *METALOGIC* (*An Introduction to the Metatheory of Standard First-Order Logic*), University of California Press.
Japaridze, G. and Dick de Jongh [1998] *The Logic of Provability*, in vol. S.R. Buss [1998], 475-546.
Jeroslow, R.G. [1973] "Redundancies in the Hilbert-Bernays derivability conditions for Gödel's second incompleteness theorem", *Journal of Symbolic Logic*, 38, 359-367.
Kalmár, Laszlo [1935] "Über die Axiomatisierbarkeit des Aussagenkalküs", *Acta Scientiarum Mathematicarum*, vol. 7, 222-243.
Kleene, S.C. [1952] *Introduction to Metamathematics*, Van Nostrand, North-Holland and Noordhoff.
[1966] *Mathematical Logic*, John Wiley & Sons, Inc., New York.
Ketland, J. [1999] "Deflationism and Tarski's Paradise", *Mind*, 108, 69-94.
[2005] "Deflationism and the Gödel Phenomena: Reply to Tennant", *Mind*, 114, 75-88.
Kleene, S.C. [1943] "Recursive predicates and quantifiers", *Trans. Amer. Math. Soc.*, vol. 53, 41-73.
[1952] *Introduction to Metamathematics*, Van Nostrand, Princeton; fourth reprint [1964], North-Holland Publishing Co., Amsterdam, P. Noordhoff N.V., Gröningen.

[1967] *Mathematical Logic*, John Wiley & Sons, Inc., New York.
[1981] "Origins of recursive function theory", *Annals of the History of Computing*, 3, 52-67.
[1987] "Reflections on Church's Thesis", *Notre Dame Journal of Formal Logic*, 28, 490-498.
Kreisel, G. and A. Levy [1968] "Reflection principles and their use for establishing the complexity of axiomatic systems", *Zeitschrift für mathematische Logic und Grundlagen der Mathematik*, 14, 97-142.
Kreisel, G., G. Takeuti [1974] "Formally self-referential propositions in cut-free classical analysis and related systems", *Dissertationes Mathematicae*, 118, 1-50.
Kripke, S. [1959] "A completeness theorem in modal logic", *Journal of Symbolic Logic*, 24, 1-14.
[1963] "Semantical analysis of modal logic I. Normal modal propositional calculi", *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 9, 67-96.
[1963] "Semantical considerations on modal logic", *Acta Philosophica Fennica*, 16, 83-94.
[1965] "Semantical analysis of modal logic II. Non-normal modal propositional calculi", in J.W. Addison, L. Henkin and A. Tarski eds. [1965].
[1996] *Elementary Recursion Theory with Applications to Formal Systems*, (draft).
LaForte, G., P.J. Hayes, K.M. Ford [1998] "Why Gödel's Theorem cannot refute computationalism", *Artificial Inteligence*, 104, 265-286.
Leivant, D. [1981] "On the proof theory of the modal logic for arithmetic provability", *Journal of Symbolic Logic*, 46(3), 531-538.
Lindström, P. [1997] *Aspects of Incompleteness*, Springer-Verlag, Berlin Heidelberg.
[2001] "Penrose's New Argument", *Journal of Philosophical Logic*, 30, 241-250.
Löb, M.H. [1955] "Solution of a problem of Leon Henkin", *Journal of Symbolic Logic*, 20, 115-118.
Löwenheim, L. [1915] "Über Möglichkeiten im Relativkakül", *Mathematische Annalen*, 76, 447-70; Engl. transl. in J.V. Heijenoort [1967], 232-251.

Lukasiewicz, J. and A. Tarski [1930] "Untersuchungen über den Aussangenkalkül", reprinted in A. Tarski [1956], 38-59.
[1952] "A problem concerning provability", *Journal of Symbolic Logic*, 160.
Mancosu, P. (ed.) [1988] *From Brouwer to Hilbert*, Oxford, Oxford University Press.
Manin, Yu I. [2010] *A Course in Mathematical Logic for Mathematicians*, Springer.
Markov [1961] "The theory of algorithms", *Tr. Math. Inst.*, XLII: 375.
Mendelson, E. [1964] *Introduction to Mathematical Logic*, Van Nostrand, Princeton.
Montagna, F. [1984] "The predicate modal logic of provability", *Notre Dame Journal of Formal Logic*, 25.
Montague, R. [1974] "Syntactical treatments of modality, with corollaries on reflection principles and finite axiomatizability", in R. Montague, *Formal Philosophy*, ed. R.H. Thomason, Yale University Press, 1974.
Mostowski, A. [1952] *Sentences undecidable in formalized arithmetic*, North-Holland Publishing Company, Amsterdam.
[1961] "A generalization of the incompleteness theorem", *Fundamenta Mathematicae*, 49, 205-232.
Nelson, R.J. [1987] "Church's Thesis and cognitive sciences", *Notre Dame Journal of Formal Logic*, 28, 581-614.
Paris, J., L.A. Harrington [1977] "The incompleteness theorems", in J. Barwise (ed.) [1977], 1133-1142.
Péter, R. [1951] *Rekursive Funktionen*, Budapest, Akadémiai Kiadó.
Post, E.L. [1921] "Introduction to a general theory of elementary propositions", *Americal Journal of Mathematics*, 43, 163-185.
[1944] "Recursively enumerable sets of positive integers and their decision problems", *Bulletin of the American Mathematical Society*, 50, 284-316.
Putnam, H. [1960] "Minds and machines", in H. Putnam.
[1975] *Philosophical Papers* 2: *Mind, Language and Reality*, Cambridge, Cambridge University Press, 362-385.
[2000] "Nonstandard models and Kripke's proof of the Gödel Theorem", *Notre Dame Journal of Formal Logic*, 41, 53-58.
Quine, W.V.O. [1938] "Completeness of the propositional calculus", *Journal of Symbolic Logic*, 3, 37-40.

[1940] *Mathematical Logic*, Harvard University Press.
[1950] *Methods of Logic*, New York.
[1966] *The Ways of Paradox*, Random.
[1972] *Methods of Logic*, Holt, Rinehart and Winston, Inc.
Rasiowa, H. and R. Sikorski [1951] "A proof of the completeness theorem of Gödel", *Fundamenta Mathematicae*, 37, 193-200.
[1951] "A proof of the completeness theorem", *Fundam. Math.*, 37, 193-200.
[1952] "A proof of the Skolem-Löwenheim theorem", *Fundamenta Mathematicae*, 38, 230-232.
Rautenberg, W. [2006] *A Concise Introduction to Mathematical Logic*, Springer.
Reidhaar-Olson, L. [1990] "A new proof of the fixed-point theorem of provability logic", *Notre Dame Journal of Formal Logic*, 31, 37-43.
Robinson, A. [1951] *On the metamathematics of algebra*, Amsterdam.
Rogers, H. [1967] *Theory of Recursive Functions and Effective Computability*, McGraw-Hill, New York.
Rosser, J.B. [1936] "Extensions of some theorems of Gödel and Church", *Journal of Symbolic Logic*, 1, 87-91.
[1937] "Gödel's Theorems for Non-constructive Logics", *Journal of Symbolic Logic*, vol. 2(3).
[1939] "An informal exposition of proofs of Gödel's Theorem and Church's Theorem", *Journal of Symbolic Logic*, vol. 4, 53-60; reprinted in M. Davis (ed.) [2004], 223-230.
[1953] *Logic for Mathematicians*, McGraw-Hill, New York.
Sambin, G. [1976] "An effective fixed point theorem in intuitionistic diagonalizable algebras", *Studia Logica*, 35, 345-61.
Sambin, G., S. Valentini [1982] "The modal logic of provability: The sequential approach", *Journal of Philosophical Logic*, 11, 311-342.
Scholz, H. and G. Hasenjaeger [1961], *Grundzüge der mathematischen Logik*, Springer-Verlag, Berlin.
Schütte, K. [1934] "Untersuchungen zum Entscheidungsproblem der mathematischen Logic", *Mathematischen Annalen*, 109, 572-603.
[1956] "Ein System des verknüpfenden Schliessens", *Archiv für mathematische Logik und Grundlagenforschung*, 2, 55-67.

[1960] *Beweistheorie*, Springer Verlag, Berlin.
Shapiro, S. [1998] "Truth and proof-through thick and thin", *Journal of Philosophy*, 95, 493-521.
[2002] "Incompleteness and Inconsistency", *Mind*, 111, No. 444, 817-832.
Shepherdson, J. [1961] "Representability of recursively enumerable sets in formal theories", *Archiv für Mathematische Logik und Grundlagenforschung*, 119-127.
Shoenfield, J.R. [1967] *Mathematical Logic*, Addison-Wesley, Reading, Mass.
Skolem, T. [1920] Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theoreme über dichte Mengen", *Vidensk. Skrifter*, 4; in Engl. in vol. Van Heijenoorth (ed.) [1967], 254-268.
Smith, P. [2006] *An Introduction to Gödel's Theorems*, University of Cambridge.
Smorynski, C. [1977] "The incompleteness theorems", in J. Barwise (ed.) [1977], 821-865.
[1981] "Fifty years of self-reference in arithmetic", *Notre Dame Journal of Formal Logic*, 22, 357-374.
[1984] "Modal logic and self-reference", in D. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic*, vol. II, D. Reidel Publishing Company, 1984, 441-495.
[1985] *Self-reference and Modal Logic*, Springer-Verlag.
Smullyan, R.M. [1961] *Theory of Formal Systems*, Princeton.
[1968], [1995] *First-Order Logic*, Dover Publications, Inc. New York.
[1992] *Gödel's Incompleteness Theorems*, Oxford UP, 1992.
[1993] *Recursion Theory for Metamathematics*, Oxford UP.
[1994] *Diagonalization and Self-Reference*, Clarendon Press, Oxford.
Solovay, R. [1976] "Provability interpretations of modal logic", *Israel Journal of Mathematics*, 25, 287-304.
Stoll, R.R. [1963] *Set Theory and Logic*, W.H. Freeman, San Francisco.
Tarksi, A. [1936] "Der Wahrheitsbegriff in den formalisierten Sprachen", *Studia Philos.*, 1, 261-405; also in [1956].
[1956] *Logic, Semantics, Matemathematics*, Clarendon Press, Oxford.

Tarski, A., A. Mostowski, R.M. Robinson [1953], *Undecidable Theories,* North-Holland, Amsterdam.

Tennant, N. [2002] "Deflationism and the Gödel phenomena", *Mind*, 111, 551-582.

[2005] "Deflationism and the Gödel phenomena: Reply to Ketland", *Mind*, 114, 89-96.

Turing, A. [1936-37] "On computable numbers, with an application to the Entscheidungsproblem", *Proc. London Math. Soc.*, 42 (230-265), 43 (544-546).

Valentini, S. [1983] "The modal logic of provability", *The Journal of Philosophical Logic*, 12, 471-476.

Van Heienoort, J. [1967] *From Frege to Gödel: A Source Book in Mathematical Logic*, 1879-1931, Harvard University Press.

Von Kutschera, F. [1967] *Elementare Logik*, Springer-Verlag, Wien, New York.

Wang, H. [1974] *From Mathematics to Philosophy*, Routledge and Kegan Paul.

Wittgenstein, L. [1921] *Logisch-philosophische Abhandlung* (*Annalen der Naturphilosophie*).

Whitehead, A.N. and B. Russell [1910-1913] *Principia Mathematica*, 3 vol., Cambridge.

Wright, C. [1995] "About "The Philosophical Significance of Gödel's Theorem": Some Issues", in C. Wright [1995], *Realism, Meaning and Truth*, Sec. Ed., Blackwell, 321-354.

Index

Virgil Drăghici, Associate Professor at Babes-Bolyai University, Cluj-Napoca, Romania. Fields of interest: *mathematical logic* (First-order Logic, Recursion Theory, K.Gödel, A.Tarski) and *German phenomenology* (Imm.Kant, E.Husserl, M.Heidegger, H.-G.Gadamer). Author of some representative books in these fields, e.g., *Logică și Adevăr. O expunere topo-logică a tematicii la E.Husserl și M.Heidegger,* Ed. Casa Cărții de Știință, Cluj-Napoca, 2003; *LOGICA* (*traditională/clasică/modală*), Ed.EFES, Cluj-Napoca, 2007; *Logica clasică*, Ed. Casa Cărții de Știință, Cluj-Napoca, 2014; and of a great number of scientific articles in both fields (mathematical logic and philosophy), published in the scientific journals and collective books.

He is also a translator of some books, relevant for mathematical logic and philosophy, e.g., W.Stegmüller, *Incompletitudine si Indecidabilitate* [Romanian translation of the German book *Unvollständigkeit und Unentscheidbarkeit*, Springer-Verlag, 1973], Presa Universitară Clujeană, 2011; R.M. Smullyan, *Teoremele gödeliene de incompletitudine* [Romanian translation of Smullyan's book *The Gödel's Incompleteness Theorems*, Oxford University Press, 1992], Ed. Casa Cărții de Știință, 2018; D. Hilbert-W. Ackermann, *Principiile Logicii Matematice* [Romanian translation of the German book *Grundzüge der theoretischen Logik*, zweite, verbesserte Auflage, Dover Publications, New York, 1946] (forthcoming).

Finally, his Habilitation Thesis is “Diagonal Structures in Logical and Philosophical Argumentation” and at present he is PhD supervisor in both fields: *mathematical logic* and *German phenomenology* at Babeș-Bolyai University, Cluj-Napoca, Romania.

ISBN: 978-606-37-1938-7