A METHOD IN PROOFS OF UNDEFINABILITY

WITH APPLICATIONS TO FUNCTIONS IN THE ARITHMETIC OF NATURAL NUMBERS

K. L. DE BOUVÈRE

A METHOD IN PROOFS OF UNDEFINABILITY

A METHOD IN PROOFS OF UNDEFINABILITY

WITH APPLICATIONS TO FUNCTIONS IN THE ARITHMETIC OF NATURAL NUMBERS

ACADEMISCH PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR IN DE WIS- EN NATUURKUNDE AAN DE UNIVERSITEIT VAN AMSTERDAM, OP GEZAG VAN DE RECTOR MAGNIFICUS DR. M. W. WOERDEMAN, HOOGLERAAR IN DE FACULTEIT DER GENEESKUNDE, IN HET OPENBAAR TE VERDEDIGEN IN DE AULA DER UNIVERSITEIT OP WOENSDAG 20 MEI 1959 DES NAMIDDAGS TE VIER UUR

DOOR

KAREL LOUIS DE BOUVÈRE S.C.J.

GEBOREN TE LEIDSCHENDAM

1959
NORTH-HOLLAND PUBLISHING COMPANY
AMSTERDAM

PRINTED IN THE NETHERLANDS

Aan de Nagedachtenis van
mijn Vader

Bij de publicatie van dit proefschrift heb ik de eer mijn oprechte dankbaarheid te betuigen aan mijn leermeester en promotor, dr. E. W. Beth. Hij heeft mijn aandacht gevestigd op de problemen der definieerbaarheid en hij heeft gedurende het gehele onderzoek hoogst waardevolle adviezen verstrekt. Zijn kennis was een schatkamer van informatie, zijn belangstelling een bron van inspiratie.

Gedurende mijn studententijd heb ik het voorrecht gehad te worden onderwezen door vele uitstekende beoefenaren der wetenschap. Het wil mij voorkomen, dat het onderwijs in de wiskunde aan de Universiteit van Amsterdam in de eerste jaren na de tweede wereldoorlog werd gekenmerkt door een ongewoon aantal mutaties in het docenten-corps. Aanvankelijk moge deze omstandigheid een enigszins verwarrende indruk hebben gemaakt, thans acht ik het van onschatbare waarde de lessen te hebben kunnen volgen van de wiskundigen dr. L. E. J. Brouwer, dr. E. M. Bruins, dr. J. G. van der Corput, dr. D. van Dantzig, dr. H. Freudenthal, dr. J. de Groot, dr. A. Heyting, dr. F. Loonstra, dr. J. A. Schouten en dr. B. L. van der Waerden. Ik dank hen ten zeerste voor de vorming, welke ik van hen mocht ontvangen.

Ook het onderricht in de natuurkunde van dr. C. J. Bakker, dr. J. de Boer en dr. A. M. J. F. Michels is voor mij, mede in verband met mijn wijsgerige vorming, van grote waarde geweest.

It is a pleasure to thank also the Professors L. A. Henkin, S. C. Kleene and A. Tarski, who spent some time as guests in Amsterdam and whose lectures it was a delight to attend.

INLEIDING EN SAMENVATTING

In een recente lezing „*Over waarheid in de wiskunde*" wees A. Heyting op de drie voornaamste componenten der klassieke wiskunde. Het is bekend hoe één ervan, de intuitieve component, is geïsoleerd door L. E. J. Brouwer, die de grondslagen heeft gelegd van de intuitionistische wiskunde. Heyting zelf is toonaangevend geweest in de verdere ontwikkeling van het *intuitionisme* en het is vooral te danken aan zijn bijdragen, zowel op het gebied van onderzoek als van onderwijs, dat de eervolle band is bestendigd tussen de intuitionistische wiskunde en het Mathematisch Instituut der Universiteit van Amsterdam. Geenszins echter heeft deze verdienste aanleiding gegeven tot eenzijdigheid in het grondslagenonderwijs aan dit instituut. In dezelfde lezing vermeldde Heyting een andere component der klassieke wiskunde, die elders is uitgegroeid tot de *semantiek*, en hij wees op belangrijke en interessante resultaten uit dit gezichtspunt verkregen. In het mathematisch en logisch onderwijs der Universiteit heeft meer in het bijzonder E. W. Beth deze tak van wetenschap ingevoerd. Het Instituut voor Grondslagenonderzoek heeft daarenboven dit gehele terrein gemakkelijker toegankelijk gemaakt.

Dit proefschrift is een voorbeeld van een semantische methode in de theorie der definitie. Met de methode, die hier wordt ontwikkeld, kan in bepaalde omstandigheden de onafhankelijkheid worden bewezen ener niet-logische constante van andere niet-logische constanten met betrekking tot een gegeven theorie. Deze onafhankelijkheid betekent, dat in het bestek der gegeven theorie de beschouwde constante niet expliciet kan worden gedefinieerd met behulp der andere voorkomende constanten. Tot nu toe had men voor dit soort bewijzen de beschikking over de methode van A. Padoa, daterend uit 1900. A. Tarski gaf in 1934 een theoretische fundering van deze methode, terwijl Beth in 1953 bewees, dat de methode algemeen is, dat wil zeggen: indien (onder bepaalde voorwaarden) met betrekking tot een theorie een niet-logische constante onafhankelijk is van de andere voorkomende niet-logische constanten, dan kan dit worden bewezen met behulp der methode van Padoa. Dat er desondanks ruimte is voor een nieuwe methode, wordt aan de ene kant gerechtvaardigd door theoretische overwegingen, welke een dubbel aspect aanwijzen, en aan de andere kant door de praktijk, waarin het niet altijd mogelijk is gebleken voor bepaalde onafhankelijkheden de geschikte toepassing van Padoa's methode rechtstreeks aan te geven.

In hoofdstuk I wordt het begrip *expliciet definieerbaar* gekarakteriseerd op model-theoretische wijze. Zowel de methode van Padoa als de stelling van Beth spelen een rol bij deze karakterisering. Bovendien wordt er onderscheid gemaakt tussen *ondefinieerbaar* en *essentieel ondefinieerbaar*. Laatstgenoemde term betekent, dat een constante niet op expliciete wijze definieerbaar is, niet alleen met betrekking tot de beschouwde theorie zelf, maar evenzeer met betrekking tot iedere consistente uitbreiding van deze theorie, die dezelfde constanten bevat.

De model-theoretische karakterisering behelst twee methoden om ondefinieerbaarheid te bewijzen: de methode van Padoa en de methode ontwikkeld in hoofdstuk II. Laatstgenoemde geeft aanleiding tot een indirect gebruik, dat analoog is met Tarski's indirecte methode in bewijzen van onbeslisbaarheid. Onder bepaalde voorwaarden stelt de methode ons namelijk in staat vast te stellen, dat een constante essentieel ondefinieerbaar is met betrekking tot een bepaalde, bijvoorbeeld eindig axiomatiseerbare, subtheorie der theorie, waarom het gaat. Hieruit volgt dan, dat de ondefinieerbaarheid zich ook uitstrekt met betrekking tot deze theorie zelf.

In hoofdstuk III wordt deze methode toegepast op enkele rekenkundige functies. De methode stelt ons bijvoorbeeld in staat vast te stellen, dat vermenigvuldiging essentieel niet op expliciete wijze kan worden gedefinieerd met behulp van eenheid en optelling met betrekking tot een eindig axiomatiseerbare subtheorie der geformaliseerde rekenkunde der natuurlijke getallen. Hieruit volgt dan, dat vermenigvuldiging niet expliciet kan worden gedefinieerd met behulp van eenheid en optelling met betrekking tot de geformaliseerde rekenkunde zelf. Dit feit was uit andere hoofde reeds bekend: de theorie met vermenigvuldiging is onbeslisbaar, terwijl de theorie met alleen optelling en eenheid beslisbaar is. Maar het pleit voor de bruikbaarheid van de methode, dat zij ons in staat stelt het bewijs rechtstreeks te leveren.

In hoofdstuk IV wordt teruggegrepen op een andere distinctie, welke in hoofdstuk I is gemaakt, namelijk die tussen *expliciete* *definieerbaarheid* en *definieerbaarheid op grond ener gelijkheid.* Deze distinctie geldt uitsluitend individuele en operationele constanten. Definieerbaarheid op grond van een gelijkheid betreft een definitie bestaande uit twee termen gescheiden door het gelijkteken. Nu is door een resultaat van K. Gödel bekend, dat bijvoorbeeld de hogere functies in de reeks van W. Ackermann arithmetisch zijn, dit wil zeggen expliciet kunnen worden gedefinieerd met behulp van optelling en vermenigvuldiging. Van de andere kant is een definitie door gelijkheid niet mogelijk. Wij passen de methode van hoofdstuk II toe op deze functies om hun expliciete ondefinieerbaarheid te bewijzen met betrekking tot bepaalde subtheorieën der rekenkunde. Deze toepassing blijkt tegelijk een methode op te leveren om de onmogelijkheid van een definitie door gelijkheid te bewijzen met betrekking tot de gehele geformaliseerde rekenkunde.

Het onderzoek is in dit proefschrift beperkt tot theorieën met standaard-formalisering. Dit zijn theorieën geformaliseerd binnen het bestek der logica van de eerste orde met identiteit. Uitbreiding tot andere logische systemen lijkt zeker mogelijk.

PREFACE

This study is a sample of a semantic [1]) method in the theory of definition. Already in 1900 A. Padoa [2]) proposed a semantic method in proofs of undefinability. G. Peano [3]) was the first to recognize the value of this method, followed by A. Tarski, A. Lindenbaum, J. C. C. McKinsey [4]) and others. A new step was made by E. W. Beth [5]) in 1953, when he showed that Padoa's method is a general method, in other words, that whenever (under certain conditions) a non-logical constant is undefinable explicitly with respect to a theory from the other non-logical constants occurring in this theory, there is a way to prove this fact by Padoa's method. At the same time this was a first application of logical methods related to G. Gentzen's formalization to a problem in the theory of models.

In this work the concept of explicit definability is characterized in terms of models. Both Padoa's method and Beth's theorem appear to be tools for this characterization. The study is confined to theories with standard formalization, but the results could be extended. The characterization in terms of models entails a new method in proofs of explicit undefinability which seems to have certain advantages over Padoa's method. A distinction is made between explicit undefinability and essential explicit undefinability. The latter term means that a certain non-logical constant is undefinable explicitly from certain other non-logical constants not only with respect to the theory in question, but also with respect to every consistent extension of this theory. Under certain conditions the method developed in this study enables us to state that a given non-logical constant is essentially undefinable explicitly from the other non-logical constants occurring with respect to a certain, say finitely axiomatizable, subtheory of the theory concerned. This implies that the same non-logical constant is undefin-

---

1) For a brief description of semantic methods, cf. e.g. [8]. The numbers in square brackets refer to *References*.

2) Cf. [13], [14] and [15].

3) Cf. [14] and [15].

4) Cf. [20], [21] and [12].

5) Cf. [2].

able explicitly from the same other non-logical constants with respect to the theory itself. For example, the method enables us to state that multiplication is essentially undefinable explicitly from addition and unity with respect to a finitely axiomatizable subtheory of the formalized arithmetic of natural numbers. It follows that multiplication is undefinable explicitly from unity and addition with respect to the formalized arithmetic of natural numbers. This fact was well-known for indirect reasons: the theory with multiplication is undecidable, whereas the theory with only unity and addition is decidable. But the method provides us with a direct proof of this fact, at the same time suggesting its own usefulness.

A further distinction is made between explicit definability and equational definability of operation constants. The latter term points to a definition by means of equality. It is known by a result of K. Gödel [6]) that e.g. the higher functions in the sequence of W. Ackermann [7]) are arithmetical, i.e. definable explicitly and validly from addition and multiplication with respect to the formalized arithmetic of natural numbers. At the same time they are undefinable equationally from the lower ones in the sequence with respect to the same theory. Application of the method developed in this study to show the explicit undefinability of these functions with respect to certain subtheories of the formalized arithmetic of natural numbers embodies a method to show their equational undefinability with respect to the complete theory of the formalized arithmetic of natural numbers.

It is the author's conviction that the suggestions made in this study are not more than a start for further investigations.

---

[6]) Cf. [6].

[7]) Cf. [1].

CONTENTS

Preface xiii

Chapter I. Definability

I, 1. Introduction 1
I, 2. Explicit definability of relation constants 2
I, 3. Explicit definability of operation constants (and individual constants) 5
I, 4. Equational definability of operation constants (and individual constants) 9

Chapter II. A method in proofs of undefinability

II, 1. Introductory remarks 14
II, 2. Direct proofs of undefinability 16
II, 3. Comparison with Padoa's method 20
II, 4. Indirect method 22

Chapter III. Addition and multiplication in the formalized arithmetic of natural numbers

III, 1. Introductory remarks 25
III, 2. Undefinability of "+" from "1" and "*S*" 32
III, 3. Undefinability of "·" from "1" and "+" 35
III, 4. Comments 39

Chapter IV. Results concerning equational undefinability

IV, 1. The Ackermann sequence 41
IV, 2. Equational undefinability 49

References 59

Index 61

CHAPTER I

DEFINABILITY

I, 1. *Introduction*

In the following investigations we consider theories with standard formalization (theories which are formalized within the first-order predicate logic – with identity, without variable predicates) [1]). In the informal discussion we refer to the non-logical constants in general of these theories as $c_1 \dots c_k$ and d (eventually $d_1 \dots d_m$), to the individual constants as $i, i_1 \dots i_l$, to the operation constants as $o, o_1 \dots o_m$ and to the predicates (relation constants) as $r, r_c, r_1 \dots r_n$; we refer to terms as $P, Q, R, \dots$, to formulas as $U, U', V, V', W, W', \dots$, and to sentences as $C, D, S, S', S_1 \dots S_k, \dots$. The theories are supposed to be formalized in the familiar logical notation with "$\vee$", "&", "–", "$\rightarrow$", "$\leftrightarrow$", "$=$", "x", "y", "z", "u", ... "x_1", "x_2", ... "(x)", "(y)", ... "(Ex)", "(Ey)", ...; where confusion is practically excluded we use the same symbols for their names in the informal discussion.

Finite concatenations of symbols of these theories are called expressions. Among expressions we distinguish terms and formulas. The simplest, so-called atomic, terms are the individual constants and the variables. A compound term is obtained by applying a k-ary operation constant on k simpler terms. The simplest, so-called atomic, formulas are obtained by combining two arbitrary terms by means of "$=$" (throughout considered as a logical constant) or by combining k arbitrary terms by means of a k-ary predicate. Compound formulas are built from simpler ones by means of the negation "–", the sentential connectives "$\vee$", "&", "$\rightarrow$" and "$\leftrightarrow$" and quantifier expressions "(x)", "(y)", "(z)", ... "(Ex)", "(Ey)", "(Ez)", ... A formula in which no variable occurs free is called a sentence. Technical symbols such as parentheses and commas are used in the familiar way; we could dispense with them as usual.

---

1) For a description of these theories and the meanings of various terms used here see e.g. [22], pp. 5 ff.

I, 2. *Explicit definability of relation constants*

Although it is possible to use relation constants without knowing, intuitively speaking, whether they apply to all individuals, to none or to some, we confine ourselves in the theories under consideration to such relation constants r (assumed e.g. to be k-ary) for which the sentences:

$$\overline{(x_1) \ldots (x_k)\, r(x_1 \ldots x_k)},$$

$$\overline{(x_1) \ldots (x_k)\, \overline{r(x_1 \ldots x_k)}},$$

are valid in the theories concerned. We call these sentences the sentences *characteristic* for the relation constants r.

A sentence D is called an explicit definition of the k-ary relation constant r from the individual constants $i_1 \ldots i_l$, the operation constants $o_1 \ldots o_m$ and the relation constants $r_1 \ldots r_n$, if D has the shape [2]):

$$(x_1) \ldots (x_k)\, (r\,(x_1 \ldots x_k) \leftrightarrow U(x_1 \ldots x_k, i_1 \ldots i_l, o_1 \ldots o_m, r_1 \ldots r_n)),$$

where $U(x_1 \ldots x_k, i_1 \ldots i_l, o_1 \ldots o_m, r_1 \ldots r_n)$ is a formula containing the free variables "x_1" ... "x_k", no other free variables and no other non-logical constants than $i_1 \ldots i_l, o_1 \ldots o_m$ and $r_1 \ldots r_n$.

We make use of two senses in which a relation constant r can be said to be definable explicitly from other non-logical constants $c_1 \ldots c_k$ with respect to a theory T:

i) a strong sense: if there is an explicit definition D of r from $c_1 \ldots c_k$ which is *valid* in T. In this case we say that r is *definable explicitly and validly* from $c_1 \ldots c_k$ with respect to T. (This is the usual sense of "definable explicitly".)

ij) a weak sense: if there is an explicit definition D of r from $c_1 \ldots c_k$ which is *compatible* with T. In this case we say that r is *definable explicitly and compatibly* from $c_1 \ldots c_k$ with respect to T.

For a k-ary relation constant r, given as such in a consistent theory, to be definable explicitly and validly or compatibly from other non-logical constants with respect to the theory concerned it follows from the restriction expounded above (the sentences characteristic for a relation constant) that the explicit definition

[2]) Cf. e.g. [4], p. 31.

in question has to fulfil certain conditions. If D is

$$(x_1) \dots (x_k)(r(x_1 \dots x_k) \leftrightarrow U),$$

the sentences:

$$(x_1) \dots (x_k)U,$$
$$(x_1) \dots (x_k)\overline{U},$$

may not be valid in T or compatible with $\mathsf{T} \cup D$, the union of T and D. The former of these sentences together with D would imply the validity (compatibility) of

$$(x_1) \dots (x_k)\, r(x_1 \dots x_k),$$

and the latter together with D would imply the validity (compatibility) of

$$(x_1) \dots (x_k)\, \overline{r(x_1 \dots x_k)},$$

contrary to the characteristic sentences.

Thus the idea of definability is guarded against quite trivial definitions.

Given a consistent theory T containing the non-logical constants $c_1 \dots c_k$ and r, the question arises whether r is definable explicitly from $c_1 \dots c_k$ with respect to T. A relation constant r is called definable explicitly from $c_1 \dots c_k$ with respect to T or undefinable explicitly from $c_1 \dots c_k$ with respect to T according as the solution of this problem is positive or negative. If r cannot be defined explicitly and validly from $c_1 \dots c_k$ with respect to T, we say simply that r is *undefinable explicitly* from $c_1 \dots c_k$ with respect to T; if r cannot be defined explicitly and compatibly from $c_1 \dots c_k$ with respect to T, we say that r is *essentially undefinable explicitly* from $c_1 \dots c_k$ with respect to T.

The problem whether or not r is definable explicitly from the other non-logical constants $c_1 \dots c_k$ with respect to T, and in what sense, we call the *definition problem for r with respect to* T.

If r is definable explicitly and validly from $c_1 \dots c_k$ with respect to T, then it is obvious that r is definable explicitly and compatibly from $c_1 \dots c_k$ with respect to T. If r is essentially undefinable explicitly from $c_1 \dots c_k$ with respect to T, then it is obvious that r is undefinable explicitly from $c_1 \dots c_k$ with respect to T.

It is possible to characterize the idea of explicit definability in terms of models.

Let T be a consistent theory containing the non-logical constants $c_1 \dots c_k$ and r. Let T_0 be the subtheory of T containing all sentences of T in which r does not occur (a sentence being valid in T_0 if and only if it is valid in T).

> For r to be definable explicitly and validly from $c_1 \dots c_k$ with respect to T it is necessary and sufficient that any model M_0 of T_0 can be extended, and extended in one way only, with a notion $\mathbf{r}$ as an interpretation of r to a model M of T.

i) The condition is necessary. If r is definable explicitly and validly from $c_1 \dots c_k$ with respect to T, there is an explicit definition D of r from $c_1 \dots c_k$ such that T is $T_0 \cup D$, the union of T_0 and D. Let M_0 be an arbitrary model $\langle \mathbf{U}; \mathbf{c}_1 \dots \mathbf{c}_k \rangle$ of T_0, $\mathbf{U}$ being the universe and $\mathbf{c}_1 \dots \mathbf{c}_k$ the interpretations of $c_1 \dots c_k$ respectively. M_0 can always be extended with an interpretation $\mathbf{r}$ of r to M, – a model $\langle \mathbf{U}; \mathbf{c}_1 \dots \mathbf{c}_k, \mathbf{r} \rangle$ of T; for the explicit definition D of r from $c_1 \dots c_k$ can be interpreted as an explicit definition of $\mathbf{r}$ from $\mathbf{c}_1 \dots \mathbf{c}_k$. On the other hand, M_0 cannot be extended in more than one way with interpretations of r, e.g. with $\mathbf{r}$ and $\mathbf{r}'$. Two different extensions would be contrary to the fact that r is definable explicitly and validly from $c_1 \dots c_k$ with respect to T, according to the method of Padoa [3]).

ij) The condition is sufficient. If any model M_0 of T_0 can be extended, and extended in one way only, with an interpretation $\mathbf{r}$ of r to a model M of T, then it is impossible that there is one model M_0 of T_0 which can be extended into two different models of T, namely $\langle \mathbf{U}; \mathbf{c}_1 \dots \mathbf{c}_k, \mathbf{r} \rangle$ and $\langle \mathbf{U}; \mathbf{c}_1 \dots \mathbf{c}_k, \mathbf{r}' \rangle$. This, however, has to be the case according to Beth's result [4]), whenever r cannot be defined explicitly and validly from $c_1 \dots c_k$ with respect to T. It follows that r is definable explicitly and validly from $c_1 \dots c_k$ with respect to T.

For r to be definable explicitly and compatibly from $c_1 \dots c_k$ with respect to T it is necessary and sufficient that r is definable

[3]) Cf. [13], pp. 321, 322; [14], pp. 250, 254 ff.; [15], p. 91.

[4]) Cf. [2] and [4], pp. 29 ff.

explicitly and validly from $c_1 \dots c_k$ with respect to some consistent extension of T having the same non-logical constants as T. This follows directly from the notions involved. So we can state:

> For r to be definable explicitly and compatibly from $c_1 \dots c_k$ with respect to T it is necessary and sufficient that there is some consistent extension T* of T, having the same non-logical constants as T, such that, when T_0* is the subtheory of T* containing all sentences of T* in which r does not occur, any model M_0^* of T_0* can be extended, and extended in one way only, with a notion **r** as an interpretation of r to a model M^* of T*.

I, 3. *Explicit definability of operation constants (and individual constants)*

In the theories under consideration for a k-ary operation constant o the following characteristic sentences are assumed to be valid:

$$(x_1)\dots(x_k)(Ex_{k+1})(o(x_1\dots x_k)=x_{k+1}),$$
$$(x_1)\dots(x_k)(x_{k+1})(x_{k+2})((o(x_1\dots x_k)=x_{k+1} \,\&\, o(x_1\dots x_k)=x_{k+2}) \rightarrow \rightarrow x_{k+1}=x_{k+2}).$$

A sentence D is called an explicit definition of the k-ary operation constant o from the individual constants $i_1 \dots i_l$, the operation constants $o_1 \dots o_m$ and the relation constants $r_1 \dots r_n$ if D has the shape:

$$(x_1) \dots (x_k)(x_{k+1})(o(x_1 \dots x_k)=x_{k+1} \leftrightarrow U(x_1 \dots x_{k+1}, i_1 \dots i_l, o_1 \dots o_m, r_1 \dots r_n)),$$

where $U(x_1 \dots x_{k+1}, i_1 \dots i_l, o_1 \dots o_m, r_1 \dots r_n)$ is a formula containing the free variables "x_1" ... "x_k", "x_{k+1}", no other free variables and no other non-logical constants than $i_1 \dots i_l$, $o_1 \dots o_m$ and $r_1 \dots r_n$.

Likewise, a sentence D is called an explicit definition of the individual constant i from the individual constants $i_1 \dots i_l$, the operation constants $o_1 \dots o_m$ and the relation constants $r_1 \dots r_n$ if D has the shape:

$$(x)(i=x \leftrightarrow U(x, i_1 \dots i_l, o_1 \dots o_m, r_1 \dots r_n)),$$

where $U(x, i_1 \dots i_l, o_1 \dots o_m, r_1 \dots r_n)$ is a formula containing the free variable "x", no other free variables and no other non-logical constants than $i_1 \dots i_l$, $o_1 \dots o_m$ and $r_1 \dots r_n$ [5]).

We can treat in this context individual constants as 0-ary operation constants.

An operation constant (individual constant) o is said to be *definable explicitly and validly* from other non-logical constants $c_1 \dots c_k$ with respect to a theory T, if there is an explicit definition D of o from $c_1 \dots c_k$ which is valid in T. It is said to be *definable explicitly and compatibly* from $c_1 \dots c_k$ with respect to T, if there is an explicit definition D which is compatible with T.

For a k-ary operation constant, given as such in a consistent theory T, to be definable explicitly and validly or compatibly from other non-logical constants with respect to the theory concerned it follows from the validity of the characteristic sentences that the explicit definition in question has to fulfil certain conditions. If D is

$$(x_1) \dots (x_k)(x_{k+1})(o(x_1 \dots x_k) = x_{k+1} \leftrightarrow U),$$

the sentence:

$$(x_1) \dots (x_k)(Ex_{k+1})(U \;\&\; (y)(U' \rightarrow x_{k+1} = y)),$$

has to be valid in T (compatible with $\mathsf{T} \cup D$), where "y" is an arbitrary variable (different from "x_1" ... "x_k", "x_{k+1}") which does not occur in U, and U' is the formula obtained from U by replacing "x_{k+1}" everywhere by "y" [6]).

In the same way as in the case of a relation constant we speak of the definition problem for o with respect to T, distinguishing between "undefinable explicitly" and "essentially undefinable explicitly".

If o is definable explicitly and validly from $c_1 \dots c_k$ with respect to T, then it is obvious that o is definable explicitly and compatibly from $c_1 \dots c_k$ with respect to T. If o is essentially undefinable explicitly from $c_1 \dots c_k$ with respect to T, then it is obvious that o is undefinable explicitly from $c_1 \dots c_k$ with respect to T.

---

5) Cf. e.g. [21], p. 299.

6) Cf. [22], p. 20.

There is a structural difference between the explicit definition of a relation constant and the explicit definition of an operation constant (individual constant). Both explicit definitions are contextual, but in the former the *definitum* contains with the *definiendum* only variables, in the latter the *definitum* contains also the symbol "=" [7]).

The *definitum* in the explicit definition of an operation constant is in fact a relation. For any k-ary operation constant o in T can there be introduced into T a corresponding $(k+1)$-ary relation constant r_c by the explicit definition C:

$$(x_1) \ldots (x_k)(x_{k+1})(r_c(x_1 \ldots x_k\, x_{k+1}) \leftrightarrow o(x_1 \ldots x_k) = x_{k+1}).$$

It follows from the sentences characteristic for an operation constant that in $\mathsf{T} \cup C$ the sentences:

$$(x_1) \ldots (x_k)(Ex_{k+1})\, r_c(x_1 \ldots x_k\, x_{k+1}),$$
$$(x_1) \ldots (x_k)(x_{k+1})(x_{k+2})\,((r_c(x_1 \ldots x_k x_{k+1})\, \& r_c(x_1 \ldots x_k x_{k+2})) \rightarrow x_{k+1} = x_{k+2}),$$

are valid. Hence r_c fulfils in $\mathsf{T} \cup C$ the conditions characteristic for a relation constant.

From a structural point of view in the case of the explicit definition of an operation constant (individual constant) it would seem more consistent to speak of the explicit definition of the corresponding relation constant. However, the characterization in terms of models fully justifies the common terminology to speak of the explicit definition of the operation constant itself.

Let T be a consistent theory containing the non logical constants $c_1 \ldots c_k$ and o. Let T_0 be the subtheory of T containing all sentences of T in which o does not occur (a sentence being valid in T_0 if and only if it is valid in T).

> For o to be definable explicitly and validly from $c_1 \ldots c_k$ with respect to T it is necessary and sufficient that any model M_0 of T_0 can be extended, and extended in one way only, with a notion $\mathbf{o}$ as an interpretation of o to a model M of T.

---

[7]) Cf. e.g. [5], p. 211. From a structural point of view the term *definitum* means the whole left-hand member of the equivalence; the term *definiendum* concerns only the defined constant in question.

(The common terminology would seem inadequate if the last part of this characterization had to be read: ... that any model M_0 of T_0 can be extended and extended in one way only with a notion $\mathbf{r}_c$ as an interpretation of the with o corresponding relation constant r_c to a model M of T.)

i) The condition is necessary. If o is definable explicitly anp validly from $c_1 \dots c_k$ with respect to T, then there is an explicit definition D of o from $c_1 \dots c_k$ such that T is $\mathsf{T}_0 \cup D$. This is not so obvious as in the analogous case of the relation constant, for T may contain sentences where o occurs in another context than $o(x_1 \dots x_k) = x_{k+1}$, so that it cannot be replaced immediately with the help of D. Nevertheless, for any sentence of T containing o in a context different from the definitum there is an equivalent sentence of T containing o in the required context. This may be illustrated by a simple example. Let e.g. a sentence S of T contain once $o(x_1 \dots x_k)$ in an atomic formula V different from the definitum. According to the sentences characteristic for an operation constant we may replace $o(x_1 \dots x_k)$ in V by a new variable "y", different from "x_1" ... "x_k", which does not occur in S. Indicating the thus changed atomic formula as $V(y)$ we replace V by

$$(Ey)(o(x_1 \dots x_k) = y \;\&\; V(y)).$$

It is possible to specify the rules for a uniform process by which to any atomic formula V, containing o once or several times, a formula V' can be effectively constructed, containing o only in atomic formulas of the shape $o(x_1 \dots x_k) = y$, and having the property that V and V' are equivalent with respect to T [8]). We then argue: for any sentence S of T containing o there is an equivalent sentence S' of T_0 (without o) such that S is derivable from S' and D in $\mathsf{T}_0 \cup D$. Or, what amounts to the same thing, T is $\mathsf{T}_0 \cup D$. From what is said about relation constants it follows that any model M_0 of T_0 can be extended, and extended in one way only, with an interpretation of the relation $o(x_1 \dots x_k) = y$ to a

[8]) Two formulas $U(x_1 \dots x_k)$ and $V(x_1 \dots x_k)$ of a theory T, containing the same free variables $x_1 \dots x_k$ and no other free variables, are said to be equivalent with respect to T if and only if the sentence:

$$(x_1) \dots (x_k)(U \leftrightarrow V),$$

is valid in T.

model M of T, which is $T_0 \cup D$. When M_0 could not be extended with an interpretation **o** of o to a model M of T, then M_0 could not be extended with an interpretation of $o(x_1 \dots x_k)=y$. When M_0 could be extended in more than one way with interpretations of o, say **o** and **o**′, then M_0 could be extended in more than one way with interpretations of $o(x_1 \dots x_k)=y$. Thus is it necessary that M_0 can be extended, and extended in one way only, with an interpretation **o** of o to a model M of T.

ij) The condition is sufficient. If any model M_0 of T_0 can be extended, and extended in one way only, with an interpretation **o** of o to a model M of T, then any model M_0 of T_0 can be extended, and extended in one way only, with an interpretation of $o(x_1 \dots x_k)=y$ to a model M of T. This is sufficient for o to be definable explicitly and validly from $c_1 \dots c_k$ with respect to T.

Likewise we state:

> For o to be definable explicitly and compatibly from $c_1 \dots c_k$ with respect to T it is necessary and sufficient that there is some consistent extension T^* of T, having the same non-logical constants as T, such that, when $T_0{}^*$ is the subtheory of T^* containing all sentences of T^* in which o does not occur, any model $M_0{}^*$ of $T_0{}^*$ can be extended, and extended in one way only, with a notion **o** as an interpretation of o to a model M^* of T^*.

I, 4. *Equational definability of operation constants* (*and individual constants*)

In order to avoid confusion we distinguish in this study between explicit definitions and equational definitions of operation constants and individual constants. A sentence D is called an equational definition of the k-ary operation constant o from the individual constants $i_1 \dots i_l$ and the operation constants $o_1 \dots o_m$ if D has the shape [9]):

$$(x_1) \dots (x_k)(o(x_1 \dots x_k)=Q(x_1 \dots x_k, i_1 \dots i_l, o_1 \dots o_m)),$$

where $Q(x_1 \dots x_k, i_1 \dots i_l, o_1 \dots o_m)$ is a term containing the free variables "x_1" ... "x_k" (and no other variables) and no other

9) Cf. e.g. [9], pp. 292, 293.

(logical or non-logical) constants than $i_1 \dots i_l$ and $o_1 \dots o_m$ [10]).

Likewise a sentence D is called an equational definition of the individual constant i from the individual constants $i_1 \dots i_l$ and the operation constants $o_1 \dots o_m$ if D has the shape $i = Q(i_1 \dots i_l, o_1 \dots o_m)$, where $Q(i_1 \dots i_l, o_1 \dots o_m)$ is a term containing no variables and no other (logical or non-logical) constants than $i_1 \dots i_l$ and $o_1 \dots o_m$.

In this context also we may treat individual constants as 0-ary operation constants.

An operation constant (individual constant) o is said to be *definable equationally and validly* from other non-logical constants $c_1 \dots c_k$ with respect to a theory T, if there is an equational definition D of o from $c_1 \dots c_k$ which is valid in T. It is said to be *definable equationally and compatibly* from $c_1 \dots c_k$ with respect to T, if there is an equational definition D of o from $c_1 \dots c_k$ which is compatible with T.

Extending the definition problem to equational definitions we distinguish between "*undefinable equationally*" and "*essentially undefinable equationally*" in the same way as in the case of explicit definability.

If o is definable equationally and validly from $c_1 \dots c_k$ with respect to T, then it is obvious that o is definable equationally and compatibly from $c_1 \dots c_k$ with respect to T. If o is essentially undefinable equationally from $c_1 \dots c_k$ with respect to T, then it is obvious that o is undefinable equationally from $c_1 \dots c_k$ with respect to T.

Let T be a consistent theory containing the non-logical constants (individual constants and operation constants only) $c_1 \dots c_k$ and o. Let T_0 be the subtheory of T containing all sentences of T in which o does not occur (a sentence being valid in T_0 if and only if it is valid in T).

For o to be definable equationally and validly from $c_1 \dots c_k$ with respect to T it is necessary but not sufficient that any

[10]) A sentence D having the shape:

$$(x_1) \dots (x_k)(o(x_1 \dots x_k) = Q(i_1 \dots i_l, o_1 \dots o_m)),$$

wherein $Q(i_1 \dots i_l, o_1 \dots o_m)$ is a term containing no variables at all, can also be called an equational definition of o. In this treatment, however, these so-called constant functions are not under consideration. The same applies when Q contains only some of the free variables "x_1" ... "x_k".

model M_0 of T_0 can be extended, and extended in one way only, with a notion **o** as an interpretation of o to a model M of T.

Here lies the first model-theoretical difference between explicit definability and equational definability (we shall see another difference in IV, 2).

i) The condition is necessary. If o is definable equationally and validly from $c_1 \dots c_k$ with respect to T, then there is an equational definition D of o from $c_1 \dots c_k$ such that T is $\mathsf{T}_0 \cup D$. Let M_0 be an arbitrary model $\langle \mathbf{U}; \mathbf{c}_1 \dots \mathbf{c}_k \rangle$ of T_0. M_0 can always be extended with an interpretation **o** of o to a model M of T, for the equational definition D of o from $c_1 \dots c_k$ can be interpreted as an equational definition of **o** from $\mathbf{c}_1 \dots \mathbf{c}_k$. On the other hand, M_0 cannot be extended in more than one way with interpretations of o, e.g. with **o** and **o**′. The same argumentation as Padoa's can be applied directly to equational definitions.

ij) The condition is not sufficient. If every model M_0 of T_0 can be extended, and extended in one way only, with an interpretation **o** of o to a model M of T, it follows that o is definable explicitly and validly from $c_1 \dots c_k$ with respect to T. Thus there is an explicit definition D:

$$(x_1) \dots (x_k)(x_{k+1})(o(x_1 \dots x_k) = x_{k+1} \leftrightarrow U),$$

valid in T, where U is a formula as mentioned above. It does not follow that U has the shape:

$$x_{k+1} = Q,$$

where Q is a term as required in the right-hand member of an equational definition, and it does not follow either that U is a formula equivalent with $x_{k+1} = Q$. Hence it does not follow that o is definable equationally and validly from $c_1 \dots c_k$ with respect to T. We shall see counterexamples in Chapter IV.

Likewise we state:

For o to be definable equationally and compatibly from $c_1 \dots c_k$ with respect to T it is necessary but not sufficient that there is some consistent extension T^* of T, having the same non-logical constants as T, such that, when $\mathsf{T}_0{}^*$ is the subtheory of T^* containing all sentences of T in which o does not occur, any model $M_0{}^*$ of $\mathsf{T}_0{}^*$ can be extended, and extended in one way only, with a notion $\mathbf{o}$ as an interpretation of o to a model M^* of T^*.

It may be useful to remark, although trivial in itself, that it follows from the characterizations in terms of models that, if an operation constant (individual constant) o is definable equationally and validly (or compatibly) from $c_1 \ldots c_k$ with respect to T, it is also definable explicitly and validly (or compatibly, respectively) from $c_1 \ldots c_k$ with respect to T.

This can also be argued in the following simple way: if o is definable equationally and validly (compatibly) from $c_1 \ldots c_k$ with respect to T, then is there an equational definition D valid in (compatible with) T. If D is

$$(x_1) \ldots (x_k)(o(x_1 \ldots x_k) = Q),$$

(where Q is a term without o and without other variables than "x_1" ... "x_k") then the sentence:

$$(x_1) \ldots (x_k)(x_{k+1})(o(x_1 \ldots x_k) = x_{k+1} \leftrightarrow x_{k+1} = Q),$$

is valid in (compatible with) T. Hence o is definable explicitly and validly (compatibly) from $c_1 \ldots c_k$ with respect to T.

In connection with the following investigations we wish to remark that, if an operation constant (individual constant) o is undefinable explicitly, then it is undefinable equationally, and if o is essentially undefinable explicitly, then it is essentially undefinable equationally. Therefore, in order to show that an operation constant o is undefinable equationally or essentially undefinable equationally from $c_1 \ldots c_k$ with respect to T, it is sufficient to show that o is undefinable explicitly, respectively essentially undefinable explicitly from $c_1 \ldots c_k$ with respect to T. If, however, o is definable explicitly (validly or compatibly) from other non-logical constants with respect to T, it does not follow that o is definable equationally (validly or compatibly) from other individual constants and operation constants with respect to T, for it is possible that in the explicit definition of o the right-hand member of the equivalence neither has the shape $x_{k+1} = Q$ nor is equivalent with such a formula. We shall see counterexamples in Chapter IV.

If i is the only individual constant of a consistent theory T it follows from the concepts involved that i is essentially undefinable equationally from other non-logical constants with respect to T. This is trivial. It does not follow, however, that in this case i is undefinable explicitly or essentially undefinable explicitly from other non-logical constants with respect to T, for equational undefinability does not imply explicit undefinability. If o is the only operation constant in T it follows in a trivial way from the concepts involved that o is essentially undefinable equationally from other non-logical constants with respect to T. Again this implies nothing about explicit undefinability of o.

CHAPTER II

A METHOD IN PROOFS OF UNDEFINABILITY

II, 1. *Introductory remarks*

In this chapter we attempt to establish a method for obtaining a negative solution to the definition problem for a non-logical constant *d* (relation constant, operation constant or individual constant) with respect to a consistent theory T. Dealing with a negative solution we confine the terminology of the method to explicit definability. A non-logical constant *d* being undefinable explicitly or essentially undefinable explicitly from other non-logical constants with respect to a theory T is automatically undefinable equationally or essentially undefinable equationally with respect to T. On the other hand, in case *d* is definable explicitly and validly or compatibly from other non-logical constants with respect to T the method cannot be used to prove eventually that *d* is undefinable equationally or essentially undefinable equationally from the other non-logical constants with respect to T. Therefore, in the lemmas concerning the method, we only speak of "undefinable explicitly" and "essentially undefinable explicitly". The crucial case of proving the equational undefinability of a non-logical constant which is definable explicitly, will be discussed later (in IV, 2, pp. 49 ff.).

Padoa's [1]) well-known general method in proofs of undefinability can be based on one side of the characterization of explicit definability in terms of models [2]), namely on the necessity that any model M_0 (cf. I, 2 and I, 3) can be extended *in one way only* with an interpretation of the non-logical constant concerned. The method consists in giving an M_0 that can be extended with more than one interpretation, thus proving that the non-logical constant concerned is undefinable explicitly from the other non-logical constants with respect to the theory concerned.

---

1) Cf. [13], pp. 321 ff.

2) Cf. I, 2, p. 4.

The method in proofs of undefinability, which is developed in this study, is based on the other side of the characterization of explicit definability in terms of models, namely on the necessity that any model M_0 *can be extended* with an interpretation of the non-logical constant concerned. The method consists in giving an M_0 that cannot be extended in any way with such an interpretation, thus proving that the non-logical constant concerned is (essentially) undefinable explicitly from the other non-logical constants with respect to the theory concerned.

We shall develop this method by some lemmas in II, 2. Afterwards we shall discuss the advantages and disadvantages of the method as compared to the method of Padoa.

This introduction may be concluded with some general remarks about explicit undefinability.

If T is a complete theory, all sentences compatible with T are valid in T and all sentences not compatible with T are not valid in T. Let T be complete and contain the non-logical constants $c_1 \ldots c_k$ and d, then d is not only definable explicitly and compatibly from $c_1 \ldots c_k$ with respect to T whenever d is definable explicitly and validly from $c_1 \ldots c_k$ with respect to T, but the converse holds also. Likewise in the case of a complete theory d is not only undefinable explicitly from $c_1 \ldots c_k$ with respect to T whenever d is essentially undefinable explicitly from $c_1 \ldots c_k$ with respect to T, but the converse holds also. Hence we state:

> For a complete theory T the condition: *d is definable explicitly and validly from $c_1 \ldots c_k$ with respect to* T is equivalent to the condition: *d is definable explicitly and compatibly from $c_1 \ldots c_k$ with respect to* T; likewise the condition: *d is undefinable explicitly from $c_1 \ldots c_k$ with respect to* T is equivalent to the condition: *d is essentially undefinable explicitly from $c_1 \ldots c_k$ with respect to* T.

Another introductory remark having some interest in itself is the following: let d be essentially undefinable explicitly from $c_1 \ldots c_k$ with respect to a theory T. Since in this case d is undefinable explicitly from $c_1 \ldots c_k$ with respect to every consistent extension of T, d certainly is undefinable explicitly from $c_1 \ldots c_k$

with respect to every consistent and complete extension of T, which has the same constants as T. On the other hand, let D be an explicit definition of d from $c_1 \ldots c_k$ which is compatible with T and let T be consistent. Every consistent theory has a consistent and complete extension, which has the same constants, as we know by theorems of Lindenbaum and Tarski [3]. $\mathsf{T} \cup D$ has a consistent and complete extension with the same constants, where of course D is valid, so that d is definable explicitly and validly from $c_1 \ldots c_k$ with respect to this consistent and complete extension. Hence we state:

> For a non-logical constant d to be essentially undefinable explicitly from non-logical constants $c_1 \ldots c_k$ with respect to a theory T it is necessary and sufficient that d is undefinable explicitly from $c_1 \ldots c_k$ with respect to every consistent and complete extension of T, which has the same constants as T [4].

II, 2. *Direct proofs of undefinability*

Let T be a consistent theory containing no other non-logical constants than $c_1 \ldots c_k$ and d. Assume d to be definable explicitly and validly from $c_1 \ldots c_k$ with respect to T. Let T_0 be the subtheory of T containing all sentences of T in which d does not occur (a sentence being valid in T_0 if and only if it is valid in T). Let M_0 be an arbitrary model of T_0.

According to the characterization of explicit definability in terms of models M_0 admits the introduction of a notion **d**, which is an interpretation of the d of T.

Conversely, when into M_0 there can in no way be introduced a notion **d** as an interpretation of the d of T, then d cannot be defined explicitly and validly from $c_1 \ldots c_k$ with respect to T:

Lemma 1: Let T be a consistent theory containing the non-logical constants $c_1 \ldots c_k$ and d. Let T_0 be the subtheory of T containing the non-logical constants $c_1 \ldots c_k$ and all sentences of T in which d does not occur. If T_0 has a model M_0 into which there can in no way be introduced a notion **d** as an interpretation

---

3) Cf. [22], pp. 15, 16.

4) This fact justifies the terminology, cf. [22], p. 14, where A. Tarski speaks of "essentially undecidable".

of the d of T, then d is undefinable explicitly from $c_1 \dots c_k$ with respect to T [5]).

Corollary: Let T be a consistent theory containing the non-logical constants $c_1 \dots c_k$ and $d_1 \dots d_m$. Let T_0 be the subtheory of T containing the non-logical constants $c_1 \dots c_k$ and all sentences of T in which $d_1 \dots d_m$ do not occur. If T_0 has a model M_0 into which there can in no way be introduced simultaneously notions $\mathbf{d}_1 \dots \mathbf{d}_m$ as interpretations of $d_1 \dots d_m$ of T, then at least one of $d_1 \dots d_m$ is undefinable explicitly from $c_1 \dots c_k$ with respect to T.

The argument of Lemma 1 is not sufficient to read this lemma with "essentially undefinable" instead of "undefinable". Assuming d to be definable explicitly and compatibly from $c_1 \dots c_k$ with respect to T, without being definable explicitly and validly from $c_1 \dots c_k$ with respect to T, then no explicit definition of d from $c_1 \dots c_k$ is valid in T. But there is, in this case, at least one such a definition which is compatible with T. Let D be such an explicit definition, which is compatible with T. In general the valid sentences of T containing d, translated by D to sentences without d, although all compatible with T_0, are not all valid in T_0. Hence T is not necessarily a subtheory of $T_0 \cup D$ and although any model of T_0 can be extended to a model of $T_0 \cup D$, it is not implied that any model of T_0 can be extended to a model of T. The same can be said for all explicit definitions of d from $c_1 \dots c_k$ which are compatible with T.

If, however, T is complete all sentences compatible with T are valid in T. So we state:

Lemma 2: Let T be a consistent and complete theory containing the non-logical constants $c_1 \dots c_k$ and d. Let T_0 be the subtheory of T containing the non-logical constants $c_1 \dots c_k$ and all sentences of T in which d does not occur. If T_0 has a model M_0 into which there can in no way be introduced a notion $\mathbf{d}$ as an interpretation of the d of T, then d is essentially undefinable explicitly from $c_1 \dots c_k$ with respect to T.

---

[5]) As an illustration cf. the construction in [4], p. 34, (2). The analysis of this example provided the starting point of the present investigation.

Corollary: The same as the corollary to Lemma 1, with "consistent and complete" instead of "consistent" and "essentially undefinable" instead of "undefinable".

If T is not complete but T_0 is complete and D is an explicit definition of d from $c_1 \ldots c_k$ compatible with T, then all valid sentences of T translated by D to sentences of T_0 are not only compatible with T_0 but valid in T_0, as this theory is complete. So any model of T_0 can be extended to a model of T and the argument of Lemma 1 holds:

Lemma 3: Let T be a consistent theory containing the non-logical constants $c_1 \ldots c_k$ and d. Let T_0 be the subtheory of T containing the non-logical constants $c_1 \ldots c_k$ and all sentences of T in which d does not occur. If T_0 is complete and has a model M_0 into which there can in no way be introduced a notion $\mathbf{d}$ as an interpretation of the d of T, then d is essentially undefinable explicitly from $c_1 \ldots c_k$ with respect to T.

Corollary: The same as the corollary to Lemma 1 with "If T_0 is complete and has a model M_0" instead of "If T_0 has a model M_0" and "essentially undefinable" instead of "undefinable".

If a non-logical constant d is undefinable explicitly from $c_1 \ldots c_k$ with respect to T, the question arises whether there exists always a model M_0 of T_0 into which there can in no way be introduced a notion $\mathbf{d}$ as an interpretation of the d of T. The answer does not follow from the characterization in terms of models. For it is possible that any model M_0 of T_0 can be extended with a notion $\mathbf{d}$, an interpretation of d, to a model M of T, provided that there is at least one M_0 which can be extended with two different interpretations of d to two different models of T, namely $\langle \mathbf{U}; \mathbf{c}_1 \ldots \mathbf{c}_k \mathbf{d} \rangle$ and $\langle \mathbf{U}; \mathbf{c}_1 \ldots \mathbf{c}_k \mathbf{d}' \rangle$. According to Beth's result [6]), the last condition is always fulfilled: when d is undefinable explicitly from $c_1 \ldots c_k$ with respect to T, then there are always two models $\langle \mathbf{U}; \mathbf{c}_1 \ldots \mathbf{c}_k \mathbf{d} \rangle$ and $\langle \mathbf{U}; \mathbf{c}_1 \ldots \mathbf{c}_k \mathbf{d}' \rangle$ of T. It follows that the model-theoretical point of view of this study does not imply the necessity of the existence of a model M_0 of T_0 into which there can in no way be introduced a notion $\mathbf{d}$ as an interpretation of d, a non-logical

6) Cf. [4], pp. 29 ff.

constant which is undefinable explicitly from $c_1 \dots c_k$ with respect to T.

In fact, the answer to the general question is negative, as can be demonstrated by an example. Let T be a theory with standard formalization containing no other non-logical constants than "$<$" and "$+1$" ("$=$" is throughout considered as a logical constant), which represents a suitable part of the arithmetic of integers; e.g. let T be the theory described in [4], p. 30, example 16. T is axiomatically built on the axioms:

$$(x)(y)(\overline{x=y} \leftrightarrow (x<y \vee y<x));$$
$$(x)(y)(\overline{x<y} \leftrightarrow (x=y \vee y<x));$$
$$(x)(y)(z)((x<y \ \& \ y<z) \rightarrow x<z);$$
$$(x)(x<x+1);$$
$$(z)(Ey)(y=z+1);$$
$$(z)(Ey)(y+1=z);$$
$$(x)(y)(z)((x<y \ \& \ y<z) \rightarrow x+1<z).$$

As Beth (l.c.) shows by Padoa's method "$<$" is undefinable explicitly from "$+1$" with respect to T. But any model M_0 of T_0 (which is the subtheory of T containing all sentences of T where "$<$" does not occur) can be extended with an interpretation $\boldsymbol{<}$ of the "$<$" of T, for any infinite set satisfying the conditions of T_0 can be ordered in the required way. According to a result of Henkin [7]) every set can be simply ordered. A simple ordering fulfilling the conditions:

$$(x)(x<x+1),$$
$$(x)(y)(z)((x<y \ \& \ y<z) \rightarrow x+1<z),$$

can be obtained in the following way. Let $\mathbf{U}_0$ be the universe of M_0. Let $\mathbf{r}$ be the binary relation in $\mathbf{U}_0$ such that, for two arbitrary elements $\mathbf{a}$ and $\mathbf{b}$ of $\mathbf{U}_0$, $\mathbf{r}(\mathbf{a}, \mathbf{b})$ if and only if $\mathbf{b}$ is one of the elements:

$$\dots \ \mathbf{a-2, \ a-1, \ a, \ a+1, \ a+2}, \ \dots .$$

---

[7]) Cf. [7]. The ordering principle 0_1.

It is obvious that **r** is a reflexive, symmetric and transitive relation, dividing $\mathbf{U}_0$ in classes **A**, **B**, **C**, ... Within each class there is a simple ordering fulfilling the required conditions. Furthermore[8]), the family {**A**, **B**, **C**, ...} can be simply ordered according to Henkin. The result is a simply ordered set fulfilling the required conditions[9]).

II, 3. *Comparison with Padoa's method*

The two-models method of Padoa and the one-model method developed in this work, to prove the independence of a non-logical constant from other non-logical constants with respect to a consistent theory T, form together a pendant of the well-known method to show the independence of a sentence S from other sentences $S_1 \dots S_k$ with respect to a theory T (i.e. that S is not derivable from $S_1 \dots S_k$ in T). Roughly speaking, this method consists in producing a model M_0 which is a model of $S_1 \dots S_k$ without being a model of S; the theorem of Löwenheim–Skolem–Gödel guarantees the existence of such a model.

The one-model method to prove the undefinability of a non-logical constant from others with respect to a theory T is analogous to the method to prove the independence of a sentence from others with respect to a theory T in this sense, that both methods make use of one single model[10]). However, the one-model method for non-logical constants seems not to be supported by a simple pendant of the theorem of Löwenheim–Skolem–Gödel.

The two-models method of Padoa to prove the undefinability of a non-logical constant from others with respect to a theory T is supported by Beth's theorem[11]), which is quite analogous to the theorem of Löwenheim–Skolem–Gödel. Beth's theorem guarantees the existence of two models to prove the undefinability in question. The method itself, however, is not analogous to the method showing the independence of sentences.

---

[8]) Cf. [19], where Th. Skolem proves the existence of non-classical models.

[9]) Cf. [23]. A Tarski's ordering principles 0_2 and 0_3 are more complicated than the one developed here.

[10]) Recently J. G. Kemeny approached the problem of the independence of sentences in a way very similar to the way developed in this study concerning the independence of notions. Cf. [10] and this work Chapter III.

[11]) Cf. [4], pp. 29 ff.

From a theoretical point of view Padoa's two-models method is wider than the one-model method: whenever a non-logical constant is undefinable explicitly from other non-logical constants with respect to a theory T, it is possible to prove this situation with Padoa's method. The same cannot be said about the one-model method. In practice it does not make much difference, since the real problem in most cases consists in finding out whether or not a non-logical constant is definable explicitly from the other non-logical constants with respect to T. From a heuristic point of view the construction of one model might have advantages over the construction of two models. This can be illustrated by one of the few cases where we know for other reasons that a certain non-logical constant is undefinable explicitly from the others with respect to a theory T. The case in question concerns the explicit undefinability of multiplication ("·") from addition ("+") with respect to a theory describing the elementary properties of the arithmetic of natural numbers with addition and multiplication. Since the theory with "1" and "+" is decidable (Presburger [12]) and the theory with "1", "+" and "·" is undecidable (Barkley Rosser [13]) it follows that "·" is undefinable explicitly from "1" and "+" with respect to the last mentioned theory [14]. Hence, it is sure that there exist two models, different in the multiplication only, to prove this undefinability by Padoa's method. However, as far as the author knows, two such models have not yet been directly constructed. On the other hand, it is possible to prove the undefinability in question directly with the one-model method, as will be done in III, 3, where we construct a model into which in no way multiplication can be introduced.

When one does not know beforehand whether or not a non-logical constant is definable explicitly, the attempts to construct the two models in order to apply Padoa's method take on the risk of being vain, since the non-logical constant might be definable explicitly. Attempts to construct one model in order to apply the one-model method run the same risk but from a double source: the non-logical constant in question might be definable explicitly and there might be no model to demonstrate an eventual unde-

12) Cf. [16].

13) Cf. [18].

14) Cf. [11], p. 407 and [4], p. 34.

finability. As illustrated in the example of multiplication, the greater risk does not imply less success.

The one-model method is essentially stronger than Padoa's method because of Lemma 3. By this lemma it can be decided that a non-logical constant d is not only undefinable explicitly from other non-logical constants $c_1 \dots c_k$ with respect to a certain incomplete theory T, but even essentially undefinable i.e. not only undefinable explicitly with respect to theory T itself but also with respect to every consistent extension of T. Thus the method constitutes an analogy between undefinability with respect to a theory and undecidability of a theory [15]). It will be shown in III, 3 that Lemma 3 can be used in practice. In III, 4 we shall also consider more closely the analogy mentioned. First we deal more in particular with the indirect method in proofs of undefinability. The reader acquainted with Tarski's indirect method in proofs of undecidability will recognize the analogy in question.

II, 4. *Indirect method*

The Lemmas 1, 2 and 3 constitute a direct method to obtain a negative solution of the definition problem for a non-logical constant d with respect to a theory T. This direct method can be combined with an indirect method, which consists in reducing the definition problem for d with respect to a theory T_2 to the definition problem for d with respect to some other theory T_1, for which the problem has previously been solved.

If d is essentially undefinable explicitly from $c_1 \dots c_k$ with respect to a theory T, then d can be defined explicitly and validly from $c_1 \dots c_k$ neither with respect to theory T itself, nor with respect to any consistent extension of T. On the other hand, let T contain the non-logical constants $c_1 \dots c_k$, d and let d be undefinable explicitly from $c_1 \dots c_k$ with respect to theory T, then d is undefinable explicitly from $c_1 \dots c_k$ with respect to every subtheory of T containing d. We might refer to this property by saying that a non-logical constant d, which is undefinable explicitly from other non-logical constants $c_1 \dots c_k$ with respect to a theory T, is hereditarily undefinable explicitly from $c_1 \dots c_k$ with respect to T [16]).

[15]) Cf. [22] passim; "essentially undecidable theory", p. 14.

[16]) The terms "essentially undefinable" and "hereditarily undefinable"

Let T_1 and T_2 be consistent theories containing the non-logical constants $c_1 \dots c_k$ and d, and let the latter be an extension of the former. Assume d to be definable explicitly and compatibly from $c_1 \dots c_k$ with respect to theory T_2. Then d is definable explicitly and compatibly from $c_1 \dots c_k$ with respect to theory T_1 also. For a sentence compatible with a theory is a fortiori compatible with a subtheory of it. Conversely, if d is essentially undefinable explicitly from $c_1 \dots c_k$ with respect to T_1, then d is essentially undefinable explicitly from $c_1 \dots c_k$ with respect to T_2. Likewise it follows from the notions involved: if d is definable explicitly and validly from $c_1 \dots c_k$ with respect to T_2, then d is definable explicitly and compatibly from $c_1 \dots c_k$ with respect to T_1; if d is undefinable explicitly from $c_1 \dots c_k$ with respect to T_2, then d is undefinable explicitly from $c_1 \dots c_k$ with respect to T_1; if d is essentially undefinable explicitly from $c_1 \dots c_k$ with respect to T_2, then d is undefinable explicitly from $c_1 \dots c_k$ with respect to T_1; and if d is definable explicitly and validly from $c_1 \dots c_k$ with respect to T_1, then d is definable explicitly and validly from $c_1 \dots c_k$ with respect to T_2. To the most practical of these statements we refer as:

Lemma 4: If a theory T_1 is a subtheory of a consistent theory T_2, both containing the non-logical constants $c_1 \dots c_k$, d, and if d is essentially undefinable explicitly from $c_1 \dots c_k$ with respect to T_1, then d is essentially undefinable explicitly from $c_1 \dots c_k$ with respect to T_2.

Let T_1 and T_2 be two consistent theories, both containing the non-logical constants $c_1 \dots c_k$ and d. Let T_1 be compatible with T_2. If d is definable explicitly and validly from $c_1 \dots c_k$ with respect to T_2, then d is definable explicitly and compatibly from $c_1 \dots c_k$ with respect to T_1. Conversely, if d is essentially undefinable explicitly from $c_1 \dots c_k$ with respect to T_2, then d is undefinable explicitly from $c_1 \dots c_k$ with respect to T_1. We refer to this statement as:

---

as used here are analogous to these in [22] of "essentially undecidable" and "hereditarily undecidable". "Undefinable" applies to a non-logical constant with respect to a theory, whereas "undecidable" applies to a theory; the aspects of extensionality and intensionality are the same. Cf. footnote 4 of this chapter.

Lemma 5: Let T_1 and T_2 be two compatible theories, both containing the non-logical constants $c_1 \dots c_k$, d and let d be essentially undefinable explicitly from $c_1 \dots c_k$ with respect to T_1, then d is undefinable explicitly from $c_1 \dots c_k$ with respect to T_2.

Let T_1 contain the non-logical constants $c_1 \dots c_k$, d and let T_1 be a subtheory of a consistent theory T_2. If d is undefinable explicitly from $c_1 \dots c_k$ with respect to T_1, it does not follow that d is undefinable explicitly from $c_1 \dots c_k$ with respect to T_2, for an explicit definition D which is not valid in T_1, might be valid in T_2. If, however, T_2 contains with $c_1 \dots c_k$, d other non-logical constants $c_{k+1} \dots c_{k+m}$ and if T_1 contains all sentences of T_2 in which $c_{k+1} \dots c_{k+m}$ do not occur, the situation is different: an explicit definition D of d from $c_1 \dots c_k$ which is valid in T_2, belongs to the valid sentences of T_1. Hence we state:

Lemma 6: If a theory T_1 containing the non-logical constants $c_1 \dots c_k$, d is a subtheory of a consistent theory T_2 containing the non-logical constants $c_1 \dots c_k$, d, $c_{k+1} \dots c_{k+m}$, and if T_1 contains all sentences of T_2 in which $c_{k+1} \dots c_{k+m}$ do not occur, whereas d is undefinable explicitly from $c_1 \dots c_k$ with respect to theory T_1, then d is undefinable explicitly from $c_1 \dots c_k$ with respect to theory T_2.

CHAPTER III

ADDITION AND MULTIPLICATION IN THE FORMALIZED ARITHMETIC OF NATURAL NUMBERS

III, 1. *Introductory remarks*

We consider the arithmetic of natural numbers formalized with standard formalization and with only five non-logical constants: an individual constant "1", a unary operation constant "S", a binary relation constant "$>$" and two binary operation constants "$+$" and "$\cdot$". Hence every possible realization of this theory is a system R, or $\langle \mathbf{X}; \mathbf{i}, \mathbf{o}_1, \mathbf{r}, \mathbf{o}_2, \mathbf{o}_3 \rangle$, in which $\mathbf{X}$ is an arbitrary set, $\mathbf{i}$ is an element of $\mathbf{X}$, $\mathbf{o}_1$ is a unary operation on $\mathbf{X}$ to $\mathbf{X}$, $\mathbf{r}$ is a binary relation in $\mathbf{X}$ and $\mathbf{o}_2$ and $\mathbf{o}_3$ are binary operations on $\mathbf{X} \times \mathbf{X}$ to $\mathbf{X}$. To define the validity in this theory we consider that special realization $\langle \mathbf{N}; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{+}, \boldsymbol{\cdot} \rangle$, where all the symbols have their usual arithmetical meaning; the elements of $\mathbf{N}$ will be denoted as *regular* natural numbers. A sentence is said to be valid in this theory of arithmetic of natural numbers if and only if it holds in $\langle \mathbf{N}; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{+}, \boldsymbol{\cdot} \rangle$. We refer to this theory as $\mathsf{T}(\cdot)$.

$\mathsf{T}(\cdot)$ is complete since the set of all valid sentences of $\mathsf{T}(\cdot)$ coincides with the set of sentences which are satisfied in a single model. However, theory $\mathsf{T}(\cdot)$ is known to be essentially undecidable and not axiomatizable [1]).

Let $\mathsf{T}(+)$ be the subtheory of $\mathsf{T}(\cdot)$ containing all sentences of the latter in which "$\cdot$" does not occur, a sentence being valid in $\mathsf{T}(+)$ if and only if it is valid in $\mathsf{T}(\cdot)$. Obviously theory $\mathsf{T}(+)$ is complete. Presburger [2]) has shown that the complete theory of formalized arithmetic of integers, formalized with standard

[1]) Cf. e.g. [22], p. 60, Theorem 9 (reading theory $\mathsf{T}(\cdot)$ for theory N), in connection with p. 14, Theorem 1. Theory N of [22] has "0" as a non-logical constant instead of our "1", but that can be adapted easily. Further, theory N does without "$>$", but "$>$" can be introduced by an explicit definition. The same applies to "S".

[2]) Cf. [16], p. 395 and pp. 92 ff.

formalization and containing no other non-logical constants than "0", "1", "+" and ">" is axiomatizable and decidable. In fact he gives a recursive set of axioms in the symbolism of this theory and shows that the theory based on these axioms is complete and decidable. It is possible to adapt the axiom system of Presburger to an axiom system of T(+) by well-known methods. It follows that T(+) is decidable and hence axiomatizable.

For our purpose it is sufficient to know that

i) T(+) is complete and axiomatizable;

ij) T(+) has *the closure property* for "S" and "+".

The latter of these statements needs some explication. Let $\langle \mathbf{U}; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{+}\rangle$ be an arbitrary model of T(+), **U** being the universe and **1**, **S**, **>**, **+** the interpretations of "1", "S", ">", "+" respectively. Let **i** be an arbitrary element of **U** and let **U′** be the closure of {**i**} under the operations **S** and **+**. We shall prove that for every model $\langle \mathbf{U}; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{+}\rangle$ of T(+) and for every element **i** of **U**, $\langle \mathbf{U}'; \mathbf{i}, \mathbf{S}, \mathbf{>}, \mathbf{+}\rangle$ is also a model of T(+) [3]. We refer to this fact by saying that T(+) has the closure property for "S" and "+". Later (cf. IV, 2, p. 51) we shall deal more in general with this property. Here we confine ourselves to what is needed in this chapter. It can be proved in an easy way that T(+) has the

---

[3] **U′** *is the closure of* {**i**} *under the operations* **S** *and* **+** means, that **U′** contains with **i** no more elements of **U** than those obtainable from **i** by (repeated) application of the operations **S** and **+**.

Among the sentences describing the generating properties of **S** and **+** we reckon here the sentences:

$$(x)(Ey)(x > 1 \rightarrow Sy = x),$$
$$(x)(y)(Ez)(x > y \rightarrow x = y + z).$$

Together with the sentences:

$$(x)(y)(z)((Sy = x \ \& \ Sz = x) \rightarrow y = z),$$
$$(x)(y)(z)(u)((x = y + z \ \& \ x = y + u) \rightarrow z = u),$$

they justify the operation constants "–1" and "–", which could be introduced into T(+) by the explicit definitions:

$$(x)(y)(x - 1 = y \leftrightarrow Sy = x),$$
$$(x)(y)(z)(x - y = z \leftrightarrow x = y + z).$$

We do not add these explicit definitions to T(+), but we wish to stress that in this context the closure of {i} under the operations **S** and **+** includes the operations **—1** and **—** in their familiar arithmetical meaning.

closure property for "S" and "$+$". It is sufficient to show that each axiom of one chosen complete axiom system for $\mathsf{T}(+)$ holds as well in a model thus "deflated" as it does in the original model. Without going into details we argue that e.g. Presburger's axiom system adapted for $\mathsf{T}(+)$ satisfies this condition. Presburger's axioms are complete and form a recursive set of which each axiom has either the shape:

$$(x) \ldots (y)V,$$

or the shape:

$$(x) \ldots (y)(Ez) \ldots (Eu)V,$$

where V is quantifier-free. The transformation to an axiom system for $\mathsf{T}(+)$ affects neither these shapes nor the recursiveness. Thus it can be verified in an easy way that each axiom contains only conditions which are satisfied in the closure mentioned. An axiom of the shape:

$$(x) \ldots (y)V,$$

holding for the original model holds automatically for the deflated model (a condition holding for all elements of a set holds for some elements of it); an axiom of the shape:

$$(x) \ldots (y)(Ez) \ldots (Eu)V,$$

holds in the deflated model, since the axioms of this kind of the system mentioned describe exactly the generating properties of $\mathbf{S}$ and $\mathbf{+}$, and nothing more [3]).

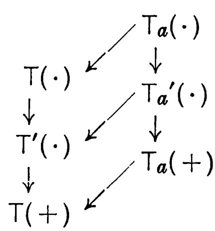
We shall also consider a theory $\mathsf{T}'(\cdot)$, an axiomatizable subtheory of $\mathsf{T}(\cdot)$ and an extension of $\mathsf{T}(+)$ with some axioms concerning "$\cdot$". We shall see that the new axioms concerning "$\cdot$" are such, that $\mathsf{T}'(\cdot)$ has the closure property for "S", "$+$" and "$\cdot$". In this case this means that for every model $\langle \mathbf{U}; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{+}, \mathbf{\cdot} \rangle$ of $\mathsf{T}'(\cdot)$ and for every element $\mathbf{i}$ of $\mathbf{U}$, $\langle \mathbf{U}'; \mathbf{i}, \mathbf{S}, \mathbf{>}, \mathbf{+}, \mathbf{\cdot} \rangle$ is also a model of $\mathsf{T}'(\cdot)$, where $\mathbf{U}'$ is the closure of $\{\mathbf{i}\}$ under the operations $\mathbf{S}$, $\mathbf{+}$ and $\mathbf{\cdot}$.

Furthermore, theory $\mathsf{T}(+)$ is extended to a theory $\mathsf{T}_a(+)$ by adding to the axioms of $\mathsf{T}(+)$ the following recursive set of axioms containing a new individual constant "a":

(1a)	$a > 1$;
(2a)	$a > S1$;
(3a)	$a > S(S1)$;
....	....
....	.........[4].

Obviously $\mathsf{T}_a(+)$ is axiomatizable. Moreover, from the fact that $\mathsf{T}(+)$ has the closure property for "S" and "$+$", it follows that $\mathsf{T}_a(+)$ has the closure property for "S" and "$+$". In this case this means that for every model $\langle \mathbf{U}; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{+}, \mathbf{a} \rangle$ of $\mathsf{T}_a(+)$ and for every two elements $\mathbf{i}_1$ and $\mathbf{i}_2$ of $\mathbf{U}$, which are suitable interpretations of "1" and "a" respectively (i.e. $\mathbf{i}_2 \mathbf{>} \mathbf{i}_1$, $\mathbf{i}_2 \mathbf{>} \mathbf{S}\mathbf{i}_1$, $\mathbf{i}_2 \mathbf{>} \mathbf{S}(\mathbf{S}\mathbf{i}_1), \ldots$) $\langle \mathbf{U}'; \mathbf{i}_1, \mathbf{S}, \mathbf{>}, \mathbf{+}, \mathbf{i}_2 \rangle$ is also a model of $\mathsf{T}_a(+)$, where $\mathbf{U}'$ is the closure of $\{\mathbf{i}_1, \mathbf{i}_2\}$ under the operations $\mathbf{S}$ and $\mathbf{+}$. In proving that $\mathsf{T}_a(+)$ has the closure property for "S" and "$+$" we can apply the same argument as for $\mathsf{T}(+)$, since the axioms (1a), (2a), (3a), ... have no influence.

In the same way we extend $\mathsf{T}'(\cdot)$ to $\mathsf{T}_a'(\cdot)$ by adding the axioms (1a), (2a), (3a), ... to the axioms of $\mathsf{T}'(\cdot)$. Obviously $\mathsf{T}_a'(\cdot)$ is axiomatizable and in the same way as expounded above for $\mathsf{T}_a(+)$, it follows from the fact that $\mathsf{T}'(\cdot)$ has the closure property for "S", "$+$" and "$\cdot$", that $\mathsf{T}_a'(\cdot)$ has this property also (in the sense of $\mathsf{T}_a(+)$, reading $\mathbf{S}$, $\mathbf{+}$ and $\cdot$ instead of $\mathbf{S}$ and $\mathbf{+}$). Further we might extend in the same way (in order to complete the scheme, although we do not need it) $\mathsf{T}(\cdot)$ to $\mathsf{T}_a(\cdot)$ by adding (1a), (2a), (3a), ... as valid sentences to the valid sentences of $\mathsf{T}(\cdot)$. Obviously $\mathsf{T}_a(\cdot)$ is not axiomatizable; later we shall see that $\mathsf{T}(\cdot)$ has not the closure property for "S", "$+$" and "$\cdot$" (cf. IV, 1, p. 49), neither has $\mathsf{T}_a(\cdot)$. We obtain the following scheme:



[4]) According to a recent paper of J. G. Kemeny the method to add sentences

$$a \neq 1,\ a \neq 2,\ a \neq 3,\ \ldots$$

to a theory of this kind is due to L. A. Henkin. Cf. [10], p. 164.

Every theory in this scheme is an extension of the theories which can be reached from it following the arrows. For our purpose (to show the independence of "$\cdot$" from "1", "S", "$>$" and "$+$" with respect to the theories $\mathsf{T}'(\cdot)$ and $\mathsf{T}(\cdot)$, cf. III, 3) we need only the five theories

$$\mathsf{T}(\cdot),\ \mathsf{T}'(\cdot),\ \mathsf{T}(+),\ \mathsf{T}_a'(\cdot) \text{ and } \mathsf{T}_a(+),$$

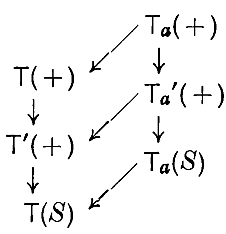
related as:

$$\mathsf{T}(+) \subset \mathsf{T}'(\cdot) \subset \mathsf{T}(\cdot),$$
$$\mathsf{T}(+) \subset \mathsf{T}'(\cdot) \subset \mathsf{T}_a'(\cdot),$$
$$\mathsf{T}(+) \subset \mathsf{T}_a(+) \subset \mathsf{T}_a'(\cdot).$$

In these theories there occur no other individual constants than "1" and "a", no other compound terms (cf. I, 1, p. 1) than those formed from atomic terms by (repeated) application of the unary operation constant "S" and the binary operation constants "$+$" and "$\cdot$", and no other atomic formulas than those obtained by combining two arbitrary terms by means of "$=$" (throughout considered as a logical constant) or the binary relation constant "$>$".

$\mathsf{T}(+)$ is consistent, since it has the model $\langle \mathbf{N}; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{+}\rangle$. So $\mathsf{T}_a(+)$ is consistent, since any finite subtheory of it has the model $\langle \mathbf{N}; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{+}, \mathbf{a}_0\rangle$ (where $\mathbf{a}_0$ belongs to $\mathbf{N}$) and is consistent. $\mathsf{T}(\cdot)$ and $\mathsf{T}'(\cdot)$ are consistent, since both have the model $\langle \mathbf{N}; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{+}, \boldsymbol{\cdot}\rangle$. So $\mathsf{T}_a'(\cdot)$ is consistent, since any finite subtheory has the model $\langle \mathbf{N}; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{+}, \boldsymbol{\cdot}, \mathbf{a}_0\rangle$ (where $\mathbf{a}_0$ belongs to $\mathbf{N}$) and is consistent.

Starting with $\mathsf{T}(+)$ we can form in a similar way another scheme of theories:



Every theory in this scheme is an extension of the theories which can be reached from it following the arrows. For our purpose (to show the independence of "+" from "1", "S" and ">", cf. III, 2) we need only the five theories

$$\mathsf{T}(+),\ \mathsf{T}'(+),\ \mathsf{T}(S),\ \mathsf{T}_a'(+) \text{ and } \mathsf{T}_a(S),$$

related as:

$$\mathsf{T}(S) \subset \mathsf{T}'(+) \subset \mathsf{T}(+),$$
$$\mathsf{T}(S) \subset \mathsf{T}'(+) \subset \mathsf{T}_a'(+),$$
$$\mathsf{T}(S) \subset \mathsf{T}_a'(S) \subset \mathsf{T}_a'(+).$$

$\mathsf{T}(+)$ is as expounded above. $\mathsf{T}(S)$ is the subtheory of $\mathsf{T}(+)$ containing all sentences of the latter in which "+" does not occur, a sentence being valid in $\mathsf{T}(S)$ if and only if it is valid in $\mathsf{T}(+)$. $\mathsf{T}(S)$ has the properties i) and ij) of p. 26 reading $\mathsf{T}(S)$ for $\mathsf{T}(+)$ and **S** for **S** and **+**. $\mathsf{T}'(+)$ is a subtheory of $\mathsf{T}(+)$ and an extension of $\mathsf{T}(S)$. For the main purpose of III, 2 we could dispense with $\mathsf{T}'(+)$ and work instead immediately with $\mathsf{T}(+)$; but for the unity of treatment, for clearness' sake in connection with the analogy to Tarski's undecidability and in order to show the strength of the method, it is useful to apply the extra link of $\mathsf{T}'(+)$. $\mathsf{T}_a'(+)$ and $\mathsf{T}_a(S)$ are obtained by adding to the axioms of these theories the recursive set of axioms (1a), (2a), (3a), ... Since $\mathsf{T}(S)$ has the closure property for "S", $\mathsf{T}_a(S)$ has this property also, as argued for $\mathsf{T}(+)$ and $\mathsf{T}_a(+)$.

As the latter scheme of theories is a continuation downwards of the former, it is obvious that all theories under consideration in the latter are consistent.

About constants, terms and formulas in the theories of the latter scheme, cf. what is said for the theories of the former scheme, reading "+" instead of "+" and "·".

In order to fix the thoughts and to refer in an easy way we give first a list of some (well-known [5])) valid sentences of $\mathsf{T}(\cdot)$:

i) sentences with no other non-logical constants than ">":

(1>) $(x)(y)(\overline{x=y} \leftrightarrow (x>y \vee y>x))$;

(2>) $(x)(y)(\overline{x>y} \leftrightarrow (x=y \vee y>x))$;

(3>) $(x)(y)(z)((x>y \mathbin{\&} y>z) \rightarrow x>z)$;

[5]) All these sentences can be interpreted as well-known arithmetical theorems.

ij) sentences with no other non-logical constants than "1", "S" and "$>$":

(1S) $(x)(Ey)(Sx=y)$;

(2S) $(x)(y)(z)((Sx=y \;\&\; Sx=z) \rightarrow y=z)$;

(3S) $(x)(Ey)(x>1 \rightarrow Sy=x)$;

(4S) $(x)(y)(z)((Sy=x \;\&\; Sz=x) \rightarrow y=z)$;

(5S) $(x)(\overline{x=1} \rightarrow x>1)$;

(6S) $(x)(Sx>x)$;

(7S) $(x)(y)(x>y \rightarrow Sx>Sy)$;

(8S) $(x)(y)(z)((x>y \;\&\; y>z) \rightarrow x>Sz)$;

iij) sentences with no other non-logical constants than "1", "S", "$>$" and "$+$":

(1+) $(x)(y)(Ez)(x+y=z)$;

(2+) $(x)(y)(z)(u)((x+y=z \;\&\; x+y=u) \rightarrow z=u)$;

(3+) $(x)(x+1=Sx)$;

(4+) $(x)(y)(x+Sy=S(x+y))$;

(5+) $(x)(y)((x>1 \;\&\; y>1) \rightarrow (x+y>Sx \;\&\; x+y>Sy))$;

(6+) $(x)(y)(z)(x>y \rightarrow (x+z>y+z))$;

(7+) $(x)(y)(z)(y>z \rightarrow (x+y>x+z))$;

(8+) $(x)(y)(z)(x+(y+z)=(x+y)+z)$;

iv) sentences with no other non-logical constants than "1", "S", "$>$", "$+$" and "$\cdot$":

(1·) $(x)(y)(Ez)(x\cdot y=z)$;

(2·) $(x)(y)(z)(u)((x\cdot y=z \;\&\; x\cdot y=u) \rightarrow z=u)$;

(3·) $(x)(x\cdot 1=x)$;

(4·) $(x)(y)(x\cdot Sy=x+(x\cdot y))$;

(5·) $(x)(y)((x>1 \;\&\; y>1) \rightarrow (x\cdot y>Sx \;\&\; x\cdot y>Sy))$;

(6·) $(x)(y)(z)(x>y \rightarrow x\cdot z>y\cdot z)$;

(7·) $(x)(y)(z)(y>z \rightarrow x\cdot y>x\cdot z)$;

(8·) $(x)(y)(z)(x\cdot(y\cdot z)=(x\cdot y)\cdot z)$.

It is practical to remark that from the sentences (1>), (2>), (3>) and (1S) ... (8S) the sentences:

$$(x)((x>1 \to U(x)) \leftrightarrow U(Sx));$$
$$(x)((x>\mathrm{S}1 \to U(x)) \leftrightarrow (x>1 \to U(Sx)));$$
$$\cdots\cdots\cdots\cdots\cdots$$

are derivable in the theories under consideration. These equivalences provide us with practical rules of inference for the arithmetical theories concerned:

(R) $$\frac{(x)(x>1 \to U(x))}{(x)\,U(Sx)} \quad \text{and conversely,}$$

$$\frac{(x)(x>S1 \to U(x))}{(x)(x>1 \to U(Sx))} \quad \text{and conversely,}$$

and so on.

We refer to these rules as (R).

III, 2. *Undefinability of "+" from "1" and "S"*

Proceeding to the definition problem for "+" from "1", "S" and ">" with respect to theory $\mathsf{T}(+)$ we consider the theories

$$\mathsf{T}(S),\ \mathsf{T}_a(S),\ \mathsf{T}'(+),\ \mathsf{T}_a'(+) \text{ and } \mathsf{T}(+),$$

as mentioned above [6]).

$\mathsf{T}(S)$ is known to be finitely axiomatizable [7]). For our purpose it is of no interest which axiom system in particular is chosen for $\mathsf{T}(S)$, provided it is complete, contains the non-logical constants "1", "S" and ">" and has $\langle \mathbf{N}; \mathbf{1}, \mathbf{S}, \mathbf{>}\rangle$, where all the symbols have their familiar arithmetical meaning, as a model. In every system the sentences (1>), (2>), (3>), (1S) ... (8S) are provable, because they are valid in $\mathsf{T}(S)$. $\mathsf{T}(S)$ has the closure property for "S", as can be shown in a similar way as for $\mathsf{T}(+)$ (cf. III, 1, p. 26, dropping "+" and $\mathbf{+}$).

---

6) In this particular case of $\mathsf{T}(+)$ the method could be applied in a less complicated way (cf. III, 1, p. 30). However, the scheme followed here can be applied in more cases as will be shown in III, 3 and IV.

7) Cf. e.g. [3], pp. 57 ff. E. W. Beth's theory M is a theory of integers, but can be adapted.

$\mathsf{T}'(+)$ is obtained by adding as axioms to the axioms of $\mathsf{T}(S)$ the sentences (1+), (2+), (3+), (4+) and (7+) [8]).

Theory $\mathsf{T}_a'(+)$ contains a.o. the following provable sentences:

(1a+)	$a+a>Sa$;
(2a+)	$a+a>S(Sa)$;
(3a+)	$a+a>S(S(Sa))$;
....	....
....	....

Proof: $a+a>a+1$, (1a) and (7+);
$a+1=Sa$, (3+);
so $a+a>Sa$.

Further,

$a+a>a+S1$, (2a) and (7+);
$a+S1=S(a+1)$, (4+);
$S(a+1)=S(Sa)$, (3+);
so $a+a>S(Sa)$.

Similarly for (3a+) and so on.

Since $\mathsf{T}_a(S)$ is consistent it has a model. This model contains interpretations for "1", "S", "$>$" and "a", which we designate by **1**, **S**, $\boldsymbol{>}$ and **a**. We consider only the closure of {**1**, **a**} under the operation **S**. The universe thus deflated we designate by $\mathbf{A}^S$. $\langle \mathbf{A}^S; \mathbf{1}, \mathbf{S}, \boldsymbol{>}, \mathbf{a}\rangle$ with the universe thus deflated is still a model of $\mathsf{T}_a(S)$ for the reasons mentioned in III, 1, p. 30, ($\mathsf{T}_a(S)$ has the closure property for "S"). $\mathbf{A}^S$ can be represented as:

$$\mathbf{1}, \mathbf{2}, \mathbf{3}, \ldots\ldots \quad \ldots\ldots \mathbf{a}-\mathbf{2},\ \mathbf{a}-\mathbf{1},\ \mathbf{a},\ \mathbf{a}+\mathbf{1},\ \mathbf{a}+\mathbf{2}, \ldots\ldots$$

where we borrow the symbols **2**, **3**, ... and **—** from the familiar arithmetic; in $\langle \mathbf{A}^S; \mathbf{1}, \mathbf{S}, \boldsymbol{>}, \mathbf{a}\rangle$ the symbols **1**, **S** and $\boldsymbol{>}$ have

---

[8]) The reason why the list of p. 31 contains also the sentences (5 +), (6 +) and (8 +), not needed here, is a matter of uniformity with IV, 1. The sentence (2 +) is not used directly in the arguments of this section, but it is needed as one of the characteristic sentences for an operation constant, cf. I, 3, p. 5. The axiom system for $\mathsf{T}'(+)$ as given here, is not independent, e.g. "S" can be defined explicitly and validly from "1" and "+" with respect to $\mathsf{T}'(+)$ and thus be eliminated ("S" is "+ 1"). This fact is of no interest for our purpose.

their usual arithmetical meaning and the ordering of $\mathbf{A}^S$ is given by the order in which the elements are written down from left to right. $\mathbf{A}^S$ contains two enumerable subsets $\mathbf{A}_1{}^S$ and $\mathbf{A}_2{}^S$, $\mathbf{A}_1{}^S$ having the order-type of the natural numbers, $\mathbf{A}_2{}^S$ having the order-type of the integers; $\mathbf{n}_a \mathbf{>} \mathbf{n}$ holds between every $\mathbf{n}_a$ belonging to $\mathbf{A}_2{}^S$ and every $\mathbf{n}$ belonging to $\mathbf{A}_1{}^S$.

$\mathbf{A}^S$ does not contain an element $\mathbf{a}^*$ with the properties:

$$\mathbf{a}^* \mathbf{>} \mathbf{a};$$
$$\mathbf{a}^* \mathbf{>} \mathbf{Sa};$$
$$\mathbf{a}^* \mathbf{>} \mathbf{S(Sa)};$$
$$\cdots\cdots$$

since this $\mathbf{a}^*$ would exceed all elements given by the closure of $\{\mathbf{1}, \mathbf{a}\}$ under the operation $\mathbf{S}$.

The model $\langle \mathbf{A}^S; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{a}\rangle$ is also a model of $\mathsf{T}(S)$, since this theory is a subtheory of $\mathsf{T}_a(S)$.

Theorem 1: The non-logical constant "+" is undefinable explicitly and essentially undefinable explicitly from the non-logical constants "1", "S" and "$>$" with respect to theory $\mathsf{T}'(+)$.

Proof: We apply Lemma 1 to $\mathsf{T}(S)$ and $\mathsf{T}'(+)$ with $\langle \mathbf{A}^S; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{a}\rangle$ as a model of the former. In no way into this model can there be introduced a notion $\mathbf{+}$ as an interpretation of the "+" of $\mathsf{T}'(+)$.

For such an introduction would imply that $\langle \mathbf{A}^S; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{+}, \mathbf{a}\rangle$ would be a model of $\mathsf{T}_a'(+)$. But according to (1+) $\mathbf{A}^S$ would then contain an element $\mathbf{a} + \mathbf{a}$; according to (1a+), (2a+), (3a+), ... this element would have the properties:

$$\mathbf{a} + \mathbf{a} \mathbf{>} \mathbf{Sa};$$
$$\mathbf{a} + \mathbf{a} \mathbf{>} \mathbf{S(Sa)};$$
$$\mathbf{a} + \mathbf{a} \mathbf{>} \mathbf{S(S(Sa))};$$
$$\cdots\cdots\cdots$$

As we have seen $\mathbf{A}^S$ does not contain such an element $\mathbf{a}^*$. So "+" is undefinable explicitly from "1", "S" and "$>$" with respect to theory $T'(+)$. Moreover, $T(S)$ is complete and so by Lemma 3 we obtain that "+" is essentially undefinable explicitly from "1", "S" and ">" with respect to $T'(+)$.

Theorem 2: The non-logical constant "+" is essentially undefinable explicitly from the non-logical constants "1", "S" and ">" not only with respect to $T'(+)$, but also with respect to every consistent extension of $T'(+)$, in particular with respect to $T(+)$.

Proof: The theorem is an immediate consequence of Theorem 1 and Lemma 4.

Theorem 3: The non-logical constant "+" is undefinable explicitly from "1", "S" and ">" with respect to every subtheory of $T(+)$ containing "1", "S", ">" and "+". The same holds with respect to every consistent theory which is compatible with $T'(+)$ and which contains the mentioned non-logical constants.

Proof: The first part of the theorem follows from the remarks preceding Lemma 4. The second part results from Lemma 5.

Corollary: What is said about "+" being (essentially) undefinable explicitly from "1", "S" and ">" with respect to the theories mentioned in Theorems 1, 2 and 3, applies a fortiori to the (essential) undefinability of "+" from "1" and "S" with respect to the same theories.

Corollary: If ">" is definable explicitly and validly from "1", "S" and "+" with respect to a theory mentioned in the Theorems 2 and 3, then "+" is (essentially) undefinable not only from "1", "S" and ">" with respect to this theory, but also from "1" and "S" with respect to the subtheory obtained from this theory by dropping all sentences containing ">".

The corollaries are trivial. Since ">" is definable explicitly and validly from "+" with respect to $T(+)$ [9]), the corollaries justify the title of this section in a double aspect.

III, 3. *Undefinability of "·" from "1" and "+"*

Proceeding to the definition problem for "·" from "1", "S",

[9]) $(x)(y)(x > y \leftrightarrow (Ez)(y + z = x))$.

">" and "+" with respect to theory $\mathsf{T}(\cdot)$ we consider the theories

$$\mathsf{T}(+),\ \mathsf{T}_a(+),\ \mathsf{T}'(\cdot),\ \mathsf{T}_a'(\cdot) \text{ and } \mathsf{T}(\cdot),$$

as mentioned above.

$\mathsf{T}(+)$ is known to be finitely axiomatizable [10]. For our purpose it is of no interest which axiom system in particular is chosen for $\mathsf{T}(+)$, provided it is complete, contains the non-logical constants "1", "S", ">" and "+" ("S" may be included in "+" as "+1"), and has $\langle \mathbf{N}; \mathbf{1}, \mathbf{S}, \boldsymbol{>}, \boldsymbol{+}\rangle$ as a model. In every system the sentences (1>) .... (3>), (1S) .... (8S), (1+) .... (8+) are provable because they are valid in $\mathsf{T}(+)$. $\mathsf{T}(+)$ has the closure property for "S" and "+" as pointed out in III, 1, p. 26 and so has $\mathsf{T}_a(+)$. $\mathsf{T}'(\cdot)$ is obtained by adding as axioms to the axioms of $\mathsf{T}(+)$ the sentences (1·), (2·), (3·), (4·) and (5·) [11].

The following sentence is provable in $\mathsf{T}'(\cdot)$:

(9·) $\qquad (x)(y)((x>1 \ \&\ y>S1) \rightarrow x\cdot y>x+y).$

Proof: $(x)(y)(y>1 \rightarrow Sx\cdot y>Sy)$, (5·) and (R);
$(x)(y)(y>1 \rightarrow Sx+Sx\cdot y>Sx+Sy)$, (7+);
$(x)(y)(y>1 \rightarrow Sx\cdot Sy>Sx+Sy)$, (4·);
$(x)(y)((x>1 \ \&\ y>S1) \rightarrow x\cdot y>x+y)$, (R).

The following sentences are provable in theory $\mathsf{T}_a(\cdot)$:

(1a·) $\qquad a\cdot a>a+a;$

(2a·) $\qquad a\cdot a>a+(a+a);$

(3a·) $\qquad a\cdot a>a+(a+(a+a));$

.... $\qquad$ . . . .

....

---

10) In III, 1 we quoted M. Presburger's infinite axiomatization (for historical reasons). The argument which E. W. Beth gives for a theory of integers with "1", "S" and ">" can also be applied here: the decision method involves only a finite number of non-logical principles, hence the theory is finitely axiomatizable.

Cf. [3], pp. 57 ff.

11) The sentences (6·), (7·) and (8·) in the list of p. 31 are not needed here. They are mentioned in the list for the sake of uniformity with IV, 1. The axiom system for $\mathsf{T}'(\cdot)$ as given here is not independent, cf. footnote 8 of this chapter.

Proof: $a \cdot a > a + a$, (1a), (2a) and (9·).

Further, according to (3S) and (1a), (2a), (3a), ... we introduce auxiliary constants:

"$a-1$", "$a-2$", "$a-3$", ... characterized by

$$S(a-1) = a;$$
$$S(a-2) = a-1;$$
$$S(a-3) = a-2;$$
$$\ldots\ldots\ldots\ldots$$

(Obviously $a-1>1$, $a-2>1$, $a-3>1$, ...).

We then argue:

$a \cdot a = a + (a + a \cdot (a-2))$, (4·) twice;

$a \cdot (a-2) > a$, (5·);

$a \cdot a > a + (a + a)$, (7+) twice.

Similarly (3a·) can be proved, and so on.

Since $\mathsf{T}_a(+)$ is consistent, it has a model. This model contains interpretations for "1", "S", "$>$", "$+$" and "a", which we designate by **1**, **S**, **>**, **+** and **a**. We consider only the closure of {**1**, **a**} under the operations **S** and **+** as described in $\mathsf{T}(+)$. Since $\mathsf{T}_a(+)$ has the closure property as well as $\mathsf{T}(+)$, the model ⟨$\mathbf{A}^+$; **1**, **S**, **>**, **+**, **a**⟩, where $\mathbf{A}^+$ designates the deflated universe, is still a model of $\mathsf{T}_a(+)$. $\mathbf{A}^+$ contains enumerably many enumerable subsets and can be represented as:

1, **2**, **3**, ... **a − 2**, **a − 1**, **a**, **a + 1**, **a + 2**, ...

.... **2a − 2**, **2a − 1**, **2a**, **2a + 1**, **2a + 2**, ...

.... **3a − 2**, **3a − 1**, **3a**, **3a + 1**, **3a + 2**, ...

where we borrow the symbols **2**, **3**, ... and **−** from the familiar arithmetic and where we write **2a** for **a + a**, **3a** for **a + a + a** and so on as usual.

$\mathbf{A}^+$ does not contain an element $\mathbf{a}^*$ with the properties:

$$\mathbf{a}^* > \mathbf{a};$$
$$\mathbf{a}^* > \mathbf{a} + \mathbf{a};$$
$$\mathbf{a}^* > \mathbf{a} + \mathbf{a} + \mathbf{a};$$
$$\dots\dots\dots$$
$$\dots\dots\dots\dots$$

Proof: Every element of $\mathbf{A}^+$ is either one of the elements $\mathbf{a}, \mathbf{a} + \mathbf{a}, \mathbf{a} + \mathbf{a} + \mathbf{a}, \dots.$ or is exceeded by one of these elements. For we can replace first $-$ by $+$ wherever it occurs and then $\mathbf{k}$ by $\mathbf{a}$ wherever it occurs ($\mathbf{k}$ is one of the regular natural numbers $\mathbf{1}, \mathbf{2}, \mathbf{3}, \dots.$).

So $\mathbf{A}^+$ does not contain such an element $\mathbf{a}^*$.

The model $\langle \mathbf{A}^+; \mathbf{1}, \mathbf{S}, >, +, \mathbf{a} \rangle$ is also a model of $\mathsf{T}(+)$, since this theory is a subtheory of $\mathsf{T}_a(+)$:

Theorem 4: The non-logical constant "·" is undefinable explicitly and essentially undefinable explicitly from the non-logical constants "1", "*S*", ">" and "+" with respect to theory $\mathsf{T}'(\cdot)$.

Proof: We apply Lemma 1 to $\mathsf{T}(+)$ and $\mathsf{T}'(\cdot)$ with $\langle \mathbf{A}^+; \mathbf{1}, \mathbf{S}, >, +, \mathbf{a} \rangle$ as a model of the former. In no way into this model can there be introduced a notion $\cdot$ as an interpretation of the "·" in $\mathsf{T}'(\cdot)$. For then $\langle \mathbf{A}^+; \mathbf{1}, \mathbf{S}, >, +, \cdot, \mathbf{a} \rangle$ would be a model of $\mathsf{T}'(\cdot)$, and hence according to (1·) $\mathbf{A}^+$ would contain an element $\mathbf{a} \cdot \mathbf{a}$; however, according to (1a·), (2a·), (3a·), ... this element would have the properties:

$$\mathbf{a} \cdot \mathbf{a} > \mathbf{a} + \mathbf{a};$$
$$\mathbf{a} \cdot \mathbf{a} > \mathbf{a} + (\mathbf{a} + \mathbf{a});$$
$$\mathbf{a} \cdot \mathbf{a} > \mathbf{a} + (\mathbf{a} + (\mathbf{a} + \mathbf{a}));$$
$$\dots\dots\dots\dots$$
$$\dots\dots\dots\dots\dots$$

and as we have seen $\mathbf{A}^+$ does not contain such an element $\mathbf{a}^*$. So "·" is undefinable explicitly from "1", "*S*", ">" and "+" with respect to theory $\mathsf{T}'(\cdot)$. Moreover, $\mathsf{T}(+)$ is complete and so by Lemma 3 we obtain that "·" is essentially undefinable explicitly from "1", "*S*", ">" and "+" with respect to theory $\mathsf{T}'(\cdot)$.

Theorem 5: The non-logical constant "·" is essentially undefinable explicitly from the non-logical constants "1", "S", "$>$" and "$+$" not only with respect to $\mathsf{T}'(\cdot)$, but also with respect to every consistent extension of $\mathsf{T}'(\cdot)$, in particular with respect to $\mathsf{T}(\cdot)$.

Proof: The theorem is an immediate consequence of Theorem 4 and Lemma 4.

Theorem 6: The non-logical constant "·" is undefinable explicitly from "1", "S", "$>$" and "$+$" with respect to every subtheory of $\mathsf{T}(\cdot)$ containing "1", "S", "$>$" and "$+$". The same holds with respect to every consistent theory which is compatible with $\mathsf{T}'(\cdot)$ and contains the mentioned non-logical constants.

Proof: The first part of the theorem follows from the remarks preceding Lemma 4. The second part results from Lemma 5.

Corollary: What is said about "·" being (essentially) undefinable explicitly from "1", "S", "$>$" and "$+$" with respect to the theories mentioned in the Theorems 4, 5 and 6 applies a fortiori to the (essential) undefinability of "·" from "1" and "$+$" with respect to the same theories.

Corollary: If "S" and "$>$" are definable explicitly and validly from "1" and "$+$" with respect to a theory mentioned in the Theorems 4, 5 and 6, then "·" is (essentially) undefinable not only from "1", "S", "$>$" and "$+$" with respect to this theory, but also from "1" and "$+$" with respect to the subtheory obtained from this theory by dropping all sentences containing "S" and "$>$".

The corollaries are trivial. In fact the non-logical constants "S" and "$>$" are definable explicitly and validly from "1" and "$+$" with respect to $\mathsf{T}'(\cdot)$ and all theories which are consistent extensions of $\mathsf{T}'(\cdot)$. As to subtheories of $\mathsf{T}(\cdot)$ and theories compatible with $\mathsf{T}'(\cdot)$ (Theorem 6), it depends. The corollaries justify the title of this section.

III, 4. *Comments*

Since theory $\mathsf{T}(+)$ is finitely axiomatizable, theory $\mathsf{T}'(\cdot)$ is finitely axiomatizable. The strength of Theorem 4 lies in the fact that "·" is essentially undefinable explicitly from "1", "S", "$>$"

and "+" with respect to a finitely axiomatizable subtheory of $\mathsf{T}(\cdot)$.

The results in the Theorems 5 and 6 are in accordance with the facts that $\mathsf{T}'(\cdot)$ is essentially undecidable and so is every consistent extension of $\mathsf{T}'(\cdot)$, – that every subtheory of $\mathsf{T}(\cdot)$ having the same non-logical constants is undecidable, – that every theory compatible with $\mathsf{T}'(\cdot)$ and having the same non-logical constants is undecidable, – whereas $\mathsf{T}(+)$ and every subtheory of $\mathsf{T}(+)$ with the same non-logical constants is decidable [12]. Let T be a consistent theory containing the non-logical constants $c_1 \ldots c_k$, d and T_0 the subtheory of T containing all sentences of T in which d does not occur; if d is definable explicitly and validly from $c_1 \ldots c_k$ with respect to T, then d can be eliminated from T. It follows that if T_0 is decidable, T should be decidable as well.

At the same time the analogy between the undefinability of a non-logical constant from other non-logical constants with respect to a theory T and the undecidability of a theory T (mentioned in II, 3, p. 22), is illustrated by the following example.

Theory $\mathsf{T}'(\cdot)$ plays the same part in proving indirectly that "$\cdot$" is (essentially) undefinable explicitly from the other non-logical constants with respect to various theories as theory Q of [22] [13] plays in proving indirectly the (essential) undecidability of various theories.

It is beyond the scope of this work to consider some of these various theories in detail.

It results from the Theorems 2 and 3 that the mentioned analogy is real and not the effect of a hidden correlation. Theory $\mathsf{T}'(+)$ plays the same part in proving indirectly that "+" is (essentially) undefinable explicitly from the other non-logical constants as theory Q plays in proving indirectly the (essential) undecidability of various theories. However, the theories of the Theorems 2 and 3 are all decidable and the indirect proof of undefinability with the help of $\mathsf{T}'(+)$ has no connection with undecidability.

---

12) Cf. [22], pp. 16 ff.; pp. 60 ff.

13) Cf. [22], pp. 51 ff.

CHAPTER IV

RESULTS CONCERNING EQUATIONAL UNDEFINABILITY

IV, 1. *The Ackermann sequence*

Turning to a more general treatment we consider the arithmetic of natural numbers formalized with standard formalization and with $n+3$ non-logical constants: an individual constant "1", a unary operation constant "S", a binary relation constant "$>$" and n binary operation constants "p_1" ... "p_n". Hence, every possible realization of this arithmetical theory is a system R, or $\langle \mathbf{X}; \mathbf{i}, \mathbf{o}_1, \mathbf{r}, \mathbf{o}_2 \dots \mathbf{o}_{n+1} \rangle$, in which $\mathbf{X}$ is an arbitrary set, $\mathbf{i}$ is an element of $\mathbf{X}$, $\mathbf{o}_1$ is a unary operation on $\mathbf{X}$ to $\mathbf{X}$, $\mathbf{r}$ is a binary relation in $\mathbf{X}$ and $\mathbf{o}_2 \dots \mathbf{o}_{n+1}$ are n binary operations on $\mathbf{X} \times \mathbf{X}$ to $\mathbf{X}$. To define the validity in this theory we consider a special realization $\langle \mathbf{N}; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{p}_1 \dots \mathbf{p}_n \rangle$ where all the symbols have their usual arithmetical meaning, $\mathbf{p}_1 \dots \mathbf{p}_n$ coinciding with the sequence of Ackermann [1]) given by

$$\mathbf{p}_1(\mathbf{k}, \mathbf{1}) = \mathbf{Sk};$$
$$\mathbf{p}_1(\mathbf{k}, \mathbf{Sl}) = \mathbf{S}(\mathbf{p}_1(\mathbf{k}, \mathbf{l}));$$
$$\mathbf{p}_2(\mathbf{k}, \mathbf{1}) = \mathbf{k};$$
$$\mathbf{p}_2(\mathbf{k}, \mathbf{Sl}) = \mathbf{p}_1(\mathbf{k}, \mathbf{p}_2(\mathbf{k}, \mathbf{l}));$$

and for m larger than 2

$$\mathbf{p}_m(\mathbf{k}, \mathbf{1}) = \mathbf{p}_{m-1}(\mathbf{k}, \mathbf{1});$$
$$\mathbf{p}_m(\mathbf{k}, \mathbf{Sl}) = \mathbf{p}_{m-1}(\mathbf{k}, \mathbf{p}_m(\mathbf{k}, \mathbf{l})).$$

A sentence is said to be valid in this theory of arithmetic of natural numbers if and only if it holds in $\langle \mathbf{N}; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{p}_1 \dots \mathbf{p}_n \rangle$.

We refer to this theory as $\mathscr{T}_n$.

Obviously, $\mathbf{p}_1(\mathbf{k}, \mathbf{l})$ coincides with $\mathbf{k} + \mathbf{l}$, $\mathbf{p}_2(\mathbf{k}, \mathbf{l})$ with $\mathbf{k} \cdot \mathbf{l}$, $\mathbf{p}_3(\mathbf{k}, \mathbf{l})$ with $\mathbf{k} \exp \mathbf{l}$ or $\mathbf{k}^{\mathbf{l}}$, and so on in accordance with the sequence of Ackermann. Each theory under consideration here is assumed to

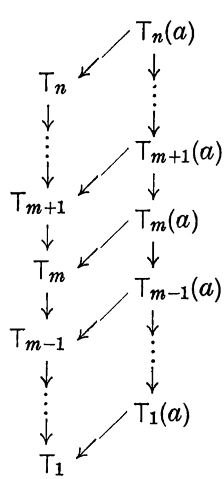
[1]) Cf. [1], pp. 119 ff.

contain a finite number of binary operation constants "$\mathbf{p}_1$" ... "$\mathbf{p}_n$".

$\mathcal{T}_n$ is complete since the set of all valid sentences of $\mathcal{T}_n$ coincides with the set of sentences which are satisfied in one single model. It is obvious that, for n larger than 1, $\mathcal{T}_n$ is undecidable and not axiomatizable [2]).

We shall consider axiomatic subtheories of $\mathcal{T}_n$ referred to as T_n, T_{n-1}, ... T_2, T_1. In this sequence each of the theories is a subtheory of the preceding ones and T_n is a subtheory of $\mathcal{T}_n$. All theories of the sequence contain the non-logical constants "1", "S" and "$>$" and further T_1 contains "p_1", T_2 contains "p_1" and "p_2", ..., T_{n-1} contains "p_1", "p_2", ... "p_{n-1}", and T_n contains all "p_1" ... "p_n". All theories T_n ... T_1 are axiomatic with a finite set of non-logical axioms, all have the closure property for the operation constants they contain, as described for T_1 (i.e. $\mathsf{T}(+)$) in III, 1, p. 26. T_{n-1} contains all sentences of T_n in which "p_n" does not occur, T_{n-2} all sentences of T_{n-1} in which "p_{n-1}" does not occur, ... T_1 all sentences of T_2 in which "p_2" does not occur.

We shall also consider the axiomatic theories $\mathsf{T}_n(a)$, $\mathsf{T}_{n-1}(a)$, ... $\mathsf{T}_2(a)$, $\mathsf{T}_1(a)$ of which respectively T_n, T_{n-1}, ... T_2, T_1 are subtheories. They are the extensions of these theories with the individual constant "a" and the recursive set of axioms (1a), (2a), (3a),... (cf. p. 28). They are built in such a way that they all have the closure property for the operation constants they contain, as described for $\mathsf{T}_1(a)$ (i.e. $\mathsf{T}_a(+)$) in III, 1, p. 28. Thus we obtain the following scheme:



2) For n larger than 1, $\mathcal{T}_n$ contains $\mathsf{T}(\cdot)$. Cf. [22], Corollary 10, p. 62.

Every theory in this scheme is an extension of the theories, which can be reached from it following the arrows.

About expressions, terms, formulas, sentences of these theories the same remarks hold as in III, 1, p. 29 with "p_1" ... "p_m" instead of "+" and "·". Parentheses and commas are used in the familiar way as technical symbols and could be dispensed with.

T_n has the model $\langle \mathbf{N}; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{p}_1 \dots \mathbf{p}_n \rangle$ and is consistent. So $\mathsf{T}_n(a)$ is consistent, since any finite subtheory of it has the model $\langle \mathbf{N}; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{p}_1 \dots \mathbf{p}_n, \mathbf{a}_0 \rangle$ (where $\mathbf{a}_0$ belongs to $\mathbf{N}$) and is consistent. So for all m (m over 1 ... n) $\mathsf{T}_m(a)$ is consistent, since $\mathsf{T}_m(a)$ is a subtheory of $\mathsf{T}_n(a)$ (cf. scheme).

The non-logical axioms of T_n are those of $\mathsf{T}(+)$ of III, 1, p. 26, somehow finitely axiomatized [3] (where "+" has to be read as "p_1"), the axioms $(1\cdot)$ ... $(8\cdot)$ (where "·" has to be read as "p_2"), and further for each "p_m" (m larger than 2 and not larger than n) the axioms:

(1m) $(x)(y)(Ez)(p_m(x, y) = z)$;

(2m) $(x)(y)(z)(u)((p_m(x, y) = z \ \& \ p_m(x, y) = u) \rightarrow z = u)$;

(3m) $(x)(p_m(x, 1) = p_{m-1}(x, 1))$;

(4m) $(x)(y)(p_m(x, Sy) = p_{m-1}(x, p_m(x, y)))$;

(5m) $(x)(y)((x > 1 \ \& \ y > 1) \rightarrow (p_m(x,y) > Sx \ \& \ p_m(x,y) > Sy))$;

(6m) $(x)(y)(z)(x > y \rightarrow p_m(x, z) > p_m(y, z))$;

(7m) $(x)(y)(z)((x > 1 \ \& \ y > z) \rightarrow p_m(x, y) > p_m(x, z))$;

(8m) $(x)(y)(z)(y > 1 \rightarrow p_m(x, p_m(y, z)) \geqq p_m(p_m(x, y), z))$ [4].

The axioms of T_n are not all independent and the system can be reduced to a simpler one. So e.g. "S" can be defined explicitly and validly from "1" and "+" with respect to T_n and thus be eliminated. This fact, however, is of no interest for our purpose.

We still have to prove that T_n is really a subtheory of $\mathscr{T}_n$. The finite character of the axiom system for T_n enables us to do so in an easy way. Knowing that T_2 (being $\mathsf{T}'(\cdot)$) is a subtheory of $\mathscr{T}_2$ (being $\mathsf{T}(\cdot)$), cf. III, 1, we only have to show that the sentences (1m) ... (8m) (for m larger than 2 and not larger than n) are valid sentences of $\mathscr{T}_n$. We can do this metamathematically, deducing

---

[3] Cf. III, footnote 10.

[4] As an abbreviation we write $x \geqq y$ for $x = y \vee x > y$.

in the arithmetic of natural numbers the arithmetical theorems which are the interpretations of the sentences under consideration. In this metamathematical deduction we may use the whole apparatus of the familiar arithmetic of natural numbers (e.g. complete induction). Insofar as these deductions are not generally known, they can be carried out in an easy way.

The finite character of the axiom system for T_n enables us also to check in an easy way that each T_m has the closure property for "p_1" ... "p_m". The axioms all have the shapes:

$$(x) \ldots (y) V,$$
$$(x) \ldots (y)(Ez) \ldots (Eu) V,$$

where V is quantifier-free. The same argument as in III, 1, p. 26 can be applied. In a similar way as for $T_a(+)$ it follows that $T_m(a)$ has the closure property for "p_1" ... "p_m".

Theory T_1 being $T(+)$ of III is complete. T_2 being $T'(\cdot)$ of III is not complete. So T_m with m larger than 1 is not complete.

We wish to prove that in the sequence "p_1", "p_2", ... "p_{n-1}", "p_n" each constant is undefinable explicitly from the preceding ones and "1", "S", "$>$" with respect to theory T_n, and that "p_1" is undefinable explicitly from "1", "S" and "$>$" only with respect to T_n. Let m be one of 1, 2, ... n, then it is sufficient to prove that "p_m" is undefinable explicitly from "1", "S", "$>$", "p_1" ... "p_{m-1}" with respect to theory T_m. Then, according to Lemma 6, the statement holds also with respect to theory T_n. For "p_1" and "p_2" this proof has been established in III, 2 and III, 3. For "p_m" with m larger [5]) than 2 and not larger than n we apply Lemma 1 to T_{m-1} and T_m with a model of the former into which in no way there can be introduced a notion $\mathbf{p}_m$ as an interpretation of the "p_m" in T_m.

For each m (m larger than 1) the following sentence is provable in T_m:

(9m) $\quad (x)(y)((x > 1 \ \& \ y > S1) \rightarrow p_m(x, y) > p_{m-1}(x, y)).$

Proof: For m equals 2, see III, 3, (9·).

---

[5]) The argument is such that the case "m equal to 2" is included again.

For m larger than 2 the argument is similar:

$$(x)(y)(y>1 \rightarrow p_m(Sx, y)>Sy), \text{ (5m) and (R);}$$
$$(x)(y)(y>1 \rightarrow p_{m-1}(Sx, p_m(Sx, y))>p_{m-1}(Sx, Sy)), \text{ (7m);}$$
$$(x)(y)(y>1 \rightarrow p_m(Sx, Sy)>p_{m-1}(Sx, Sy)), \text{ (4m);}$$
$$(x)(y)((x>1 \ \& \ y>S1) \rightarrow p_m(x, y)>p_{m-1}(x, y)), \text{ (R).}$$

For each m (m larger than 1) the following sentences are provable in $\mathsf{T}_m(a)$:

(1am) $p_m(a, a)>p_{m-1}(a, a)$;

(2am) $p_m(a, a)>p_{m-1}(a, p_{m-1}(a, a))$;

(3am) $p_m(a, a)>p_{m-1}(a, p_{m-1}(a, p_{m-1}(a, a)))$;

....

Proof: For m equals 2, see III, 3, (1a·), (2a·), (3a·), ...
For m larger than 2 the argument is similar. We borrow the symbols "$a-1$", "$a-2$", "$a-3$", ..., as in III, 3. We then argue:

$$a>1 \text{ and } a>S1, \text{ (1a) and (2a);}$$

so $p_m(a, a)>p_{m-1}(a, a)$, (9m).
Further $p_m(a, a)=p_{m-1}(a, p_{m-1}(a, p_m(a, a-2)))$, (4m) twice; and $p_m(a, a-2)>a$, (5m), (6S) and (3);
so $p_m(a, a)>p_{m-1}(a, p_{m-1}(a, a))$, (7m).
Similarly we can prove (3am), and so on.

Since $\mathsf{T}_m(a)$ is consistent, it has a model. This model contains interpretations for "1", "S", "$>$", "p_1" ... "p_m" and "a", which we designate **1**, **S**, $\succ$, $\mathbf{p}_1 \ldots \mathbf{p}_m$ and **a**. We consider only the closure of $\{\mathbf{1}, \mathbf{a}\}$ under the operations $\mathbf{S}$, $\mathbf{p}_1 \ldots \mathbf{p}_m$. The universe thus deflated we designate by $\mathbf{A}^m$. We remark that $\langle \mathbf{A}^m; \mathbf{1}, \mathbf{S}, \succ, \mathbf{p}_1 \ldots \mathbf{p}_m, \mathbf{a}\rangle$ with the universe thus deflated is still a model of $\mathsf{T}_m(a)$, for T_m has the closure property for "p_1" ... "p_m" and so has $\mathsf{T}_m(a)$. Further we remark that $\langle \mathbf{A}^m; \mathbf{1}, \mathbf{S}, \succ, \mathbf{p}_1 \ldots \mathbf{p}_m, \mathbf{a}\rangle$ is also a model of T_m, a subtheory of $\mathsf{T}_m(a)$.

The ordered set $\mathbf{A}^m$ does not contain an element $\mathbf{a}^*$ with the properties:

$$\mathbf{a}^* > \mathbf{p}_m(\mathbf{a}, \mathbf{a});$$
$$\mathbf{a}^* > \mathbf{p}_m(\mathbf{a}, \mathbf{p}_m(\mathbf{a}, \mathbf{a}));$$
$$\mathbf{a}^* > \mathbf{p}_m(\mathbf{a}, \mathbf{p}_m(\mathbf{a}, \mathbf{p}_m(\mathbf{a}, \mathbf{a})));$$

For the sake of the argument we introduce $\mathbf{p}_m{}^k(\mathbf{a})$-sequences. A $\mathbf{p}_m{}^k(\mathbf{a})$-sequence is structurally defined as a term containing no other [6]) symbols than $k-1$ times the symbol $\mathbf{p}_m$ and k times the symbol **a**; the first symbol of the sequence is $\mathbf{p}_m$ and the last two are **a**, further the symbol $\mathbf{p}_m$ and the symbol **a** alternate. Thus we obtain:

$$\mathbf{p}_m{}^1(\mathbf{a}) = \mathbf{a};$$
$$\mathbf{p}_m{}^2(\mathbf{a}) = \mathbf{p}_m(\mathbf{a}, \mathbf{a});$$
$$\mathbf{p}_m{}^3(\mathbf{a}) = \mathbf{p}_m(\mathbf{a}, \mathbf{p}_m(\mathbf{a}, \mathbf{a}));$$

We have to prove now that $\mathbf{A}^m$ does not contain an element $\mathbf{a}^*$ with the properties:

$$\mathbf{a}^* > \mathbf{p}_m{}^2(\mathbf{a});$$
$$\mathbf{a}^* > \mathbf{p}_m{}^3(\mathbf{a});$$
$$\mathbf{a}^* > \mathbf{p}_m{}^4(\mathbf{a});$$
$$\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot$$

Proof: For m equal to 1, see III, 3, p. 38; for m larger than 1: elements of $\mathbf{A}^m$ which can be written with **1**, **a**, $\mathbf{p}_1 \ldots \mathbf{p}_m$ only (**S** is included in $\mathbf{p}_1$) are called polynomials in **1**, **a**, $\mathbf{p}_1 \ldots \mathbf{p}_m$. Not all elements of $\mathbf{A}^m$ are such polynomials, e.g.

$$\mathbf{a} - 1,\ \mathbf{a} - 2,\ \mathbf{a} - 3,\ \ldots$$
$$\mathbf{p}_1(\mathbf{a}, \mathbf{a}) - 1,\ \mathbf{p}_1(\mathbf{a}, \mathbf{a}) - 2,\ \mathbf{p}_1(\mathbf{a}, \mathbf{a}) - 3,\ \ldots$$

and such more.

---

[6]) Parentheses and commas are used as technical symbols only. They could as well be left out.

For every element that is not a polynomial in $\mathbf{1}, \mathbf{a}, \mathbf{p}_1 \ldots \mathbf{p}_m$ obviously there is a polynomial which exceeds it, simply obtainable by changing $-$ in $+$ (i.e. $\mathbf{p}_1$) wherever it occurs. We then argue that every polynomial is either a $\mathbf{p}_m{}^k(\mathbf{a})$-sequence or it is exceeded by such a sequence. When the polynomial is not a $\mathbf{p}_m{}^k(\mathbf{a})$-sequence, then we obtain an element that exceeds it by replacing first all symbols $\mathbf{1}$ by $\mathbf{a}$ ((6m) and (7m) or (6·) and (7·)), and then all $\mathbf{p}_1 \ldots \mathbf{p}_{m-1}$ by $\mathbf{p}_m$, ((9m)). Further we replace successively all sub-terms of the shape $\mathbf{p}_m(\mathbf{p}_m{}^r(\mathbf{a}), \mathbf{p}_m{}^s(\mathbf{a}))$ by $\mathbf{p}_m{}^{r+s}(\mathbf{a})$, a process by which the term can only grow, for

$$\mathbf{p}_m(\mathbf{p}_m{}^r(\mathbf{a}), \mathbf{p}_m{}^s(\mathbf{a})) = \mathbf{p}_m(\mathbf{p}_m(\mathbf{a}, \mathbf{p}_m{}^{r-1}(\mathbf{a})), \mathbf{p}_m{}^s(\mathbf{a})) \leqq$$
$$\leqq \mathbf{p}_m(\mathbf{a}, \mathbf{p}_m(\mathbf{p}_m{}^{r-1}(\mathbf{a}), \mathbf{p}_m{}^s(\mathbf{a}))), \quad ((8m) \text{ or } (8\cdot));$$

and so on.

So there is no element $\mathbf{a}^*$ that exceeds all $\mathbf{p}_m{}^k(\mathbf{a})$-sequences.

Theorem 7: The non-logical constant "p_m" is undefinable explicitly from the non-logical constants "1", "S", "$>$", "p_1" ... ... "p_{m-1}" with respect to theory T_m.

Proof: For m equal to 1, see III, 2. For m larger than 1: we apply Lemma 1 to T_{m-1} and T_m with $\langle \mathbf{A}^{m-1}; \mathbf{1}, \mathbf{S}, \succ, \mathbf{p}_1 \ldots \mathbf{p}_{m-1}, \mathbf{a} \rangle$ as a model of the former. If into this model there could be introduced a notion $\mathbf{p}_m$, the interpretation of the "p_m" of T_m, then $\langle \mathbf{A}^{m-1}; \mathbf{1}, \mathbf{S}, \succ, \mathbf{p}_1 \ldots \mathbf{p}_{m-1}, \mathbf{p}_m, \mathbf{a} \rangle$ would be a model of $\mathsf{T}_m(a)$. But according to (1m) or (1·) $\mathbf{A}^{m-1}$ would then contain an element $\mathbf{p}_m(\mathbf{a}, \mathbf{a})$. This element, however, would have the properties:

$$\mathbf{p}_m(\mathbf{a}, \mathbf{a}) \succ \mathbf{p}_{m-1}(\mathbf{a}, \mathbf{a});$$
$$\mathbf{p}_m(\mathbf{a}, \mathbf{a}) \succ \mathbf{p}_{m-1}(\mathbf{a}, \mathbf{p}_{m-1}(\mathbf{a}, \mathbf{a}));$$
$$\mathbf{p}_m(\mathbf{a}, \mathbf{a}) \succ \mathbf{p}_{m-1}(\mathbf{a}, \mathbf{p}_{m-1}(\mathbf{a}, \mathbf{p}_{m-1}(\mathbf{a}, \mathbf{a})));$$
$$\cdot \quad \cdot \quad \cdot \qquad \cdot \quad \cdot \quad \cdot$$

according to (1am), (2am), (3am), ... But as we have seen $\mathbf{A}^{m-1}$ does not contain an element which exceeds all $\mathbf{p}_{m-1}{}^k(\mathbf{a})$-sequences.

Theorem 8: Each non-logical constant in the sequence "p_1" ... "p_n" is undefinable explicitly from the preceding ones together with "1", "S" and "$>$" with respect to theory T_n (for "p_1" this means from "1", "S" and "$>$" only).

Proof: The theorem is an immediate consequence of Theorem 7 and Lemma 6.

Corollary: Each non-logical constant in the sequence ("1"), "p_1" ... p_n is undefinable explicitly from the preceding ones with respect to theory T_n, and also with respect to the subtheory of T_n obtained from T_n by dropping all sentences containing "S" and "$>$".

The corollary is trivial. If the non-logical constant d is undefinable explicitly from the non-logical constants $c_1 \dots c_k$ with respect to a theory T, then a fortiori d is undefinable explicitly from some of $c_1 \dots c_k$ with respect to T. Further, since T_n is an extension of $\mathsf{T}(+)$, "S" and "$>$" can be defined explicitly and validly from "1" and "$+$" with respect to T_n and so be eliminated. It is even possible to go on in this direction and to eliminate more of these non-logical constants [7]), but this is beyond the scope of this study.

Corollary: Each non-logical constant in the sequence "p_1" ... "p_n" is undefinable explicitly from the preceding ones together with "1", "S" and "$>$" ("p_1" from "1", "S" and "$>$" only) with respect to every subtheory of T_n containing these non-logical constants.

This corollary is an immediate application of the property that an undefinable non-logical constant is hereditarily undefinable (cf. II, 4, p. 22). One could make more corollaries of this kind combining the two mentioned, but all this is trivial.

Theorem 9: Each non-logical constant in the sequence "p_1" ... "p_n" is undefinable explicitly from the preceding ones together with "1", "S" and "$>$" ("p_1" from "1", "S" and "$>$" only) with respect to every consistent extension $\mathsf{T}_n{}^*$ of T_n fulfilling the conditions

i) $\mathsf{T}_n{}^*$ is a subtheory of $\mathscr{T}_n$;

ij) $\mathsf{T}_{n-1}{}^*$, the subtheory of $\mathsf{T}_n{}^*$ containing all sentences of $\mathsf{T}_n{}^*$ in which "p_n" does not occur, has the closure property for "p_1" ... "p_{n-1}".

---

[7]) Cf. [17].

Proof: Let $\mathsf{T}_n{}^*$ be an extension of T_n and a subtheory of $\mathscr{T}_n$, and let $\mathsf{T}_{n-1}{}^*$ have the closure property for "p_1" ... "p_{n-1}". Let $\mathsf{T}_m{}^*$ be the subtheory of $\mathsf{T}_n{}^*$ containing all sentences of $\mathsf{T}_n{}^*$ in which "p_{m+1}" ... "p_{n-1}", "p_n" do not occur. Obviously $\mathsf{T}_m{}^*$ is an extension of T_m and obviously $\mathsf{T}_{m-1}{}^*$ has the closure property for "p_1" ... "p_{m-1}". The same argument as for T_m in Theorem 7 holds for $\mathsf{T}_m{}^*$. Further the same argument as for T_n in Theorem 8 holds for $\mathsf{T}_n{}^*$.

For m larger than 2 it cannot be concluded that "p_m" is essentially undefinable explicitly from "1", "S", "$>$", "p_1" ... "p_{m-1}" with respect to theory T_m, since neither T_m nor T_{m-1} is complete. Hence it cannot be concluded that each non-logical constant of the sequence "p_3" ... "p_n" is essentially undefinable explicitly from the preceding ones together with "1", "S", "$>$", "p_1" and "p_2" with respect to theory T_n. Thus it does not follow that each constant of this sequence is undefinable explicitly from the preceding ones together with "1", "S", "$>$", "p_1" and "p_2" with respect to e.g. theory $\mathscr{T}_n$. This is in accordance with the fact that the non-logical constants "p_3", "p_4", ... "p_n" are known to be arithmetical, i.e. definable explicitly and validly from "1", ("S", "$>$",) "p_1" and "p_2" with respect to $\mathscr{T}_n$ [8]). In connection with Theorem 9 it follows from Gödel's result that, for n larger than 1, $\mathscr{T}_n$ has not the closure property for "p_1" ... "p_n" [9]).

IV, 2. *Equational undefinability*

In the foregoing section we met the operation constants "p_3" ... "p_n", which according to Gödel's result were definable explicitly and validly from "1", "S", "$>$", "p_1" and "p_2" with respect to theory $\mathscr{T}_n$. At this point the question arises whether or not a constant of the sequence "p_3" ... "p_n" is definable equationally and validly [10]) from the preceding ones together with "1", "S", "$>$", "p_1" and "p_2" with respect to theory $\mathscr{T}_n$. It is

---

[8]) Cf. [6], Satz VII, pp. 191 ff.

[9]) Cf. [10], p. 168, where Kemeny states that $\mathscr{T}_2$ has not the closure property for "+" and "·".

[10]) We say "validly" because $\mathscr{T}_n$ is complete and all sentences containing no more non-logical constants than "1", "S", "$>$", "p_1" ... "p_n" which are compatible with $\mathscr{T}_n$ are valid in $\mathscr{T}_n$.

well-known that the answer to this question is negative. In $\mathscr{T}_n$, one says, "p_m" grows faster than any polynomial in "p_1" ... "p_{m-1}". E.g. in $\mathscr{T}_n$ "p_3" grows faster than any polynomial in "p_1", "p_2", i.e. some sentence of the shape [11]):

$$(x)(y)(Ez)((x>z \ \& \ y>z) \rightarrow p_3(x, y) > Q(x, y)),$$

is valid in $\mathscr{T}_n$, where $Q(x, y)$ is an arbitrary term, containing "x" and "y" as the only free variables and no other operation constants than "p_1" (including "S") and "p_2", (since Q is a term, it does not contain logical constants or relation constants). The argument is: if "p_3" were definable equationally and validly from "1", "p_1" and "p_2" with respect to $\mathscr{T}_n$, then for a certain $Q(x, y)$ a sentence of the shape:

$$(x)(y)(p_3(x, y) = Q(x, y)),$$

would be valid in $\mathscr{T}_n$. But the latter sentence is not compatible with the former.

This can be generalized in a slight way by saying: in $\mathscr{T}_n$ a sentence of the shape S:

$$(Ex)(Ey)\overline{(p_3(x, y) = Q(x, y))},$$

is valid for any $Q(x, y)$; therefore "p_3" is not definable equationally and validly from "1", "p_1" and "p_2" with respect to $\mathscr{T}_n$. It is obvious that "p_3" is not definable equationally and compatibly from "1", "p_1" and "p_2" with respect to any subtheory of $\mathscr{T}_n$ which contains the non-logical constants "1", "p_1", "p_2" and "p_3" and in which some sentence of the shape S is valid. In other words, if T is a subtheory of $\mathscr{T}_n$ containing the non-logical constants mentioned, and if a sentence of the shape S is valid in T, then "p_3" is essentially undefinable equationally from "1", "p_1" and "p_2" with respect to T. Thus, in order to prove that "p_3" is undefinable equationally from "1", "p_1" and "p_2" with respect to $\mathscr{T}_n$, all we have to do is to prove that e.g. in T_3 of the foregoing section a sentence of the shape S is valid for any $Q(x, y)$.

Proceeding we shall argue: the condition T_3 *has not the closure property for* "p_1" *and* "p_2" is equivalent to the condition *a sentence*

---

11) "$p_3(x, y)$" can be written as "$x \exp y$" or "x^y"; "$p_1(x, y)$" and "$p_2(x, y)$" as "$x + y$" and "$x \cdot y$" respectively.

of the shape S is valid in T_3 *for any* $Q(x, y)$. Thus it will be sufficient to prove that T_3 has not the closure property for "p_1" and "p_2". This is exactly what we did implicitly in the argument of Theorem 7 in the foregoing section, where we proved that

$$\langle \mathbf{A}^{m-1}; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{p}_1 \dots \mathbf{p}_{m-1}, \mathbf{p}_m, \mathbf{a} \rangle$$

is not a model of $\mathsf{T}_m(a)$, which should be the case if $\mathsf{T}_m(a)$ had the closure property for "p_1" ... "p_{m-1}".

In this section we deal with equational undefinability in a general way which is closely connected with the method developed in chapter II to prove the (explicit) [12] undefinability of non-logical constants. In fact, the examples of IV, 1 reveal a unity of treatment: there we had operation constants which are definable explicitly and validly from other non-logical constants with respect to theory $\mathscr{T}_n$. At the same time they were undefinable explicitly from these other non-logical constants with respect to theory T_n, a certain subtheory of $\mathscr{T}_n$. This implied that they were undefinable equationally from the other non-logical constants in question with respect to theory T_n (explicit undefinability implies equational undefinability). We shall see now that the arguments given in IV, 1 entail the possibility to state that the non-logical constants in question are essentially undefinable equationally from the other non-logical constants in question with respect to theory T_n, and so undefinable equationally with respect to $\mathscr{T}_n$. Thus in one stroke we show the explicit undefinability with respect to certain subtheories of $\mathscr{T}_n$ and the equational undefinability with respect to $\mathscr{T}_n$ itself. At the same time we deal with the crucial case (mentioned in II, 1, p. 14) in terms of models also.

First we have to consider the closure property in a more general way. Let T be a consistent theory containing the non-logical constants $i_1 \dots i_l$, $o_1 \dots o_m$ and $r_1 \dots r_n$. Let the model $\langle \mathbf{U}; \mathbf{i}_1 \dots \mathbf{i}_l, \mathbf{o}_1 \dots \mathbf{o}_m, \mathbf{r}_1 \dots \mathbf{r}_n \rangle$ be an arbitrary model of T, $\mathbf{U}$ being the universe and $\mathbf{i}_1 \dots \mathbf{i}_l, \mathbf{o}_1 \dots \mathbf{o}_m, \mathbf{r}_1 \dots \mathbf{r}_n$ the interpretations of $i_1 \dots i_l$, $o_1 \dots o_m$, $r_1 \dots r_n$ respectively. Let $\mathbf{U}'$ be the closure of $\{\mathbf{i}_1' \dots \mathbf{i}_l'\}$ under the operations $\mathbf{o}_i \dots \mathbf{o}_j$, where $\mathbf{o}_i \dots \mathbf{o}_j$ are some

[12] The parentheses indicate that explicit undefinability implies equational undefinability.

of $\mathbf{o}_1 \dots \mathbf{o}_m$ [13]) and where $\mathbf{i}_1' \dots \mathbf{i}_l'$ are arbitrary elements of $\mathbf{U}$, such however, that $\mathbf{i}_1' \dots \mathbf{i}_l'$ are apt to be the interpretations of $i_1 \dots i_l$ respectively. The last remark means that, if e.g.

$$r_1(i_1 \dots i_k), \text{ (r_1 assumed to be k-ary)}$$

is a valid sentence of T, then

$$\mathbf{r}_1(\mathbf{i}_i' \dots \mathbf{i}_k'),$$

is true in $\langle \mathbf{U}; \mathbf{i}_1 \dots \mathbf{i}_l, \mathbf{o}_1 \dots \mathbf{o}_m, \mathbf{r}_1 \dots \mathbf{r}_n \rangle$, and likewise for all mutual relations concerning $i_1 \dots i_l$ as given in T. $\mathbf{U}'$ being the closure of $\{\mathbf{i}_1' \dots \mathbf{i}_l'\}$ under the operations $\mathbf{o}_i \dots \mathbf{o}_j$ means that $\mathbf{U}'$ contains in addition to $\mathbf{i}_1' \dots \mathbf{i}_l'$ no more elements of $\mathbf{U}$ than those obtainable from $\mathbf{i}_1' \dots \mathbf{i}_l'$ by (repeated) application of the operations $\mathbf{o}_i \dots \mathbf{o}_j$. T is said to have the closure property for $o_i \dots o_j$ if and only if for any model $\langle \mathbf{U}; \mathbf{i}_1 \dots \mathbf{i}_l, \mathbf{o}_1 \dots \mathbf{o}_m, \mathbf{r}_1 \dots \mathbf{r}_n \rangle$ of T and for any suitable $\mathbf{i}_1' \dots \mathbf{i}_l'$ of $\mathbf{U}$, $\langle \mathbf{U}'; \mathbf{i}_1' \dots \mathbf{i}_l', \mathbf{o}_1 \dots \mathbf{o}_m, \mathbf{r}_1 \dots \mathbf{r}_n \rangle$ is also a model of T. E.g. in the foregoing sections we saw that T(+) or T_1 had the closure property for "S" and "+" ("p_1"), T′(·) or T_2 had the closure property for "S", "+" and "·" ("p_1" and "p_2"), T_{m-1} had the closure property for "p_1" ... "p_{m-1}". In this section we shall speak explicitly about the fact that T_m had not the closure property for "p_1" ... "p_{m-1}".

Let T be a consistent theory containing the non-logical constants $i_1 \dots i_l, o_1 \dots o_m, o, r_1 \dots r_n$. Let T_0 be the subtheory of T containing all sentences of T in which o does not occur (a sentence being valid in T_0 if and only if it is valid in T). Let T_0 have the closure property for $o_1 \dots o_m$. Let M be an arbitrary model of T, say M is $\langle \mathbf{U}; \mathbf{i}_1 \dots \mathbf{i}_l, \mathbf{o}_1 \dots \mathbf{o}_m, \mathbf{o}, \mathbf{r}_1 \dots \mathbf{r}_n \rangle$. Since T_0 is a subtheory of T, the model $\langle \mathbf{U}; \mathbf{i}_1 \dots \mathbf{i}_l, \mathbf{o}_1 \dots \mathbf{o}_m, \mathbf{r}_1 \dots \mathbf{r}_n \rangle$ is a model of T_0. Moreover, the model $\langle \mathbf{U}'; \mathbf{i}_1' \dots \mathbf{i}_l', \mathbf{o}_1 \dots \mathbf{o}_m, \mathbf{r}_1 \dots \mathbf{r}_n \rangle$ where $\mathbf{U}'$ is the closure of $\{\mathbf{i}_1' \dots \mathbf{i}_l'\}$ under the operations $\mathbf{o}_1 \dots \mathbf{o}_m$, is a model of T_0, since this theory has the closure property for $o_1 \dots o_m$. This is the case for any model M of T. On the other hand, if T has not the closure

[13]) In the particular cases of Chapter III and IV, 1 we specified that in addition to S and + the operations −1 and − were taken into consideration (cf. Chapter III, footnote 3). This particular convention does not apply to the general treatment here. If, however, in this section we refer to theories of the preceding sections as illustrations of the general treatment, the convention of footnote 3, Chapter III holds for these theories.

property for $o_1 \dots o_m$, i.e. if there is a model M of T which does not remain a model of T if the universe is deflated to the closure of a certain $\{\mathbf{i}_1' \dots \mathbf{i}_l'\}$ under the operations $\mathbf{o}_1 \dots \mathbf{o}_m$, then there is a $\langle \mathbf{U}'; \mathbf{i}_1' \dots \mathbf{i}_l', \mathbf{o}_1 \dots \mathbf{o}_m, \mathbf{o}, \mathbf{r}_1 \dots \mathbf{r}_n \rangle$, which is not a model of T. In other words, in this case there is not only a model of T_0, into which in no way there can be introduced an interpretation $\mathbf{o}$ of the o of T, but the very reason of this impossibility is given also, namely the deflation of the universe of the model. According to Lemma 1 it follows that o is undefinable explicitly (and equationally) from $i_1 \dots i_l, o_1 \dots o_m, r_1 \dots r_n$ with respect to T. It is also obvious that this is due to the valid existential sentences of T. Intuitively speaking one could say that apparently $\mathbf{U}'$ contains all elements required by the valid existential sentences of T_0, but not all elements required by the valid existential sentences of T. This applies already if T is *the minimal extension* of T_0 with the operation constant o, i.e. if to the valid sentences of T_0 are added as valid sentences only the two sentences:

$$(x) \dots (y)(Ez)(o(x, \dots, y)=z),$$
$$(x) \dots (y)(z)(u)((o(x, \dots, y)=z \;\&\; o(x, \dots, y)=u) \rightarrow z=u),$$

required for o to be an operation constant. Hence, if this T does not have the closure property for $o_1 \dots o_m$, then the sentence:

$$(x) \dots (y)(Ez)(o(x, \dots, y)=z),$$

does not hold for the deflated model. If, however, there was a sentence of the shape D:

$$(x) \dots (y)(o(x, \dots, y)=Q(x, \dots, y)),$$

(i.e. an equational definition), which was compatible with T (where $Q(x, \dots, y)$ is a polynomial in $o_1 \dots o_m$ with "x" ... "y" as the only free variables), then the deflated model would be a model of $T \cup D$, and this is contradictory. It follows that a sentence of the shape D is not compatible with T, or that o is essentially undefinable equationally from $i_1 \dots i_l, o_1 \dots o_m$ with respect to T, or that any sentence of the shape S:

$$(Ex) \dots (Ey)\overline{(o(x, \dots, y)=Q(x, \dots, y))},$$

is valid in T. The same applies to any T which contains no other valid existential sentences containing o than those derivable from the union of T_0 and the sentence:

$$(x) \ldots (y)(Ez)(o(x, \ldots, y) = z).$$

Whenever this is the case, we call T *a minimal extension* of T_0 with o.

We then summarize in the following way:

Lemma 7: Let T be a consistent theory containing the non-logical constants $i_1 \ldots i_l$, $o_1 \ldots o_m$, o, $r_1 \ldots r_n$. Let T_0 be the subtheory of T containing all sentences of T in which o does not occur. Let T be a minimal extension of T_0 with o. If T_0 has the closure property for $o_1 \ldots o_m$, whereas T has not the closure property for $o_1 \ldots o_m$, then o is undefinable explicitly from $i_1 \ldots i_l$, $o_1 \ldots o_m$, $r_1 \ldots r_n$ and essentially undefinable equationally from $i_1 \ldots i_l$, $o_1 \ldots o_m$ with respect to T.

As a first example we can state:

Theorem 10: The non-logical constant "p_m" is essentially undefinable equationally from the non-logical constant "1", "S", "p_1" ... "p_{m-1}" with respect to theory T_m of the foregoing section.

Proof: In the proof of Theorem 7 we established that T_{m-1} has the closure property for "p_1" ... "p_{m-1}", whereas T_m has not the closure property for "p_1" ... "p_{m-1}". We saw (on p. 47) that $\langle \mathbf{A}^{m-1}; \mathbf{1}, \mathbf{S}, \mathbf{>}, \mathbf{p}_1 \ldots \mathbf{p}_{m-1}, \mathbf{p}_m, \mathbf{a} \rangle$ is not a model of $\mathsf{T}_m(a)$. Hence, $\mathsf{T}_m(a)$ has not the closure property for "p_1" ... "p_{m-1}"; and hence, T_m has not the closure property for "p_1" ... "p_{m-1}" (if T_m has the closure property for "p_1" ... "p_{m-1}", then $\mathsf{T}_m(a)$ has the closure property for "p_1" ... "p_{m-1}"). Moreover T_m is a minimal extension of T_{m-1} with "p_m". We apply Lemma 7.

In a similar way as in II, 4 we can develop an indirect method in proofs of equational undefinability. The arguments are obvious (cf. II, 4), and we shall confine ourselves to giving the main results in some lemmas.

Lemma 8: If a theory T_1 is a subtheory of a consistent theory T_2, both containing the non-logical constants $i_1 \ldots i_l$, $o_1 \ldots o_m$, o, and if o is essentially undefinable equationally from $i_1 \ldots i_l$, $o_1 \ldots o_m$

with respect to T_1, then o is essentially undefinable equationally from $i_1 \dots i_l$, $o_1 \dots o_m$ with respect to T_2.

Lemma 9: Let T_1 and T_2 be two compatible theories, both containing the non-logical constants $i_1 \dots i_l, o_1 \dots o_m, o$, and let o be essentially undefinable equationally from $i_1 \dots i_l$, $o_1 \dots o_m$ with respect to T_1; then o is undefinable equationally from $i_1 \dots i_l, o_1 \dots o_m$ with respect to T_2.

Lemma 10: Let a theory T_1 containing the non-logical constants $i_1 \dots i_l$, $o_1 \dots o_m$, o be a subtheory of a consistent theory T_2 containing the non-logical constants $i_1 \dots i_l$, $i_{l+1} \dots i_{l+j}$, $o_1 \dots o_m$, o, $o_{m+1} \dots o_{m+k}$, whereas T_1 contains all sentences of T_2 in which $i_{l+1} \dots i_{l+j}$, $o_{m+1} \dots o_{m+k}$ do not occur and o is undefinable equationally from $i_1 \dots i_l$, $o_1 \dots o_m$ with respect to T_1; then o is undefinable equationally from $i_1 \dots i_l$, $o_1 \dots o_m$ with respect to T_2.

It is not necessary to rewrite Lemma 10 with "essentially undefinable equationally" instead of "undefinable equationally". Lemma 8 provides for the case that T_2 contains more non-logical constants than T_1. The non-logical constants of T_2, which are not contained in T_1, do not play a part in the definition problem under consideration in the case of Lemma 8.

As examples of the indirect method we can state:

Theorem 11: The non-logical constant "p_m" is essentially undefinable equationally from the non-logical constants "1", "S", "p_1" ... "p_{m-1}" not only with respect to T_m, but also with respect to every consistent extension of T_m, in particular with respect to $\mathscr{T}_m$.

Proof: The theorem is an immediate consequence of Theorem 10 and Lemma 8.

Theorem 12: The non-logical constant "p_m" is undefinable equationally from the non-logical constants "1", "S", "p_1" ... "p_{m-1}" with respect to any subtheory of $\mathscr{T}_m$ containing "1", "S", "p_1" ... "p_{m-1}". The same holds with respect to any consistent theory which is compatible with T_m and which contains the non-logical constants mentioned.

Proof: The first part of the theorem follows from the fact that equational undefinability in the same way as explicit undefinability is hereditary (cf. II, 4, p. 22). The second part results from Lemma 9.

Theorem 13: Each non-logical constant in the sequence "p_1" ... "p_n" is essentially undefinable equationally from the preceding ones together with "1" and "S" with respect to theory T_n. (For "p_1" this means: from "1" and "S" only.)

Proof: The theorem is an immediate consequence of Theorem 10 and Lemma 8. If T_m is extended to T_n it is true that T_n contains besides "p_1" ... "p_m" still other non-logical constants ("p_{m+1}" ... ... "p_n"), but they do not matter.

Theorem 14: Each non-logical constant in the sequence "p_1" ... "p_n" is essentially undefinable equationally from the preceding ones together with "1" and "S" with respect to any consistent extension of T_n, in particular with respect to $\mathscr{T}_n$ (for "p_1": from "1" and "S" only).

Proof: The theorem is an immediate consequence of Theorem 13 and Lemma 8.

Theorem 15: Each non-logical constant in the sequence "p_1" ... "p_n" is undefinable equationally from the preceding ones together with "1" and "S" with respect to any subtheory of $\mathscr{T}_n$ containing "1", "S", "p_1" ... "p_n". The same holds with respect to any consistent theory which is compatible with T_n and which contains the non-logical constants mentioned (for "p_1": from "1" and "S" only).

Proof: The first part of the theorem follows from the fact that equational undefinability is hereditary. The second part results from Lemma 9.

What is said about the (essential) equational undefinability of operation constants applies also to individual constants. We consider the individual constant in question as a 0-ary operation constant and argue in the same way. E.g. let T be a consistent theory containing the non-logical constants $i_1 \dots i_l$, i, $o_1 \dots o_m$, $r_1 \dots r_n$. T has the closure property for e.g. i, $o_1 \dots o_m$ if and only if for any model $\langle \mathbf{U}; \mathbf{i}_1 \ldots \mathbf{i}_l, \mathbf{i}, \mathbf{o}_1 \ldots \mathbf{o}_m, \mathbf{r}_1 \ldots \mathbf{r}_n \rangle$ of T and for any suitable $\mathbf{i}_1' \ldots \mathbf{i}_l'$ of $\mathbf{U}$, $\langle \mathbf{U}'; \mathbf{i}_1' \ldots \mathbf{i}_l', \mathbf{i}, \mathbf{o}_1 \ldots \mathbf{o}_m, \mathbf{r}_1 \ldots \mathbf{r}_n \rangle$ is also a model of T, where $\mathbf{U}'$ is the closure of $\{\mathbf{i}_1' \ldots \mathbf{i}_l'\}$ under the operations $\mathbf{i}, \mathbf{o}_1 \ldots \mathbf{o}_m$, i.e. $\mathbf{i}$ belongs to the closure of $\{\mathbf{i}_1' \ldots \mathbf{i}_l'\}$ under the operations $\mathbf{o}_1 \ldots \mathbf{o}_m$. If in this context we say that T does not have the closure property for $\mathbf{o}_1 \ldots \mathbf{o}_m$, we imply that there is a model $\langle \mathbf{U}; \mathbf{i}_1 \ldots \mathbf{i}_l, \mathbf{i}, \mathbf{o}_1 \ldots \mathbf{o}_m, \mathbf{r}_1 \ldots \mathbf{r}_n \rangle$ of T, such that a suitable $\langle \mathbf{U}'; \mathbf{i}_1' \ldots \mathbf{i}_l', \mathbf{i}, \mathbf{o}_1 \ldots \mathbf{o}_m, \mathbf{r}_1 \ldots \mathbf{r}_n \rangle$ is not a model of T. If T_0 is the subtheory of T containing all sentences of T in which i does not occur, and if T_0 has the closure property for $o_1 \ldots o_m$, whereas T is a minimal extension of T_0 with i, which does not have the closure property for $o_1 \ldots o_m$, then i is undefinable explicitly from $i_1 \ldots i_l$, $o_1 \ldots o_m$, $r_1 \ldots r_n$ and essentially undefinable equationally from $i_1 \ldots i_l$, $o_1 \ldots o_m$ with respect to T. T being a minimal extension of T_0 with i means in this context that T contains no more valid existential sentences containing i than those derivable from the union of T_0 and the sentence $(Ex)(x=i)$.

E.g. $\mathsf{T}_a(S)$ of chapter III contains the non-logical constants "1", "a", "S" and "$>$". We consider "a" as a 0-ary operation constant. $\mathsf{T}(S)$ is the subtheory of $\mathsf{T}_a(S)$ containing all sentences of $\mathsf{T}_a(S)$ in which "a" does not occur. $\mathsf{T}_a(S)$ is a minimal extension of $\mathsf{T}(S)$ with "a". $\mathsf{T}(S)$ has the closure property for "S"; $\mathsf{T}_a(S)$ has not the closure property for "S", since $\mathbf{a}$ does not belong to the closure of $\{\mathbf{1}\}$ under the operation $\mathbf{S}$. It follows that "a" is undefinable explicitly from "1", "S" and "$>$"," and essentially undefinable equationally from "1" and "S" with respect to $\mathsf{T}_a(S)$. In the same way it can be proved that "a" is undefinable explicitly from "1", "S", "$>$", "p_1" ... "p_m" and essentially undefinable equationally from "1", "p_1" ... "p_m" with respect to theory $\mathsf{T}_m(a)$.

REFERENCES

[1] W. Ackermann: Zum Hilbertschen Aufbau der reellen Zahlen. *Mathematische Annalen*, **99** (1928), pp. 118 ff.

[2] E. W. Beth: On Padoa's method in the theory of definition. *Indagationes Mathematicae*, **15** (1953), pp. 330 ff.

[3] E. W. Beth: *Les fondements logiques des mathématiques*. Deuxième édition revue et augmentée, Paris–Louvain, 1955.

[4] E. W. Beth: *L'existence en mathématiques*. Paris–Louvain, 1956.

[5] J. Dopp: Les variétés syntaxiques de la définition dans les langages rigoureux. *Les Études Philosophiques*, **11** (1956), pp. 209 ff.

[6] K. Gödel: Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, **38** (1931), pp. 173 ff.

[7] L. A. Henkin: Metamathematical theorems equivalent to the prime ideal theorems for Boolean algebras, preliminary report. *Bulletin of the American Mathematical Society*, **60** (1954), pp. 387 f. (abstract 553).

[8] A. Heyting: On truth in mathematics, in *Koninklijke Nederlandse Akademie van Wetenschappen, Speeches and Reports on may* 7 *and* 9, 1958. Amsterdam, 1958, pp. 146 ff.

[9] D. Hilbert und P. Bernays: *Grundlagen der Mathematik, erster Band*. Berlin, 1934.

[10] J. G. Kemeny: Undecidable problems of elementary number theory. *Mathematische Annalen*, **135** (1958), pp. 160 ff.

[11] S. C. Kleene: *Introduction to Metamathematics*. Amsterdam, 1953.

[12] J. C. C. McKinsey: On the independence of undefined ideas. *Bulletin of the American Mathematical Society*, **41** (1935), pp. 291 ff.

[13] A. Padoa: Essai d'une théorie algébrique des nombres entiers, précédé d'une introduction logique à une théorie déductive quelconque. *Bibliothèque du Congrès International de Philosophie*, **3** (1901). Paris, 1901, pp. 309 ff.

[14] A. Padoa: Un nouveau système irréductible de postulats pour l'algèbre. *Compte Rendu du Deuxième Congrès International des Mathématiciens tenu à Paris du* 6 *à* 12 *août* 1900. Paris, 1902, pp. 249 ff.

[15] A. Padoa: Le problème no. 2 de M. David Hilbert. *L'enseignement mathématique*, **5** (1903), pp. 85 ff.

[16] M. Presburger: Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt. *Comptes-rendus du I Congrès des Mathématiciens des Pays Slaves*. Warszawa, 1929, pp. 92 ff. and p. 395.

[17] Julia Robinson: Definability and decision problems in arithmetic. *The Journal of Symbolic Logic*, **14** (1949), pp. 98 ff.

[18] Barkley Rosser: Extensions of some theorems of Gödel and Church. *The Journal of Symbolic Logic,* **1** (1936), pp. 87 ff.

[19] Th. Skolem: Über die Nicht-characterisierbarkeit der Zahlenreihe mittels endlich oder abzählbar unendlich vieler Aussagen mit ausschlieszlich Zahlenvariablen. *Fundamenta Mathematicae,* 23 (1934), pp. 150 ff.

[20] A. Tarski et A. Lindenbaum: Sur l'indépendance des notions primitives dans les systèmes mathématiques. *Annales de la Société Polonaise de Mathématique,* **5** (1926), pp. 111 ff.

[21] A. Tarski: Some methodological investigations on the definability of concepts, in *Logic, Semantics, Metamathematics, papers from* 1923 *to* 1938 *by Alfred Tarski, translated by J. H. Woodger.* Oxford, 1956, pp. 296 ff. The paper concerned is a translation of the German version (1935) supplemented by some passages translated from the Polish original (1934).

[22] A. Tarski, A. Mostowski, R. M. Robinson: *Undecidable Theories.* Amsterdam, 1953.

[23] A. Tarski: Prime ideal theorems for set algebras and ordering principles, preliminary report. *Bulletin of the American Mathematical Society,* **60** (1954), p. 391 (abstract 563, cf. abstracts 562 and 564).

INDEX

Numbers refer to pages. Numbers following the name of an author indicate those pages on which either the author himself is mentioned or reference is made to one of his works. To avoid repetitions, words are often replaced by dashes, a dash replacing one word.

Ackermann, W., 41, 58
The — sequence, see: sequence
Addition, 21, 25
Arithmetic, 19, 21, 25, 33, 37, 41, 44
Arithmetical, 25f., 30, 32, 34, 41, 44, 49
Atomic formula, 1, 8, 29
— term, 1, 28
Auxiliary constant, 37
Axiom, 26ff., 30, 33, 36, 42ff.
non-logical —, see: non-logical
— system, 26f., 32f.
Axiomatic(ally), 19, 42
Axiomatizable, 25ff., 32, 36, 39f., 42
Axiomatization, 36
Axiomatize, 43

Bernays, P., 8, 58
Beth, E. W., 2, 4, 17ff. ,32, 36, 58
Theorem of —, see: theorem
Binary, 19, 25, 29, 41f.

Characteristic sentence, 2f., 5ff., 33
Characterization in terms of models, 4, 7f., 12, 14ff., 18
Class, 20
Family of —es, see: family
Closure, 26ff., 33f., 37, 45, 51ff., 57
The — property, 26ff., 30, 32f., 36f., 42, 44f., 48ff., 56f.
Comma, 1, 43, 46
Complete, 15ff., 25ff., 32, 35f., 38, 42, 44, 49
— induction, see: induction
Compatibility, 2
Compatible (bly) 2f., 6, 10, 12, 15ff., 22ff., 35, 39f., 49f., 53, 55f.
Definable equationally and —, see: definable
Definable explicitly and —, see: definable
Compound formula, 1
— term, 1, 29
Concatenation, 1
Consistent, *passim*
Constant
Auxiliary —, see: auxiliary
Corresponding relation —, see: corresponding
— function, see: function
Individual —, see: individual
Logical —, see: logical
Non-logical —, *passim*
Operation —, see: operation
Relation —, see: relation
Context, 8
Contextual, 7
Corresponding relation constant, 7f.
Counterexample, 11, 13

Decidable, 21, 26, 40
Decision method, see: method
Definability, 1f.
Equational —, 9, 11
Explicit —, 2, 4f., 10f., 14ff.
Definable
— equationally and compatibly, 10ff., 50
— equationally and validly, 10ff., 49f.
— explicitly, 2f., 14, 21
— explicitly and compatibly, 2ff., 9, 12, 14f., 17, 22f., 49
— explicitly and validly, 2ff., 11f., 14ff., 22f., 33, 35, 39f., 51
Definition, 2

Equational —, 9ff., 53
Explicit —, 2, 4ff., 11f., 16ff., 24ff.
— problem, see: problem
Definiendum, 7
Definitum 7f.
Deflated, 27, 33, 37, 45, 53
Deflation, 53
Direct method, see: method
Derivable, 8, 20, 32, 54, 57
Dopp, J., 7, 58

Element, 19, 25ff., 34, 38, 41, 45ff., 52f.
Eliminate, 33, 40, 43, 48
Enumerable, 34, 37
Equational
— definability, see: definability
— definition, see: definition
— undefinability, see: undefinability
Equationally
Definable — and compatibly, see: definable
Definable — and validly, see: definable
Essentionally undefinable —, see: undefinable
Undefinable —, see: undefinable
Equivalence, 7, 12, 32
Equivalent, 8, 11, 13, 50
Essential undecidability, see: undecidability
— undefinability, see: undefinability
Essentially undecidable, see: undecidable
— undefinable equationally, see: undefinable
— undefinable explicitly, see: undefinable
Exceed, 34, 38, 47
Existential, 53f., 57
Explicit definability, see: definability
— definition, see: definition
— undefinability, see: undefinability
Explicitly
Definable —, see: definable
Definable — and compatibly, see: definable
Definable — and validly, see: definable
Essentially undefinable —, see: undefinable
Hereditarily undefinable —, see: undefinable
Undefinable —, see: undefinable
Expression, 1, 43
Quantifier — 1
Extension
Minimal —, see: minimal
— of a model, 4f., 7ff., 11f., 14f., 17ff.
— of a theory, 5, 9, 11, 15f., 22f., 27ff., 35, 39f., 42f., 48f., 55f.
Extensionality, 23

Family of classes, 20
Faster than, 50
Finite(ly) 1, 29, 32, 36, 39, 42ff.
First-order predicate logic, 1
Formalization, see: standard formalization
Formalize, 1, 25, 41
Formula, 1f., 5f., 8, 11, 13, 30, 43
Atomic —, see: atomic
Coumpound —, see: compound
Free variable, 1f., 5f., 8ff., 50, 53
Function
Constant —, 10

Generating property, 26f.
Gödel, K., 49, 58
Theorem of Löwenheim-Skolem-—, see: theorem

Henkin, L. A., 19f., 28, 58
Hereditarily undecidable, see: undecidable
— undefinable explicitly, see: undefinable
Hereditary, 56
Heyting, A., 58

Hilbert, D., 9, 58
Hold, 25, 27, 41, 53

Identity, 1
Incomplete, 22
Independence, 20, 29f.
Independent, 33, 36, 43
Indirect method, see: method
Individual, 2
— constant, 1f., 5ff., 9f., 12ff., 25, 27f., 41f., 56
Induction
complete —, 44
Inference
Rules of —, 32
Infinite, 19, 36
Integer, 19, 25, 32, 34, 36
Intensionality, 23
Interpretation, 4f., 7ff., 11f., 14ff., 26, 33f., 37f., 44f., 47, 51ff.
Suitable —, see: suitable

Kemeny, J. G., 20, 28, 49, 58
Kleene, S. C., 21, 58

Lindenbaum, A., 16, 59
Logical constant, 1, 10, 19, 29, 50
Löwenheim, L.
Theorem of —-Skolem-Gödel, see: theorem

McKinsey, J. C. C., 58
Metamathematical(ly), 43f.
Method, 14f., 20, 22, 30, 51
Decision —, 36
Direct —, 22
Indirect —, 22
One-model —, 20ff.
— of Padoa, 4, 14f., 19ff.
Two-models —, 20f.
Minimal extension, 53f., 57
Model, *passim*
Characterization in terms of —s, see: characterization
Extension of a —, see: extension
Non-classical —, see: non-classical
One-— method, see: method
Two-—s method, see: method
—- theoretical, 11, 18
Mostowski, A., 1, 6, 16, 22f., 25, 40, 42, 59
Multiplication, 21f., 25

Name, 1
Natural number, 21, 25, 34, 41, 44
Negation, 1
Notation, 1
Notion, 4f., 7ff., 11ff., 16ff., 20, 34, 38, 44, 47
Number, see: natural numbers
Non-classical model, 20
Non-logical axiom, 42f.
— constant, *passim*
— principle, 36

Operation, 25ff., 33f., 37, 41, 45, 51ff., 57
— constant, 1f., 5ff., 12ff., 25f., 29, 33, 41f., 49f., 51, 53, 56f.
Ordered, 45
Ordering, 34
Simple —, 19f.
— principle, 19f.
Order-type, 34

Padoa, A., 4, 11, 14, 20, 58
Method of —, see: method
Parenthesis, 1, 43, 46
Polynomial, 46f., 50, 53
Predicate, 1
— logic, 1
variable —, see: variable
Presburger, M., 21, 25ff., 36, 58
Principle
Non-logical —, see: non-logical
Ordering —, see: ordering
Problem
Definition —, 3, 6, 10, 14, 22, 32, 35, 55
Proof of undecidability, 22, 40
— of undefinability, 14ff., 22, 54
Property,
The closure —, see: closure

Generating —, see: generating
Provable, 32f., 36, 44f.

Quantifier expression, 1
—-free, 27, 44

Realization, 25, 41
Recursive, 26f., 30, 42
Recursiveness, 27
Reflexive, 20
Regular, 25, 38
Relation, 7f., 19f., 25, 41, 52
— constant, 1ff., 14, 25, 29, 41, 50
corresponding — constant, see: corresponding
Robinson, Julia, 48, 58
Robinson, R. M., 1, 6, 16, 22f., 25, 40, 42, 59
Rosser, Barkley, 21, 59
Rules of inference, 32

Satisfy, 25, 27, 42
Sentence, *passim*
Characteristic —, see: characteristic
Sentential connective, 1
Sequence, 42, 46f.
The Ackermann —, 41, 44, 48f., 56
Set, 19f., 25ff., 30, 41f., 45
Simple ordering, see: ordering
Skolem, Th., 20, 59
Theorem of Löwenheim-—-Gödel, see: theorem
Standard formalization, 1, 19, 25f., 41
Structural(ly), 7, 46
Subset, 34, 37
Subterm, 47
Subtheory, *passim*
Suitable interpretation, 28, 52, 57
Symbol, 1, 7, 25, 32f., 37, 41, 45
Technical —, 1, 43, 46
Symmetric, 20
System, 25, 41
Axiom —, see: axiom

Tarski, A., 1, 6, 16, 20, 22f., 25, 30, 40, 42, 59
Technical symbol, see: symbol
Term, 1, 9ff., 29f., 43, 46f., 50
Atomic —, see: atomic
Compound —, see: compound
Theorem of Beth, 4, 18, 20
— of Löwenheim-Skolem-Gödel, 20
Theory, *passim*
Extension of a —, see: extension
Transformation, 27
Transitive, 20
Translate, 17f.

Unary, 25, 29, 41
Undecidability, 22, 30, 40
Essential —, 40
Proof of —, see: proof
Undecidable, 21, 23, 40, 42
Essentially —, 16, 23, 25, 40
Hereditarily —, 23
Undefinability, 20ff., 32, 35, 40, 51
Equational —, 13f., 41, 49, 51, 54, 56
Essential —, 35, 39, 56
Explicit —, 13, 15, 51, 56
Proof of —, see: proof
Undefinable
Essentially — equationally, 10, 12ff., 50 f., 53ff.
Essentially — explicitly, 3, 6, 12ff., 22ff., 34f., 38ff.
— equationally, 10, 12, 14, 50f., 53, 55f.
— explicitly, 3, 6, 12ff., 21ff., 34f., 38ff., 44, 47ff., 51, 53f., 57
Hereditarily — explicitly, 22, 48
Union, 3f., 54, 57
Unity, 21
Universe, 4, 19, 26, 33, 37, 45, 51, 53

Variable, 1, 6ff., 12
Free —, see: free
— predicate, 1
Valid(ly), *passim*
Definable equationally and —, see: definable
Definable explicitly and —, see: definable
Validity, 2, 6, 25, 41

STELLINGEN

I

Zij $f(y_1, \ldots, y_n)$ de Fourier-getransformeerde van een eindige, niet-negatieve Borel-maat in de n-dimensionale ruimte R^n (y_j reëel).

Zij verder $w_k = (w_{k1}, \ldots, w_{kn}) \in R^n$, $(k = 1, \ldots, N)$.

Neem aan, dat $g_k(t) = f(tw_{k1}, \ldots, tw_{kn})$ met $-\varepsilon \leqslant t \leqslant +\varepsilon$ kan worden voortgezet tot een functie, die continu is in de rechthoek $|\mathrm{Re}(t)| \leqslant \varepsilon$, $a_k \leqslant \mathrm{Im}(t) < b_k$, en analytisch in het inwendige ($\varepsilon > 0$, $a_k \leqslant 0$, $b_k \geqslant 0$, $k = 1, \ldots, N$). Zij tenslotte C de kleinste convexe verzameling in R^n met $a_k w_k \in C$, $b_k w_k \in C$, $(k = 1, \ldots, N)$. Dan kan $f(y_1, \ldots, y_n)$ voortgezet worden tot een functie $g(z_1, \ldots, z_n)$, die continu is in het gebied $\{[\mathrm{Im}\,(z_1), \ldots, \mathrm{Im}\,(z_n)] \in C\}$ en analytisch in het inwendige.

II

De karakterisering door N. H. McCoy:

"Een ideaal $\mathfrak{p}$ in een willekeurige ring R is dan en slechts dan een priem-ideaal in R als $aRb \in \mathfrak{p}$ impliceert $a \in \mathfrak{p}$ of $b \in \mathfrak{p}$" verdient een ruimere belangstelling in verband met de toepassingsmogelijkheden van stellingen voor commutatieve ringen op algemene ringen.

N. H. McCoy: Prime Ideals in General Rings. *American Journal of Mathematics,* 71 (1949), pp. 823 ff.

III

Ten onrechte beweert Borel, dat de bol de som is van drie disjuncte delen, A, B en C, welke voor geschikte rotaties φ en ψ voldoen aan de relaties:

$$\varphi A = B + C,$$
$$\psi A = B,$$
$$\psi B = C.$$

Bovendien is zijn bewijs van de niet-houdbaarheid van het keuze-axioma onjuist.

E. Borel: *Les paradoxes de l'infini.* Paris, 1946.

IV

Een monotoon stijgende binaire functie "p" kan niet expliciet worden gedefinieerd met behulp van "1", "S" en "$>$" met betrekking tot de rekenkunde der natuurlijke getallen geformaliseerd met standaard-formalisatie.

Zie dit proefschrift, Theorem 2, p. 35. Theorem 2 kan in de zin van deze stelling worden gegeneraliseerd.

V

Het resultaat van het eerste onderzoek met het synchro-cyclotron van CERN te Genève (experimentele bevestiging van Yukawa's theorie van 1935 aangaande het verval der π-mesonen) heeft duidelijk het belang voor de wetenschap van een internationale samenwerking aangetoond.

VI

Het "Veiligheidsbesluit ioniserende stralen" van 20 maart 1957 geldt met betrekking tot de beveiliging van arbeiders tegen de gevaren van toestellen of stoffen, die ioniserende stralen uitzenden. In artikel 10,3 wordt een maximum doseringssnelheid vastgesteld aan de oppervlakte van televisietoestellen. Het is in dit verband bevreemdend, dat een dergelijke maatregel ontbreekt geldend met betrekking tot de beveiliging van iedereen.

Staatsblad van het Koninkrijk der Nederlanden, **116** (18 april 1957).

VII

De constructie van mathematische modellen en de beschikbaarheid van electronische reken- en administratiemachines heeft het mogelijk gemaakt op vereenvoudigde wijze de activiteiten van leiders van bedrijven in een markteconomie te simuleren. Het daaruit ontwikkelde "Business Game" wordt door de ontwerpers aanbevolen als een aanvullend leermiddel bij de scholing van het hogere bedrijfskader. Meer dan de opleiding is echter de selectie met deze nieuwe mogelijkheid gediend.

Zie bijvoorbeeld: *IBM Decision-Making-Laboratory*, deel I: Handleiding voor deelnemers. Amsterdam, 1959.

VIII

De wijze, waarop H. Reichenbach de mening verdedigt, dat de relativiteitstheorie van Einstein behalve als een fysische theorie ook als een filosofische theorie moet worden beschouwd, is niet overtuigend.

H. Reichenbach: The Philosophical Significance of the Theory of Relativity, in *Albert Einstein: Philosopher-Scientist, The Library of Living Philosophers.* New York, 1949, pp. 289ff.

IX

Wijsgerige beschouwingen, die een "monde en soi" pretenderen te beschrijven, kunnen zich bij de behandeling van verschijnselen der anorganische natuur niet beroepen op resultaten der fysica, tenzij zij eerst het bewijs hebben geleverd, dat ook deze resultaten betrekking hebben op een "monde en soi". Een zodanig bewijs wordt door de resultaten der fysica op zichzelf niet geïmpliceerd.

M. Merleau-Ponty: *Phénoménologie de la perception.* Paris, 1945.

X

De beschouwing van de fysica als een *verschijnsel* in de zin der fenomenologische methode maakt het noodzakelijk de verschillende soorten van verschijnselen aan een gradatie te onderwerpen. Daarbij komt aan het verschijnsel fysica een hogere *graad* toe dan aan de verschijnselen, die door de fysica worden geordend. Deze gradatie is ook in ander opzicht vruchtbaar voor natuurfilosofische beschouwingen.

XI

Bij pogingen om tot een formalisering te komen van theorieën uit de natuurkunde en scheikunde verdient de mereologie van St. Leśniewski bijzondere aandacht.

XII

De opvatting van P. Lorenzen over de taak van de wijsbegeerte der wiskunde (opsporing en kritiek der "Vormeinungen", welke aan de hedendaagse wiskunde ten grondslag liggen) is te beperkt.

Zijn mening, dat de kritiek op het principe van het uitgesloten derde de moderne wiskunde niet op beslissende wijze aantast, is aanvechtbaar.

P. LORENZEN: Wie ist Philosophie der Mathematik möglich? *Philosophia Naturalis*, 4 (1957), pp. 192 ff.

XIII

De discussie tussen N. Rescher, J. L. Mackie en L. Goddard over het bestaan van "willekeurige individuen" vooronderstelt een probleem, dat reeds tevoren door E. W. Beth volledig tot oplossing was gebracht.

N. RESCHER: Can there be Random Individuals? *Analysis*, 18 (1957–1958), pp. 114ff.
L. GODDARD: Mr. Rescher on Random Individuals. *Analysis*, 19 (1958–1959), pp. 6f.
J. L. MACKIE: The Rules of Natural Deduction. *Analysis*, 19 (1958–1959), pp. 27ff.
E. W. BETH: *Semantic Entailment and Formal Derivability*. Amsterdam, 1955.

XIV

Bij de opleiding van leraren in de wis- en natuurkundige vakken dient aandacht te worden geschonken aan de studie van de methodeleer. Meer in het bijzonder behoort de studie van de mathematische logika en van het wiskundig grondslagenonderzoek een onderdeel te vormen van de opleiding tot leraar in de wiskunde.