The Journal of Symbolic Logic
Volume 16, Number 4, December 1951

ON THE INTERPRETATION OF NON-FINITIST PROOFS—PART I

G. KREISEL

1. The purpose of the present paper is to formulate the problem of non-finitist proofs, and to solve it for certain extensions of the predicate calculus, and for analysis with the exclusion of the theory of sets of points.[1] The corresponding problem for general formal systems is discussed in another publication [1].

To fix ideas we introduce the problem by examples from analysis. The general formulation is given in the text. Also, we shall use in the introduction the concepts *decidable*, *verifiable*, *finitist* without much formal explanation, because the reader is probably familiar with them, and they are defined early on in the text.

The paper presupposes some knowledge of the methods and results in the theory of proofs. These enable one to state the general problem rather more precisely.

Introduction. 2. To understand the difference between *finitist* and *non-finitist* proofs, recall three types of formulae in arithmetic:

(i) Formulae without variables, e.g. $6 < 2$, $5 = 3 + 2$, which can be decided systematically; they are made up of the predicate symbols $=$, $<$, and recursive terms without variables. By the theory of the propositional calculus any formula made up of such numerical formulae joined together by the operations of the propositional calculus, is also systematically decidable. The so called elementary proofs of number theory without variables, [2], I, 295, are made up of numerical formulae.

(ii) Formulae whose only variables are *free* variables for numbers and functions; these are still decidable in the sense that when numerals $0, 1, 2, \cdots$ are substituted for the individual variables, and computable functions for the function variables, the resulting formula is one of (i), and therefore decidable. If

$$A[a_1 \cdots a_n ; \cdots f(a_1 \cdots a_n) \cdots] \qquad 2.1$$

is a free variable formula, $\mathfrak{n}_1 \cdots \mathfrak{n}_n$ any set of numerals, and $\cdots f_0(a_1 \cdots a_n)$ computable functions, we call

$$A[\mathfrak{n}_1 \cdots \mathfrak{n}_n ; \cdots f_0(\mathfrak{n}_1 \cdots \mathfrak{n}_n) \cdots] \qquad 2.2$$

particular cases of 2.1. 2.1 is *verifiable* if any (= every) particular case is correct. Note that this (informal) definition of the concept 'verifiable' makes no mention of the system in which 2.1 is to be proved.

Received December 9, 1949. Passages added November 1950 are indicated in the text. The paper will be published in two parts.

[1] (*Added November 1950.*) The part of analysis that we consider can be developed in simple extensions of what is usually called *pure number theory with induction*, e.g. in [10]. While, as Gentzen points out on top of p. 533, 'non-finitist' arguments are hardly used in pure number theory, they occur constantly in elementary analysis, already with the convergence of sequences. Thus the results of the present paper derive their interest from applications to analysis rather than from the elementary theory of numbers.

(iii) Formulae containing *bound* variables, e.g.,

$$(x)(Ey)(z)[f(x, y) = 0 \vee f(x, z) \neq 0]$$

where $f(a, b)$ is a function symbol. For convenience, we reduce these formulae to their prenex form, that is

$$(x_{11}) \cdots (x_{1n_1})(Ey_{11}) \cdots (Ey_{1m_1}) \cdots (x_{n1}) \cdots (x_{nn_n})(Ey_{n1}) \cdots (Ey_{nm_n})A(x_{11} \cdots x_{nn_n} y_{11} \cdots y_{nm_n}) \qquad 2.3$$

where $A(a_{11} \cdots a_{nn_n} b_{11} \cdots b_{nm_n})$ is a (decidable) free variable formula; such a reduction is always possible in formalisms containing the predicate calculus, and only such formalisms will be considered.

3. We shall denote decidable formulae, 2(i) and 2(ii), by *Italic* capitals, others by German capitals.

4. Typical cases of *finitist arguments*, [2], I, 20–31, establish verifiable free variable formulae 2.1. Their *application* is that they are *schemata* for elementary proofs of particular cases 2.2 of 2.1. Proofs in the usual free variable formalisms with various forms of induction are finitist: that they really can be used as schemata is shown by the relevant consistency proofs in [2], I, 298–299, which describe in detail how proofs with variables of 2.1 are to be converted into an elementary proof without variables of 2.2: numbers are substituted for the variables, and an application of induction is replaced by a sequence of implications.

Our thesis is: we do not regard an argument as finitist if it contains formulae with bound variables.[2]

5. Among non-finitist proofs one distinguishes conveniently those proving free variable formulae where the *proof* contains bound variables, and those proving bound variable formulae 2.3. In the former case it is clear what is to be meant by saying that the formula is correct, namely that it is verifiable, only there is no general method at hand for converting the non-finitist proof into an elementary proof of particular cases of the formula. In the latter case it is not clear what is to be meant by saying that the formula is correct, and typical cases of non-finitist proofs mentioned in the literature are of this second type.

The *naive interpretation* of a proof of 2.3 is this: the proof should 'provide' computable[3] functions

---

[2] This is of course no characterization of the concept of *finitist proof* since it is not said what sequences of free variable formulae are to be called a finitist proof. By the familiar diagonal argument there can be no question of a recursive enumeration of finitist proofs.

The condition given, which seems natural, is not satisfied by intuitionist formalisms which contain universal formulae as premisses in implications. The reason for introducing these formulae evidently is that it was thought that no free variable formula can 'mean the same' as a formula 2.3: this is true (when made precise, e.g., by the notion of *interpretation*) if only free individual variables are used, but not if free function variables and recursive functionals are introduced.

[3] Trivially, there are functions $\mathfrak{t}$, e.g. those defined by the *selection* symbol ϵ, or by the least number symbol μ, so that $A[a_{11} \cdots a_{nn_n} \cdots \mathfrak{t}_{ij}(\cdots a_{rs} \cdots) \cdots]$ can be proved. Hence the requirement for *computable* functions. 'Computable' means quasi-recursive throughout.

$\beta_{ij}(\cdots a_{rs} \cdots), \quad 1 \leq i \leq n, \quad 1 \leq j \leq m_i, \quad 1 \leq r \leq i, \quad 1 \leq s \leq n_r$

so that

$$A[a_{11} \cdots a_{nn_n} \cdots \beta_{ij}(\cdots a_{rs} \cdots) \cdots] \qquad 5.1$$

is verifiable. Such a system of functions β_{ij} we call a computable *Erfüllung* of 2.3.

The naive interpretation is tempting because of the ordinary reading of 2.3 as: whatever values of $x_{11} \cdots x_{1n_1}$ may be chosen, there are $y_{11} \cdots y_{1m_1}$, so that whatever values of $x_{21} \cdots x_{2n_2}$ may be chosen, $\cdots A(x_{11} \cdots x_{nn_n} y_{11} \cdots y_{nm_n})$ holds.

The paradox[4] of non-finitist proofs is that the naive interpretation is *false* (for proofs in elementary classical analysis): in appendix I it is shown of a certain primitive recursive sequence of rationals that the naive interpretation of the theorem, *a bounded monotone sequence of rationals converges*, is false. In other words, the proof of 2.3 need not have the application which the ordinary reading of 2.3 suggests.

Historical remarks on the problem. 6. The well known consistency formulation of the problem of non-finitist proofs, given by Hilbert, is based on two considerations:

(i) To make precise statements about non-finitist proofs of a branch of mathematics, one has to describe the proofs to be considered; this is done by the *formalization* of proofs (as initiated by Frege).

(ii) The objectionable feature of non-finitist proofs, according to [2], I, bottom of p. 43, is that their *application* is obscure; consequently, the problem is to make clear how they are applied, and this is to be done by finitist methods.[5]

---

[4] This situation was recognized by Kronecker, Brouwer, and others long before it could be stated in the precise form given below: with the modern theory of proofs the precise formulation is easy.

[5] This precept is not always followed, even in [2]. For consider ω-consistency, as formulated in [2], II, 272: a system Σ of number theory is ω-consistent if $\neg(x)\mathfrak{A}(x)$ cannot be proved in Σ when $\mathfrak{A}(0), \mathfrak{A}(1), \cdots$ can all be proved in Σ; in symbols, let a be the number of the formula $\mathfrak{A}(b)$, $e(a)$ the number of the formula $\neg(x)\mathfrak{A}(x)$, $s(n, m)$ Gödel's substitution function, and $Prov_\Sigma(m, n)$ the recursive formula which holds if and only if m is the number of a sequence of formulae which are a proof in Σ of the formula with number n. Then ω-consistency of Σ means

$$(x)(Ey)Prov_\Sigma[y, s(x, a)] \rightarrow (z) \neg Prov_\Sigma[z, e(a)],$$

or, in prenex form,

$$(z)(Ex)(y)\{\neg Prov_\Sigma[y, s(x, a)] \vee \neg Prov_\Sigma[z, e(a)]\} \qquad 6.1$$

which is of form 2.3, that is, of the very form whose interpretation is at the root of all the trouble.

To avoid such situations we adopt the following standpoint: *All logical (=metamathematical) problems are to be so formulated that the solution is arithmetized by a free variable formula.*

Thus, instead of a proof of 6.1, we require, *when $\mathfrak{A}$ is a free variable formula*, a function $g(z, a)$ so that if z is the number of a proof of $\neg(x)A(x)$, $\neg A[g(z, a)]$ holds (naive inter-

These two considerations are used in [2], I, to motivate the consistency formulation for a system Σ: either, of two formulae $\mathfrak{A}$ and $\neg\mathfrak{A}$ at most one is provable in Σ, or, not all formulae of Σ are provable in Σ.

Now, this formulation must be considered appropriate if one is interested in the application of a mathematical system to a branch of science where with its empirical propositions are associated formulae of the mathematical system; and then true empirical propositions are to be picked out (those associated with provable formulae of the system): thus a classification of empirical propositions is given. Provided the mathematical system is *consistent*, the classification is non-trivial, that is, some empirical propositions are not picked out. It is here assumed that there is, e.g., a formula $\mathfrak{A}$ so that both $\mathfrak{A}$ and $\neg\mathfrak{A}$ are associated with empirical propositions.

Further, the consistency formulation is appropriate if one only considers proofs of free variable formulae, e.g. on the ground that for others the notion of correctness has not been defined.[6] And then in the case of a consistent system of number theory which includes elementary proofs of number theory, a proved free variable formula is verifiable: here non-finitist proofs may be considered short cuts; this is a perfectly practical application, no 'mere parlour game with symbols.'

7. **In neither of these applications of the system of analysis is the absence of an Erfüllung of proved formulae 2.3 paradoxical.** If one wishes to discuss *this* paradox (this unexpected feature of proofs in classical analysis), the consistency formulation misses the mark. Whatever the interest of consistency *proofs* may turn out to be in this connection, further work is required to make clear in what way they help with the paradox.

8. (*Added Nov. 1950.*) What the critic of the non-constructive system is after was probably stated for the first time explicitly—though not precisely—by Gentzen in the last two paragraphs on p. 594 of [10]: a *finitist sense* is to be given to proved formulae of the non-finitist system considered. Roughly speak-

---

pretation of 6.1). This replaces then the roundabout impossibility formulation and the use of universal premisses.

It is shown in a forthcoming paper that number theory Z_μ is not ω-consistent in this strong sense when $\mathfrak{A}$ may be a bound variable formula, and a weaker ω-consistency is established by applying the results of sections IV and V of the present paper.

[6] The consistency formulation is often put in the following way: one ignores bound variable formulae except as methods of proof, and shows that their use is 'harmless' because the free variable formulae proved are verifiable.—It is perhaps worth while to consider a consistency formulation of a logical problem which was solved over 100 years ago: the use of complex numbers. The problem arose because the operation with the symbol i was added to the operations of ordinary algebra, but no application was given to formulae containing i. Now suppose proofs of algebra with and without the symbol i to be formalized. Then a consistency proof of the former might show how a formula without i which has been proved by the use of i, can also be proved in the algebra without i. But nothing is said of formulae which contain i. The actual solution of the problem of complex numbers did more: when these are replaced by pairs of numbers, or matrices, and suitable definitions of addition and multiplication are introduced, *all* formulae containing i get an interpretation.

ing, it should be shown how information can be obtained from proofs of $\mathfrak{A}$ in the system, concerning free variable formulae A; it is desirable that the *forms* of the formulae $\mathfrak{A}$ and A be closely connected, that is, A should be read off easily from $\mathfrak{A}$.

Though this program was not formulated as a precise mathematical problem, several investigations are relevant.

(1) Hilbert and Bernays supplement the consistency proof for the predicate calculus by their *Kriterien der Widerlegbarkeit im reinen Prädikatenkalkul*, [2], II, 172–174. We believe that our problem of *interpretation* formulates Gentzen's problem, and it will be seen that the *Kriterien* constitute an interpretation in our sense.

(2) Gödel, in [11], gave an intuitionist sense to formulae of classical number theory, but not finitist in the sense of para. 4: $\mathfrak{A} \vee \mathfrak{B}$ is replaced by

$$\neg(\neg\mathfrak{A} \;\&\; \neg\mathfrak{B}),$$

$(Ex)\mathfrak{A}(x)$ by $\neg(x)\neg\mathfrak{A}(x)$, $\neg\mathfrak{A}$ by $\mathfrak{A} \rightarrow 1 = 0$. Then, from a proof in classical number theory of a formula $\mathfrak{A}$ one gets an intuitionist proof of the modified version of $\mathfrak{A}$. This is read off easily from $\mathfrak{A}$, but it is not a free variable formula. Note that this intuitionist version of $(Ex)A(x)$, even when $A(a)$ is a free variable formula, is $(x)\neg A(x) \rightarrow 1 = 0$, and not a formula $A(\mathfrak{n})$ with a numeral $\mathfrak{n}$.

(3) Gentzen at the end of the first version of his consistency proof suggests that his reduction rule should be considered as giving a *finitist sense* to bound variable formulae of number theory, but observes that this *finitist sense* can turn out to be pretty complicated even for simple bound variable theorems, so that the connection between the theorem and its finitist sense is not transparent. In the second version of his consistency proof in [12], he does not return to this question.

(4) S. C. Kleene and D. Nelson in a series of papers on 'realizations' of formulae of number theory are interested in a related problem, but rather from the point of view of intuitionist number theory.

9. Our aim in this paper is to give a 'finitist sense,' or as we prefer to call it, an *interpretation* to systems of number theory (various extensions of Z_μ, [2], II, 289) whose formulae are recursively enumerable.

A computable function $f(n, a)$ is called an interpretation of the system Σ if

(α) $f(n, a)$ is the number of a free variable formula A_n when a is the number of a formula $\mathfrak{A}$ of Σ (some Gödel numbering being assumed),

(β) if $\mathfrak{A}$ is proved in Σ, from the proof we find an $\mathfrak{n}$ so that $A_\mathfrak{n}$ is verifiable,

(γ) if $\neg\mathfrak{A}$ is proved in Σ, for each n we find a substitution for the (individual and function) variables of A_n which makes A_n false,

(δ) if $\mathfrak{B}$ is proved from $\mathfrak{A}$ in Σ, we find a $g(n)$ so that $B_{g(n)}$ is verifiable if A_n is verifiable.

(See para. 20 for more details.)

This definition of an interpretation of $\mathfrak{A}$ by a (disjunctive) sequence of A_n is adopted only after it is shown in appendix II that for certain informal systems Σ there is no interpretation of $\mathfrak{A}$ by a single free variable formula satisfying (α)–(δ). It is these informal systems which make the problem interesting: they are systems in which any verifiable free variable formula may be added as an axiom.

It turns out that Hilbert's ϵ-substitution method is a particularly suitable tool in the construction of interpretations.

Summary of principal results. 10. In a formalized system Σ in which proofs are recursively enumerable an interpretation of the formula $\mathfrak{A}$ is: *the formula* $\neg\mathfrak{A}$ *is not provable in* Σ, arithmetized by a single free variable formula.

(No attempt is made in this summary to state the results precisely.)

11. From Herbrand's theorem one gets an interpretation of formulae $\mathfrak{A}$ in certain extensions of the predicate calculus by sequences A_n, where $\mathfrak{A}$ and A_n are related by very simple conditions, (1)–(3) of [2], II, 151. The verification that the sequences constitute an interpretation is trivial (section II).

12. From the first ϵ-theorem of the predicate calculus one gets another interpretation of formulae of the predicate calculus. It amounts to this: from a proof of $\mathfrak{A}$ one gets a rule which enables one to find a counter-example to any system of constants and functions which might be proposed to satisfy the negation of $\mathfrak{A}$. (In the case 2.3, take $n_i = m_i = 1$, $n = 2$, form $\neg\mathfrak{A}$, reduce it to prenex form

$$(Ex_1)(y_1)(Ex_2)(y_2)\neg A(x_1x_2y_1y_2). \qquad 12.1$$

An Erfüllung of 12.1 would be a number $\mathfrak{x}_1$, and a function $\mathfrak{x}_2(a)$ so that

$$\neg A[\mathfrak{x}_1\,\mathfrak{x}_2(a_1)a_1a_2] \qquad 12.2$$

is verifiable. If $\mathfrak{A}$ is proved in the relevant extension of the predicate calculus, we find numbers $\mathfrak{y}_1$, and $\mathfrak{y}_2$ which make 12.2 false. The rule for finding these numbers for given $\mathfrak{x}_1$, and for a given function $\mathfrak{x}_2(a)$ is provided by simple *functionals of the predicate calculus.*)

The verification that this interpretation satisfies (γ) and (δ) is not quite trivial (Section III).

In both interpretations we find a primitive recursive function $f(a)$ so that $A[a, f(a)]$ is verifiable when

$$(x)(Ey)A(x, y)$$

with *free variable* $A(a, b)$ has been proved.

13. The no-counter-example interpretation of para. 12 is developed for certain extensions of number theory, the consistency proof in [3] and some (rather general) theorems on the ϵ-substitution method in section IV being the principal weapons. The functionals used are *ordinal recursive functionals* of finite order, and discussed in section V.

Applications. 14. The applications of the work to 'ordinary' mathematics depend on the fact that a large part of analysis can be developed in the extensions of Z_μ which are discussed in the text; roughly speaking, those parts where rational and not continuous *bound* variables are used; in particular, the theories of sequences or functions of a complex variable within their circles of convergence.

The practical use of the systematic elimination methods such as those of [3] or of the text can never consist in writing out a textbook proof in some formal system (more precisely: a formalized proof suggested by the textbook argument), and then applying the systematic method: it would be forbiddingly long. The practical value of these logical investigations consists rather in telling you what to look for:

E.g., if one has to prove a bound variable formula in elementary analysis, one tries to find a counter-example to any proposed Erfüllung of its opposite. Though this *reductio ad absurdum* method is often used, proofs became rather more systematic if one can state the *form* of the counter-example (functionals of section V) at the beginning of the subject.

Again, if one has proved a formula $(x)(Ey)A(x, y)$ and wishes to find a computable Erfüllung, one applies the method just described to the lemmas used in the course of the proof, and then proves the theorem from their finitist versions, when quantifiers are eliminated in the process. Often the following application of the results of para. 13 is useful: if $(x)B(x) \rightarrow (x)(Ey)A(x, y)$ has been proved, and A and B are decidable, there are computable functions $f(a)$, $g(a)$ so that $(x)[x \leq g(a) \rightarrow B(x)] \rightarrow A[a, f(a)]$ is verifiable, i.e. for a given a the premiss need hold for a finite number of arguments of B only (if one *looks* for these functions in the textbook proof, one can often define them without applying the systematic method). These techniques can be used on Littlewood's theorem that $\pi(n) - li(n)$ changes sign infinitely often, where $\pi(n)$ denotes the number of primes not exceeding n, and $li(n)$ the logarithmic integral. The proof was believed to be 'non-constructive' [4], or to require 'new ideas' of proof to make it constructive [5]. Since the proof was not developed in a formal system there can be no question of applying the results of the present work to it in a precise sense. But if one examines the official proof in [4] (and the usual proofs of the standard theorems on complex variables used), and the restatement in section VI, it might fairly be said, I think, that it applies the principles of para. 13 in a straightforward manner without introducing 'new ideas' of proof (cf. particularly the crucial theorem H of [4]).

(*Added Nov. 1950.*) The work of section VI can be made much more elementary if certain very simple lemmas are used instead of the lemmas A and B.

15. More 'logical' applications depend on the solution of an *unsolved problem.*

In the formulation of the problem of interpretation the A_n are only required to be verifiable, and no restriction is imposed on the free variable calculus in which they are to be proved. While it is known that the relevant A of paras. 11 and 12 can be proved by the elementary calculus with free variables, no analogous result has been established for full number theory. In para. 38 we set up a formalism which probably meets the case.

If it does, it would be a free variable formalism in which classical number theory can be interpreted in the sense of para. 9: from a classical proof of $\mathfrak{A}$ we get a free variable proof of A, $\mathfrak{A}$ and A being quite simply related. We would have a development of elementary analysis without bound variables. Such a development of mathematical disciplines is, I think, the main aim of intuitionists. Note that this is only one aspect of the problem in which we are interested: the intuitionist approach *ignores* classical proofs altogether, while we are interested in the *relation* between classical and free variable formalisms.

(*Added Nov. 1950.*) Also, such a formalism would throw light on the problem of *direct* proofs in the theory of numbers, e.g. [4]: it seems fairly clear that most theorems of analytic number theory could be proved in one of the extensions of Z_μ that we consider: only rational bound variables need be used, complex variables are replaced by pairs of reals, and rationals by pairs of whole numbers; thus from an analytical proof we should get by a systematic method a free variable proof using whole numbers only. But this is clearly not what is meant by a direct proof such as those of Selberg and Erdös. It would *seem* that what is meant by a direct proof is one in which no (auxiliary) functions and predicates are introduced in the proof other than those occurring already in the statement of the theorem (or possibly certain 'familiar' ones). Now it is known that in the elementary proofs with free variables such functions and predicates can always be eliminated: not only explicitly defined ones, which is trivial, but, e.g., recursively defined ones; also in Z_μ where they can be replaced by μ-symbol expressions. But in free variable formalisms of number theory this cannot be done generally, and the condition above may serve as a characterization of the notion of *direct* proof: it seems to express fairly well what Gauss meant by: throughout a direct proof you keep the problem before your mind.

For an application of the *methods* of the paper, cf. footnote 5.

Trivial interpretation. 16. Let Σ be a consistent formalized system containing negation, in which also the *Deduktionstheorem* holds, [1], II, 387. Let $\text{Prov}_\Sigma(m, n)$ be a decidable formula which holds if and only if m is the number of a proof in Σ of the formula with number n, and $e(a)$ a computable function so that $e(a)$ is the number of the formula $\neg \mathfrak{A}$ when a is the number of $\mathfrak{A}$. Then

$$\neg\ \text{Prov}_\Sigma[m, e(a)] \qquad 16.1$$

with free variable m is the formula associated with $\mathfrak{A}$: it is an interpretation, for

(β) if $\mathfrak{A}$ is proved and Σ is consistent, 16.1 is verifiable,

(γ) if $\neg\mathfrak{A}$ is proved in Σ, by a proof with number $\mathfrak{m}$, $\mathfrak{m}$ substituted in 16.1 makes 16.1 false,

(δ) if $\mathfrak{B}$ is proved from $\mathfrak{A}$, and the variables of $\mathfrak{A}$ are bound, the formula $\mathfrak{A} \to \mathfrak{B}$, and therefore $\neg\mathfrak{B} \to \neg\mathfrak{A}$ can both be proved in Σ.

By [2], II, bottom of p. 285, the formulae

$$(Ex)\ \text{Prov}_\Sigma[x, e(b)] \to (Ex)\ \text{Prov}_\Sigma[x, e(a)],$$

and thus

$$(x)(Ey)\{\text{Prov}_\Sigma[y, e(a)] \lor \neg\ \text{Prov}_\Sigma[x, e(b)]\}$$

can both be proved in Z_μ.[7] By section V there is an (ordinal) recursive function $f(n)$ so that

$$\neg \text{Prov}_\Sigma[m, e(b)] \vee \text{Prov}_\Sigma[f(m), e(a)].$$

If $\neg \text{Prov}_\Sigma[m, e(a)]$ is verifiable, so is $\neg \text{Prov}_\Sigma[m, e(b)]$.

I. *Preliminary definitions.*

17. In the present section we introduce formal systems in which a problem of non-finitist proofs arises.

Def. 1: Formal system of formulae.

Given a finite scheme for introducing constants, function symbols, and predicate symbols: then any constant is called a term, and a function symbol with a term as argument is also called a term; any expression got by substituting terms in the free places of a predicate symbol is a prime formula, and prime formulae joined together by the operations of the propositional calculus are formulae. We only consider systems which contain the symbol $=$.

By a finite scheme we mean a recursive procedure of deciding of any expression made up of the symbols used whether it is a constant, function symbol, or predicate symbol. Note that because of our formation rules, terms and formulae are recursive concepts.

Def. 2: Decidable formulae $\mathfrak{N}$.

We call a system of formulae systematically decidable if we have a procedure by which any prime formula is given the value T or F, so that if a term $\mathfrak{a}$ in a prime formula P is replaced by a term $\mathfrak{b}$ to give P', where the formula $\mathfrak{a} = \mathfrak{b}$ has the value T, then P and P' have the same truth value. The truth value of a composite formula is got from the truth values of its constituent prime formulae by the usual rules for truth functions (cf. [2] II, 130).

Def. 3: Valuation of terms.

By our definition of term the terms of $\mathfrak{N}$ can be brought into a linear order. Given a term $\mathfrak{u}$, by the decision procedure we can find the first term $\mathfrak{b}$ in the linear order so that $\mathfrak{u} = \mathfrak{b}$ gets T. $\mathfrak{b}$ is called the *value* of $\mathfrak{u}$. (cf. ausgezeichnete Gleichheitswertung of [2] II, 205.)

Def. 4: Computable functions of $\mathfrak{N}$.

Let ϕ be a computable numerical function with n arguments. The values of $\phi_\mathfrak{N}(a_1 \cdots a_n)$ are defined as follows: if the value of a_i is the $n_i{}^{\text{th}}$ term in the ordering of def. 3, then $\phi_\mathfrak{N}(a_1 \cdots a_n)$ is the $\phi(n_1 \cdots n_n)^{\text{th}}$ term.

Similarly the values of functions of functions are defined. Note that the value of $\phi_\mathfrak{N}[\psi_\mathfrak{N}(a_1 \cdots a_n)a_2 \cdots a_n]$ is independent of the order in which it is evaluated.

Def. 5: Verifiable formulae.

(i) Free individual variables: let $A(a_1 \cdots a_n)$ be an expression made up of predicate symbols of $\mathfrak{N}$, terms of $\mathfrak{N}$, and free variables, naturally not belonging

[7] The idea is obvious: if $\mathfrak{p}$ is the number of a proof $\mathfrak{P}$ of the formula $\neg\, \mathfrak{B}$, we get a proof of the formula $\neg\, \mathfrak{A}$ simply by adding the given proof of $\neg\, \mathfrak{B} \rightarrow \neg\, \mathfrak{A}$ to $\mathfrak{P}$. Thus from the number of the proof of $\mathfrak{A} \rightarrow \mathfrak{B}$ we get the number of a proof of $\neg\, \mathfrak{A}$ as a function of $\mathfrak{p}$.

to $\mathfrak{N}$, in some of the free places. If $A(\mathfrak{n}_1 \cdots \mathfrak{n}_n)$ has the value T whatever terms $\mathfrak{n}_1 \cdots \mathfrak{n}_n$ of $\mathfrak{N}$ are substituted for the free variables, then $A(a_1 \cdots a_n)$ is called verifiable with respect to $\mathfrak{N}$, or simply verifiable.

(ii) Free function variables: suppose that if the free individuals in A are replaced by terms of $\mathfrak{N}$, and free function variables by computable functions of $\mathfrak{N}$, the resulting formula has the value T, then A is called verifiable.

Def. 6: Predicate calculus of first order.

This is the 'reine Prädikatenkalkul' of [2], II, 375–380, to which the ϵ-formula [2], II, 13, and the axioms for equality J_1 and J_2 of [2], II, 28, have been added.

We call it of first order to distinguish it from predicate calculi in which bound function variables are used.

18. Classifications of systems considered:

(a) Formal systems consisting of the predicate calculus to which suitable inference schemata—in the sequel, various forms of induction—and a schema for introducing free variable formulae as axioms are added. The predicate and function symbols are assumed to belong to a system $\mathfrak{N}$.

(b) ω-consistent extensions are systems (a) to which *any* verifiable free variable formula of the system may be added as an axiom.

(c) Free function variables are added to the system, and any verifiable free variable formula containing possibly function variables and symbols for computable functions of $\mathfrak{N}$ may be added to (b).

Observe that we do not use explicit definitions: it is shown in para. 29 how explicit definitions of certain computable terms can be introduced in systems (c).

19. By the numbering of formulae of the predicate calculus [2], II the formulae and terms of the systems (a)–(c) are recursive concepts.

Systems (c), superficially connected with the idea of external consistency, are introduced for the following reason: by the theory of computable functions these cannot be enumerated, and to any system $\mathfrak{N}$ which contains primitive recursive functions we can find a primitive recursive formula $A(a, b)$ and a function $f(n)$ so that $A[n, f(n)]$ is verifiable, but for any function *symbol* $\phi(a)$ of $\mathfrak{N}$, $(x)[A[x, \phi(x)]]$ is false: (c) allows us to introduce $(x)(Ey)A(x, y)$ while it is easy to show that (b) does not. Observe that though (c) is an informal system it contains undecidable formulae of the form $(x)(Ey)(z)A(x, y, z)$.

Note. Whenever we speak of proofs in the sequel we mean proofs in one of the systems above.

20. The problem of non-finitist proofs.

Def. 7: Interpretation of proofs in 18(a)–(c).

We call a computable function $f(n, a)$ an interpretation of one of the systems (a)–(c) if the four following conditions hold:

(α) $f(n, a)$ is the number of a free variable formula A_n whenever a is the number of a formula $\mathfrak{A}$ of the system.

(β) From a proof of $\mathfrak{A}$ we find an n so that A_n is verifiable.

(γ) From a proof of $\neg \mathfrak{A}$, to any n we find a substitution for the (free) variables of A_n which makes A_n false.

(δ) If $\mathfrak{B}$ is proved from $\mathfrak{A}$ we find a function $g(n)$ so that if A_n is verifiable, so is $B_{g(n)}$.

Note. To avoid the ambiguity of the definition of ω-consistency pointed out in para. 6, we strengthen condition (δ) as follows: we have a function $h(m)$ so that if a substitution with number m for the variables in $B_{g(n)}$ makes $B_{g(n)}$ false, then $h(m)$ is the number of a substitution for the variables in A_n which makes A_n false; this result is arithmetized without bound variables. Observe also that (δ) implies (γ).

Remark. By appendix II it is not possible to associate with $\mathfrak{A}$ a single A (i.e. $f(n, a) = f(m, a)$) which satisfies the conditions (α)–(δ): this may be expressed by saying that the interpretation of $\mathfrak{A}$ *depends* on the *proof* of $\mathfrak{A}$, but the possible interpretations of $\mathfrak{A}$ are *limited* by the *form* of $\mathfrak{A}$.

II. *First interpretation of the predicate calculus.*

21. Our first interpretation is got by a simple application of Herbrand's theorem, [2], II, p. 158(b).

For definiteness we consider a system of type (b), consisting of the predicate calculus to which verifiable free variable axioms, but no inference rules, have been added. It is also possible to give an interpretation of extensions (c) along the lines below.

With any prenex formula $\mathfrak{A}$ we associate disjunctions $A_1^n \vee \cdots A_{p_n}^n$ where A_i^n is got from $\mathfrak{A}$ by removing the quantifiers of $\mathfrak{A}$, replacing the all-variables by free variables $a_1 \cdots a_k$, and the existence variables by those terms of the predicate calculus which satisfy the relations (1)–(3) of [2], II, 151: the terms are made up of the symbols $a_1 \cdots a_k$, function symbols of $\mathfrak{A}$, and symbols occurring in the axioms that have been added to the predicate calculus.

We show that this interpretation satisfies conditions (α)–(δ):

(α) By the numbering of the predicate calculus and the recursive character of (1)–(3) the numbers of the disjunctions $A_1^n \vee \cdots \vee A_{p_n}^n$ are given by a primitive recursive function of $\mathfrak{a}$ and n, where $\mathfrak{a}$ is the number of $\mathfrak{A}$.

(β) If $\mathfrak{A}$ has been proved by the predicate calculus from the free variable axioms $A_0(a_1 \cdots a_n)$, by Herbrand's theorem, we find a disjunction

$$A_1^s \vee \cdots \vee A_{p_s}^s \vee \neg A_0(\alpha_1^i \cdots \alpha_n^i), \qquad 1 \leq i \leq p_s,$$

α terms of $\mathfrak{N}$, which can be proved by the elementary calculus with free variables, and is therefore verifiable. Since $A_0(a_1 \cdots a_n)$ is verifiable, $\neg A_0(\alpha_1^i \cdots \alpha_n^i)$ is false, and therefore $A_1^s \vee \cdots \vee A_{p_s}^s$ is verifiable.

(γ) Suppose $\neg \mathfrak{A}$ has been proved from the axioms. We have to find a substitution for the variables $a_1 \cdots a_k$ in $A_1^s \vee \cdots \vee A_{p_s}^s$ by terms of $\mathfrak{N}$ which makes this disjunction false. Observe that for any s, $\mathfrak{A}$ can be proved in the predicate calculus from $A_1^s \vee \cdots A_{p_s}^s$; therefore from $\neg \mathfrak{A}$ and the axioms we can prove $(Ey_1) \cdots (Ey_n)[\neg \tilde{A}_1^s \,\&\, \cdots \,\&\, \neg \tilde{A}_{p_s}^s]$ where $\tilde{A}$ is got from A by replacing $a_1 \cdots a_k$ by $y_1 \cdots y_k$. By [2], II, 30, terms of $\mathfrak{N}$ can be found which make $A_1^s \vee \cdots \vee A_{p_s}^s$ false.

(δ) If $\mathfrak{B}$ can be proved from $\mathfrak{A}$ it can also be proved from $A_1^s \vee \cdots A_{p_s}^s$, where the latter is any one of the disjunctions associated with $\mathfrak{A}$. By (β) we find a $g(n)$ so that if $A_1^s \vee \cdots A_{p_s}^s$ is verifiable, so is $B_1^{g(s)} \vee \cdots B_{p_{g(s)}}^{g(s)}$.

Remarks. The interpretation above (and also those given below) has two useful properties in addition to those required by the general definition of interpretation:

(i) A formula without bound variables is its own interpretation.

(ii) If $(Ex_1) \cdots (Ex_n)A(x_1 \cdots x_n)$ has been proved, we find terms $\alpha_1 \cdots \alpha_n$ so that $A(\alpha_1 \cdots \alpha_n)$ is true.

Because of (i) (γ) follows from (δ); for if $\neg \mathfrak{A}$ is proved $\mathfrak{A} \rightarrow 0 = 1$ can be proved provided the variables in $\mathfrak{A}$ are bound, and hence by (δ) we find a substitution for the variables of A_n which makes A_n false.

It should be observed that the proofs of (γ) and (δ) are easy because the negation of the formulae A_n contains only existence symbols for *individuals*, since A_n does not contain free *function* variables. Superficial examination suggests that it is not possible to interpret number theory with induction without the use of (free) function variables. To get familiar with the problems arising in this way we turn to another interpretation of the predicate calculus.

III. *Second interpretation of the predicate calculus.*

22. In the present section we give the no-counter-example interpretation, which requires free variable definitions of negation and the concept 'functional of the predicate calculus.'

23. In the sequel we take $n_i = 1$ and $m_i = 1$ in 2.3 to simplify notation. If the reader prefers to follow out definitions and proofs on some special case, he may take for his standard form of 2.3 the formula

$$(x_1)(Ey_1)(x_2)(Ey_2)A(x_1 x_2 y_1 y_2).$$

Def. 1: Negation I of 2.3.

A computable Erfüllung of $\neg \mathfrak{A}$ by a constant $\mathfrak{x}$ of $\mathfrak{N}$ and computable functions $\mathfrak{x}_i(b_1 \cdots b_{i-1})$, $1 < i \leq n$ is called a negation I of 2.3, that is

$$\neg A[\mathfrak{x}\mathfrak{x}_2(b_1) \cdots \mathfrak{x}_n(b_1 \cdots b_{n-1})b_1 \cdots b_n] \qquad 23.1$$

is verifiable.

If the $\mathfrak{x}_i$ are required to belong to some formal system $\mathfrak{K}$, we say that 23.1 is a negation I in $\mathfrak{K}$ of 2.3.

Def. 2: Negation II of 2.3 *in* $\mathfrak{K}$.

A counter-example of class $\mathfrak{K}$ to any Erfüllung of 2.3, i.e. a set of functionals q_i containing the free variables $f_j(a_1 \cdots a_j)$, $1 \leq j \leq n$ is called a negation II of 2.3 if

$$\neg A[q_1 \cdots q_n f_1(q_1) \cdots f_n(q_1 \cdots q_n)] \qquad 23.2$$

is verifiable.

Note that since we have not defined computable functionals in a general way, the functionals used will always be restricted to a particular class.

Def. 3: First repeated negation of 2.3 (Negation I is false).

If p_i^k, $1 \leq i \leq n$, $1 \leq k \leq l$ are terms of $\mathfrak{K}$ containing the free individual variable a and the function variables $f_2(a_1), \cdots, f_n(a_1 \cdots a_{n-1})$ so that

$$A[a, f_2(p_1^1) \cdots f_n(p_1^1 \cdots p_{n-1}^1)p_1^1 \cdots p_n^1] \vee \cdots \vee A[a, f_2(p_1^l) \cdots f_n(p_1^l \cdots p_{n-1}^l)p_1^l \cdots p_n^l] \tag{23.3}$$

is verifiable, we say the p_i^k are a counter-example to any proposed Erfüllung of $\neg \mathfrak{A}$ by $a, f_2 \cdots f_n$.

We sometimes write $p_{x_1 \cdots x_{n-1}}[f_2(x_1) \cdots f_n(x_1 \cdots x_{n-1}); a]$ where the dummies $x_1 \cdots x_{n-1}$ indicate that the p depend on the functions f and not on their values at a fixed set of arguments.

Remark. We use a *disjunction* in 23.3: either $A(a \cdots p_n^1)$ or $\cdots$ or $A(a \cdots p_n^l)$ in order to be able to restrict our functionals to terms of the predicate calculus. Of course, if a repeated negation 23.3 has been found, it can always be put in the form

$$A[af_2(\mathfrak{p}_1) \cdots f_n(\mathfrak{p}_1 \cdots \mathfrak{p}_{n-1})\mathfrak{p}_1 \cdots \mathfrak{p}_n]$$

where the set $\mathfrak{p}_i$ is defined as the first of the sets of terms p_i^k which make 23.3 true.

Def. 4: Second repeated negation of 2.3 (Negation II is false).

If to any set of functionals $q_{i;x_1 \cdots x_n}^k[g_1(x_1) \cdots g_n(x_1 \cdots x_n); \cdots b \cdots; \cdots f \cdots]$ of $\mathfrak{K}$, where $1 \leq i \leq n$, $1 \leq k \leq l_1$, we find computable functions $\mathfrak{y}$ so that

$$A[\tilde{q}_1^1 \cdots \tilde{q}_n^1 \mathfrak{y}_1(\tilde{q}_1^1) \cdots \mathfrak{y}_n(\tilde{q}_1^1 \cdots \tilde{q}_n^1)] \ \& \cdots A[\tilde{q}_1^{l_1} \cdots \tilde{q}_n^{l_1} \mathfrak{y}_1(\tilde{q}_1^{l_1}) \cdots \mathfrak{y}_n(\tilde{q}_1^{l_1} \cdots \tilde{q}_n^{l_1})] \tag{23.4}$$

is true, where $\tilde{q}$ is got from q by replacing g by $\mathfrak{y}$, we say there is no counter-example in $\mathfrak{K}$ to every proposed Erfüllung of 2.3. The $\mathfrak{y}$ contain $\cdots b \cdots$, and $\cdots f \cdots$ as parameters.

Def. 5: Functionals of the predicate calculus, $\mathfrak{F}_\mathfrak{R}$, *with free individual variables* a_i, $1 \leq i \leq n$, *and free function variables* $f_j(b_1 \cdots b_{n_j})$, $1 \leq j \leq m$.

By induction: A term t of the predicate calculus to which terms of $\mathfrak{R}$ and function variables have been added, belongs to $\mathfrak{F}_\mathfrak{R}$ if

(i) it is one of the variables a_i, $1 \leq i \leq n$, or a term of $\mathfrak{R}$;

(ii) it is got from a term t_1 of $\mathfrak{R}$ with free places by substituting terms of $\mathfrak{F}_\mathfrak{R}$ in these places;

(iii) it is a term $f_j(\mathrm{t}_1 \cdots \mathrm{t}_{n_j})$, $1 \leq j \leq m$, where t are terms of $\mathfrak{F}_\mathfrak{R}$.

Theorem. *If two computable functions of* $\mathfrak{R}$, $f_j^0(b_1 \cdots b_{n_j})$ *and* $g_j^0(b_1 \cdots b_{n_j})$ *have the same values for a certain* finite *set of arguments, the values of a (given) term t of* $\mathfrak{F}_\mathfrak{R}$ *with* f_j^0 *or* g_j^0 *as arguments are equal.*

This is true if t is of form (i). It is true for a t of (ii) if it is true for all the terms of $\mathfrak{F}_\mathfrak{R}$ which are substituted in t_1. It is true of terms (iii) if it is true for all the terms $\mathrm{t}_1 \cdots \mathrm{t}_{n_j}$ *and* $f_j^0(\mathrm{t}_1 \cdots \mathrm{t}_{n_j}) = g_j^0(\mathrm{t}_1 \cdots \mathrm{t}_{n_j})$.

In other words, the values of terms of $\mathfrak{F}_\mathfrak{R}$ depend only on the extensions of f^0 and g^0. This remark is used below.

24. We can now state the interpretation of the predicate calculus:

If $\mathfrak{A}$ is proved we can find a counter-example of the predicate calculus to any Erfüllung of $\neg \mathfrak{A}$.

The proof depends on the following relation between the two definitions of repeated negation:

LEMMA. *If p_i^k are terms of the predicate calculus to which terms of $\mathfrak{N}$ have been added, so that*

$$A[af_2(p_1^1) \cdots f_n(p_1^1 \cdots p_{n-1}^1)p_1^1 \cdots p_n^1] \vee \cdots \vee A[af_2(p_1^l) \cdots f_n(p_1^l \cdots p_{n-1}^l)p_1^l \cdots p_n^l] \tag{24.1}$$

is verifiable, then to any functionals q_j^k of the predicate calculus we can find functions $\mathfrak{y}$ so that

$$A[\bar{q}_1^1 \cdots \bar{q}_n^1 \mathfrak{y}_1(\bar{q}_1^1) \cdots \mathfrak{y}_n(\bar{q}_1^1 \cdots \bar{q}_n^1)] \& \cdots \& A[\bar{q}_1^{l_1} \cdots \bar{q}_n^{l_1} \mathfrak{y}_1(\bar{q}_1^{l_1}) \cdots \mathfrak{y}_n(\bar{q}_1^{l_1} \cdots \bar{q}_n^{l_1})] \tag{24.2}$$

is true.

Conversely, if 2.3 has been disproved, to any formula 24.1 we find $\mathfrak{x}$ so that

$$\neg A[\mathfrak{x}\mathfrak{x}_2(\bar{p}_1^1) \cdots \mathfrak{x}_n(\bar{p}_1^1 \cdots \bar{p}_{n-1}^1)\bar{p}_1^1 \cdots \bar{p}_n^1] \& \cdots \& \neg A[\mathfrak{x}\mathfrak{x}_2(\bar{p}_1^l) \cdots \mathfrak{x}_n(\bar{p}_1^l \cdots \bar{p}_{n-1}^l)\bar{p}_1^l \cdots \bar{p}_n^l] \tag{24.3}$$

is true.

(i) Observe first, by the method of eliminating quantifiers by the ϵ-symbol, [2], II, 15–17, one can prove

$$\neg \mathfrak{A} \rightarrow \neg A[\mathfrak{e}_1 \cdots \mathfrak{e}_n(b_1 \cdots b_{n-1})b_1 \cdots b_n] \tag{24.4}$$

by repeated application of the ϵ-formula, where

$$\mathfrak{e}_1 = \epsilon_{x_1}(y_1) \cdots (Ex_n)(y_n) \neg A(x_1 \cdots x_n y_1 \cdots y_n)$$

$$\mathfrak{e}_{r+1}(a_1 \cdots a_r) = \epsilon_{x_{r+1}}(y_{r+1}) \cdots (Ex_n)(y_n) \neg A[\mathfrak{e}_1 \mathfrak{e}_2(a_1) \cdots \mathfrak{e}_r(a_1 \cdots a_{r-1})x_{r+1} \cdots x_n a_1 \cdots a_r y_{r+1} \cdots y_n]$$

and the b are free variables. Note, again by [2], II, 15–17, the quantifiers in the explicit definitions of the $\mathfrak{e}$ can be eliminated by ϵ-expressions, but the above notation is simpler.

Similarly one can prove by the ϵ-formula

$$\mathfrak{A} \rightarrow A[a_1 \cdots a_n \mathfrak{t}_1(a_1) \cdots \mathfrak{t}_n(a_1 \cdots a_n)] \tag{24.5}$$

where $\mathfrak{t}_1(a_1) = \epsilon_{y_1}(x_2)(Ey_2) \cdots (x_n)(Ey_n)A(a_1 x_2 \cdots x_n y_1 \cdots y_n)$

$$\mathfrak{t}_r(a_1 \cdots a_r) = \epsilon_{y_r}(x_{r+1})(Ey_{r+1}) \cdots (x_n)(Ey_n) A[a_1 \cdots a_r x_{r+1} \cdots x_n \mathfrak{t}_1(a_1) \cdots \mathfrak{t}_{r-1}(a_1 \cdots a_{r-1})y_r \cdots y_n]$$

where the a are free variables.

(ii) Substitute in 24.4 the term $\bar{p}_i$ of the predicate calculus for b_i, where $\bar{p}_i$ is got by replacing the function symbols f in p of 24.1 by $\mathfrak{e}$. Hence we have proofs of the implications

$$A[\mathfrak{e}_1 \cdots \mathfrak{e}_n(\bar{p}_1^k \cdots \bar{p}_{n-1}^k)\bar{p}_1^k \cdots \bar{p}_n^k] \rightarrow \mathfrak{A}, \qquad 1 \leq k \leq l,$$

and therefore of

$$\{A[\mathfrak{e}_1 \cdots \mathfrak{e}_n(\bar{p}_1^1 \cdots \bar{p}_{n-1}^1)\bar{p}_1^1 \cdots \bar{p}_n^1] \vee \cdots \vee A[\mathfrak{e}_1 \cdots \mathfrak{e}_n(\bar{p}_1^l \cdots \bar{p}_{n-1}^l)\bar{p}_1^l \cdots \bar{p}_n^l]\} \rightarrow \mathfrak{A}. \quad 24.6$$

Similarly, from 24.5

$$\neg\{A[\bar{q}_1^1 \cdots \bar{q}_n^1\mathfrak{t}_1(\bar{q}_1^1) \cdots \mathfrak{t}_n(\bar{q}_1^1 \cdots \bar{q}_n^1)] \& \cdots \& A[\bar{q}_1^{l_1} \cdots \bar{q}_n^{l_1}\mathfrak{t}_1(\bar{q}_1^{l_1}) \cdots \mathfrak{t}_n(\bar{q}_1^{l_1} \cdots \bar{q}_n^{l_1})]\} \rightarrow \neg\mathfrak{A}, \quad 24.7$$

where $\bar{q}$ is got from q in 24.2 by substituting $\mathfrak{t}$ for the function variables in q.

(iii) Consider then the two proofs by the ϵ-formula of the implications 24.6 and 24.7. Apply Hilbert's substitution method of arithmetic functions for the matrices (= Grundtyp, [2], II, 93–128) of the ϵ-terms occurring in these proofs. By this method all the formulae of these proofs will be reduced to true formulae. But since 24.1 is supposed to be verifiable, whatever functions are substituted for the $\mathfrak{e}_r(b_1 \cdots b_{r-1})$ in 24.6 the premiss will be true, and hence in the final substitution $\mathfrak{A}$ will be replaced by a true formula. Therefore a false one is substituted for $\neg\mathfrak{A}$ in 24.7, and, since the whole formula 24.7 is reduced to a true one, the premiss of 24.7 is reduced to a false one. In other words, the functions $\mathfrak{y}_r(a_1 \cdots a_r)$ which are substituted by the Hilbert method for $\mathfrak{t}_r(a_1 \cdots a_r)$ make 24.2 true.

The converse is immediate: we add to a given proof of $\neg\mathfrak{A}$ the proof, got from 24.4, of the formulae

$$\neg\mathfrak{A} \rightarrow \neg A[\mathfrak{e}_1 \cdots \mathfrak{e}_n(\bar{p}_1^k \cdots \bar{p}_{n-1}^k)\bar{p}_1^k \cdots \bar{p}_n^k], \qquad 1 \leq k \leq l.$$

This sequence of formulae constitutes a proof of the formulae

$$\neg A[\mathfrak{e}_1 \cdots \mathfrak{e}_n(\bar{p}_1^k \cdots \bar{p}_{n-1}^k)\bar{p}_1^k \cdots \bar{p}_n^k].$$

Apply the Hilbert substitution method to this proof. The number $\mathfrak{x}_1$, and the functions $\mathfrak{x}_{r+1}(b_1 \cdots b_r)$ which replace $\mathfrak{e}_1$, $\mathfrak{e}_{r+1}(b_1 \cdots b_r)$ in the final substitution establish the converse.

Note that the functions $\mathfrak{x}$ and $\mathfrak{y}$ differ from 0 for a finite set of arguments only.

Remark. Ackermann's attempt [6] to apply the substitution method to a formalism for analysis led to difficulties with functionals, which were pointed out by von Neumann [7]. It should be observed that the trouble arose when *bound* function variables were used which are hardly mentioned in the present paper. This is why von Neumann's criticism is consistent with the lemma above.

25. Using the lemma we can set up the following interpretation of extensions (c) of the predicate calculus. *Interpretation.* Given a formula $\mathfrak{A}$, 2.3, $n_i = m_i = 1$, we consider the terms of $\mathfrak{F}\mathfrak{N}$ which contain the free individual variable a, the function variables $f_2(a_1), \cdots, f_n(a_1 \cdots a_{n-1})$, and finite sets of function symbols which have been added to the predicate calculus in (c); we order finite sequences of n-tuples of such functionals linearly, calling the r^{th} sequence $(p_{r1}^1 \cdots p_{rn}^1), \cdots, (p_{r1}^{m_r} \cdots p_{rn}^{m_r})$.

Then we associate with $\mathfrak{A}$ the disjunction of the formulae

$$A[a_1 f_2(p_{r1}^k) \cdots f_n(p_{r1}^k \cdots p_{rn-1}^k)p_{r1}^k \cdots p_{rn}^k], \qquad 1 \leq k \leq m_r. \quad 25.1$$

This satisfies the conditions of para. 20.

(α) The sequences 25.1 have a recursive numbering by the numbering of terms of the predicate calculus.

(β) If $\mathfrak{A}$ has been proved from the free variable axiom $A_0(a_1 \cdots a_q)$ by Herbrand's theorem we can find an $\mathfrak{n}$ so that the disjunction of $A[af_2(p_{\mathfrak{n}1}^k) \cdots f_n(p_{\mathfrak{n}1}^k \cdots p_{\mathfrak{n}n-1}^k)p_{\mathfrak{n}1}^k \cdots p_{\mathfrak{n}n}^k] \vee \neg A_0(\alpha_1^k \cdots \alpha_q^k)$, $1 \leq k \leq m_{\mathfrak{n}}$ is verifiable. Since $A_0(a_1 \cdots a_q)$ is verifiable, 25.1 must itself be verifiable (as in para. 21).

(γ) Since formulae without variables are associated to themselves, by the remark at the end of para. 21 (γ) follows from (δ). It also follows immediately from the lemma.

(δ) Suppose $\mathfrak{B}$ is proved from $\mathfrak{A}$. Bind the free (individual) variables in $\mathfrak{A}$ and replace them by all-variables; denote the resulting formula also by $\mathfrak{A}$.

By the Deduktionstheorem, $\mathfrak{A} \rightarrow \mathfrak{B}$ can be proved. Let the prenex form of $\mathfrak{B}$ be

$$(u_1)(Ev_1) \cdots (u_r)(Ev_r)B(u_1 \cdots u_r v_1 \cdots v_r)$$

so that a prenex form of $\mathfrak{A} \rightarrow \mathfrak{B}$ is

$$(u_1)(Ev_1) \cdots (u_r)(Ev_r)(Ex_1)(y_1) \cdots (Ex_n)(y_n)$$
$$[B(u_1 \cdots u_r v_1 \cdots v_r) \vee \neg A(x_1 \cdots x_n y_1 \cdots y_n)].$$

Now, if $\mathfrak{A} \rightarrow \mathfrak{B}$ can be proved, so can its prenex form, and now by (β) we find terms p_i^k, q_j^k, $1 \leq i \leq r$, $1 \leq j \leq n$, $1 \leq k \leq l$ of the predicate calculus so that

$$B[af_2(p_1^1) \cdots f_r(p_1^1 \cdots p_{r-1}^1)p_1^1 \cdots p_r^1]$$
$$\vee \neg A[q_1^1 \cdots q_n^1 g_1(q_1^1) \cdots g_n(q_1^1 \cdots q_n^1)] \vee \cdots \vee$$
$$B[af_2(p_1^l) \cdots f_r(p_1^l \cdots p_{r-1}^l)p_1^l \cdots p_r^l]$$
$$\vee \neg A[q_1^l \cdots q_n^l g_1(q_1^l) \cdots g_n(q_1^l \cdots q_n^l)]$$

is verifiable. The only variables of p, q are the individual variable a, and the function variables $f_i(a_1 \cdots a_{i-1})$, $1 < i \leq r$, $g_j(a_1 \cdots a_j)$, $1 \leq j \leq n$.

If a disjunction A_n associated with $\mathfrak{A}$ is verifiable, by the lemma we can find computable functions $\mathfrak{y}_i(a_1 \cdots a_i)$ so that the conjunction of

$$A[\tilde{q}_1^k \cdots \tilde{q}_n^k \mathfrak{y}_1(\tilde{q}_1^k) \cdots \mathfrak{y}_n(\tilde{q}_1^k \cdots \tilde{q}_n^k)], \qquad 1 \leq k \leq l,$$

is verifiable, where the parameters of $\mathfrak{y}$ are the individual variable a and the function variables f.

Now substitute these functions for g in p_i^k to give $\tilde{p}_i^k$. Then the disjunction of

$$B[af_2(\tilde{p}_1^k) \cdots f_r(\tilde{p}_1^k \cdots \tilde{p}_{r-1}^k)\tilde{p}_1^k \cdots \tilde{p}_r^k], \qquad 1 \leq k \leq l,$$

must be verifiable. The terms got will still be terms of the predicate calculus since only a fixed number of values of $\mathfrak{y}$ are needed in $\tilde{p}_i$.

26. *Erfüllung interpretation.* Given a proof in the predicate calculus of

$$(x_1) \cdots (x_p)B(x_1 \cdots x_p) \rightarrow (x_1) \cdots (x_n)(Ey_1) \cdots (Ey_m)A(x_1 \cdots x_n y_1 \cdots y_m) \qquad 26.1$$

we get terms $p_i^k(a_1 \cdots a_n)$, $1 \leq i \leq m$, $q_j^k(a_1 \cdots a_n)$, $1 \leq j \leq p$, $1 \leq k \leq l$, so that the disjunction of

$$A(a_1 \cdots a_n p_1^k \cdots p_m^k) \vee \neg B(q_1^k \cdots q_p^k), \qquad 1 \leq k \leq l,$$

holds, i.e.

$$[B(q_1^1 \cdots q_p^1) \& \cdots B(q_1^l \cdots q_p^l)] \rightarrow [A(a_1 \cdots a_n p_1^1 \cdots p_m^1) \vee \cdots A(a_1 \cdots a_n p_1^l \cdots p_m^l)] \qquad 26.2$$

and hence we have an Erfüllung provided the premiss is true.

In other words, if the formula $(Ey)A(\alpha, y)$ can be proved from $(x)B(x)$ it can also be proved if $B(n)$ is true for a certain *finite* number of n, depending on α.

27. *Examples.* We interpret the formula

$$(x)(Ey)(z)[A(x, y) \vee \neg A(x, z)]$$

since this contains already the difficulty of non-finitist proofs. We denote the formula by $\mathfrak{A}$.

(a) Take the *first repeated negation*: the prenex form of $\neg \mathfrak{A}$ is

$$(Ex)(y)(Ez)[\neg A(x, y) \& A(x, z)].$$

An Erfüllung of $\neg \mathfrak{A}$ would be a term α and a function $f_0(b)$ so that

$$\neg A(\alpha, b) \& A[\alpha, f_0(b)] \qquad 27.1$$

is verifiable, b being the free variable. This is impossible since either $b = \alpha$ or $b = f_0(\alpha)$ is a counter example.

Conversely, as an example of 24.4, we can prove $\mathfrak{A}$ from

$$A(\alpha, \alpha) \vee \neg A[\alpha, f_0(\alpha)] \vee A[\alpha, f_0(\alpha)] \vee \neg A\{\alpha, f_0[f_0(\alpha)]\}$$

by substituting a suitable ϵ-term for f_0, and applying the step 24(i).

(b) Next consider the *second repeated negation*: an Erfüllung of $\mathfrak{A}$ is a function $f_0(a)$ so that

$$A[a, f_0(a)] \vee \neg A(a, b)$$

is verifiable. A negation II would be terms $q_x[f(x)]$, $p_x[f(x)]$ of the predicate calculus so that

$$\neg A\{q_x[f_0(x)], f_0\{q_x[f_0(x)]\}\} \& A\{q_x[f_0(x)], p_x[f_0(x)]\} \qquad 27.2$$

is true.

Take for $q_x[f(x)]$ the functional $f(\alpha)$, for $p_x[f(x)]$ the functional $f_2\{f\{f_1[f(\alpha)]\}\}$.

We have to find an $f_0(a)$ to make 27.2 false, where f_1, f_2 are taken to be symbols for computable functions.

(i) Take $f_0^{(1)}(\alpha) = f_0^{(1)}[f_1(\alpha)] = \alpha$, otherwise arbitrary.

If 27.2 is true when $f_0^{(1)}$ is substituted for f, $A(\alpha, \alpha)$ must be false, $A[\alpha, f_2(\alpha)]$ true, and therefore $f_2(\alpha) \neq \alpha$.

(ii) Define $f_0^{(2)}(\alpha) = f_2(\alpha)$, $f_0^{(2)}[f_2(\alpha)] = f_2[f_2(\alpha)]$, and if $f_1[f_2(\alpha)] \neq f_2(\alpha)$ take $f_0^{(2)}\{f_1[f_2(\alpha)]\} = f_2(\alpha)$.

If 27.2 is true when $f_0^{(2)}$ is substituted for f_0, $A\{f_2(\alpha), f_2[f_2(\alpha)]\}$ must be false, $A\{f_2(\alpha), f_2\{f_2[f_2(\alpha)]\}\}$ true.

(iii) Define $f_0^{(3)}(\alpha) = f_2(\alpha)$, $f_0^{(3)}[f_2(\alpha)] = f_2\{f_2[f_2(\alpha)]\}$, otherwise arbitrary.

Then one of the three functions $f_0^{(1)}, f_0^{(2)}, f_0^{(3)}$ makes 27.2 false. For, if $f_0^{(1)}$ does not, $A(\alpha, \alpha)$ must be false, $A[\alpha, f_2(\alpha)]$ true, and $f_2(\alpha) \neq \alpha$. Then the definitions of $f_0^{(2)}$ and $f_0^{(3)}$ allow these functions to be one valued (are consistent). If $f_0^{(3)}$ is substituted in 27.2, we get

$$\neg A\{f_2(\alpha), f_0^{(3)}[f_2(\alpha)]\} \;\&\; A[f_2(\alpha), f_2\{f_0^{(3)}\{f_1[f_2(\alpha)]\}\}]$$

i.e.

$$\neg A\{f_2(\alpha), f_2\{f_2[f_2(\alpha)]\}\} \;\&\; A[f_2(\alpha), f_2\{f_0^{(3)}\{f_1[f_2(\alpha)]\}\}]. \qquad 27.3$$

But, if $f_0^{(2)}$ makes 27.2 true, $\neg A\{f_2(\alpha), f_2\{f_2[f_2(\alpha)]\}\}$ is false, and therefore 27.3 is also false.

Note that a restriction on the class of functionals is necessary.

IV. *The Hilbert substitution method and interpretation.*

28. The work of the last section can be generalized: under quite wide conditions the no-counter-example interpretation can be given to extensions (c) of a formalism to which the Hilbert substitution method, [2], II, 93–128, has been applied. The general principles are stated in the present section, and the conditions needed are checked in the next one for extensions (c) of number theory Z.

29. Recall the substitution method: given a system (a) containing the predicate calculus with the ϵ-symbol, verifiable free variable axioms, and rules of inference in the form of ϵ-formulae, as with induction. Consider a proof in which the variables have been eliminated by ϵ-expressions so that it consists of (i) applications of the ϵ-formulae (these are the *critical* ϵ-formulae of the proof), (ii) substitutions in identical formulae of the propositional calculus or axioms, (iii) formulae obtained from preceding ones by the syllogism.

By the *Hilbert substitution method* we mean a technique of converting the critical ϵ-formulae of any such proof (= any finite set of critical ϵ-formulae) into true formulae without variables by substituting computable functions for the matrices of the ϵ-terms occurring in the proof. Thereby the whole of the proof can be reduced to a sequence of true formulae without variables, [2], II, bottom of p. 95.

Note that, by p. 128 of [2], if the substitution method can be applied to a system (a), it can also be applied to its extensions (b) and (c).

Remark 1. The proof is not necessarily turned into an elementary *proof* without variables since the (true) formulae which replace the critical formulae of the proof are generally not proved by their predecessors.

Remark 2. ϵ-terms are not generally replaced by their *values* and ϵ-matrices by the functions which they represent, even if they are computable numbers or functions.

Consider the single critical formula $3 > 5 \rightarrow \epsilon_x(x > 5) > 5$. By the substitution method of [3], e.g., $\epsilon_x(x > a)$ is replaced at the first stage of the elimination by $f(a)$, $f(a) = 0$ for all a, and $\epsilon_x(x > 5)$ by 0 in this, the first and final, step of the elimination. 0 is *not* a value of $\epsilon_x(x > 5)$, $\epsilon_x(x > 5) \neq 0$ can be proved.

We consider now an ϵ-term $\mathfrak{e}(\mathfrak{e}_1)$ when $\mathfrak{e}(a)$ represents a computable function $e(a)$ in one of the senses below: *we wish to ensure that when* $\mathfrak{e}_1$ *gets the substitution* e_1, $\mathfrak{e}(\mathfrak{e}_1)$ *gets the substitution* $e(e_1)$.

(i) Suppose there are *free variable axioms* (e.g. recursive relations) satisfied by the computable function $e(a)$ so that from these axioms one can prove *in* the formalism

$$(x)[\mathfrak{e}(x) = e(x)]$$

and hence

$$\mathfrak{e}(\mathfrak{e}_1) = e(\mathfrak{e}_1). \tag{29.1}$$

Add the proof of 29.1 to the given proof in which the term $\mathfrak{e}(\mathfrak{e}_1)$ occurs, and apply the substitution method to the extended proof. If $\mathfrak{e}_1$ gets the number e_1, then $\mathfrak{e}(\mathfrak{e}_1)$ gets the number $e(e_1)$. This observation is needed repeatedly below.

Note that the ϵ-expressions subordinate (= untergeordnet, [2], II, 23) to $\mathfrak{e}(a)$ need not be computable.

(ii) Another method is useful if $\mathfrak{e}(a)$ and the matrices of *all* ϵ-expressions subordinate to $\mathfrak{e}(a)$ are computable in the following sense: the matrix of an ϵ-expression of rank 1, $\epsilon_x A(x; a_1 \cdots a_n)$ is called computable and represents the function $f(a_1 \cdots a_n)$ if $A[f(a_1 \cdots a_n); a_1 \cdots a_n]$ is *verifiable* and $n < f(a_1 \cdots a_n) \rightarrow \neg A(n; a_1 \cdots a_n)$. ($A[f(a_1 \cdots a_n); a_1 \cdots a_n]$ *need not be provable in the system considered.*)

In the matrix of an ϵ-expressions of rank $n + 1$, substitute the computable functions represented by the subordinate ϵ-expressions for the latter, so that the resulting expression is of rank 1, [2], II, 104. In the given proof replace $\mathfrak{e}(a)$ and the matrices of the ϵ-expressions subordinate to it by the computable functions which they represent. After this modification, critical ϵ-formulae *not* belonging to ϵ-terms with matrices subordinate to $\mathfrak{e}(a)$ remain critical ϵ-formulae. Critical ϵ-formulae belonging to such ϵ-terms will be true whatever numbers are substituted for the ϵ-terms contained (= eingelagert [2], II, 23) in them.

This method is useful because any computable function can be represented in the above sense by terms of Z_μ.

This method will not be needed in the present paper, but is important in other applications (cf. footnote 5).

Remark 3. When it is said that the technique of converting critical formulae into true ones must apply to any proof in any extension (c) of the given formalism, we must have a recursive *functional*

$$\omega \ldots y_i \ldots [\cdots f_j(\cdots y_i \cdots) \cdots ; \cdots p_r \cdots] \tag{29.2}$$

which is a *bound* for the numbers that are substituted for the ϵ-terms of given proofs: the f_j are the function symbols occurring in the proof, p_r constants of the proof (e.g. the number of terms occurring in the proof or its Gödel number). In the general case of non-numerical formalisms we get a bound for the terms without variables to be substituted for the ϵ-terms.[8]

With recursive functionals, for given constants of the proof, the values of the functional will depend on the values of f_j for a *finite* number of arguments only.

[8] In the case of a system (b), without function variables (para. 18), the functional 29.2 reduces to a computable function, τ say (cf. [3]), and can be used for enumerating those computable functions which can be proved *in* the system to be computable; that is, the formula $(x)(Ey)(Ez)A(x, y, z)$ can be proved *in* the system when the computable function is $\mu_y(Ez)A(x, y, z)$ (any computable function can be thus expressed with primitive recur-

Construction of the interpretation. 30. To a formula $\mathfrak{A}$, 2.3 with $n_i = m_i = 1$ for simplicity, we associate formulae A_r

$$A[af_2(\mathfrak{y}_1^{(r)}) \cdots f_n(\mathfrak{y}_1^{(r)} \cdots \mathfrak{y}_{n-1}^{(r)})\mathfrak{y}_1^{(r)} \cdots \mathfrak{y}_n^{(r)}] \qquad 30.1$$

where $(\mathfrak{y}_1^{(r)} \cdots \mathfrak{y}_n^{(r)})$ is defined to be the first n-tuple of numbers, $\mathfrak{y}_i^{(r)}$ less than the r^{th} functional of form 29.2, for which 30.1 is true, and $\mathfrak{y}_i^{(r)} = 0$ if there is no such n-tuple. The definition is effective when computable functions are substituted for the function variables in $\mathfrak{y}_i^{(r)}$.

The sequence A_n satisfies (α).

Also it satisfies (β) because if $\mathfrak{A}$ can be proved so can

$$(Ey_1) \cdots (Ey_n)A[af_2(y_1) \cdots f_n(y_1 \cdots y_{n-1})y_1 \cdots y_n]$$

for any a and f, that is, in ϵ-notation, there are ϵ-terms $\mathfrak{e}_1 \cdots \mathfrak{e}_n$ so that

$$A[af_2(\mathfrak{e}_1) \cdots f_n(\mathfrak{e}_1 \cdots \mathfrak{e}_n)\mathfrak{e}_1 \cdots \mathfrak{e}_n]$$

can be proved in an extension (c). Hence numbers $\mathfrak{y}_1^{(r)} \cdots \mathfrak{y}_n^{(r)}$ are got by the substitution method, bounded by the relevant functional of 29.2, which make 30.1 true.

31. To check (γ) we have to show: if $\neg\mathfrak{A}$ has been proved, to any set of functionals $\mathfrak{y}_1 \cdots \mathfrak{y}_n$ of paragraph 30 containing the function symbols f, we can find computable functions $g_2(a_1), \cdots, g_n(a_1 \cdots a_{n-1})$ and a number $\mathfrak{x}$ so that

$$\neg A[\mathfrak{x}g_2(\tilde{\mathfrak{y}}_1) \cdots g_n(\tilde{\mathfrak{y}}_1 \cdots \tilde{\mathfrak{y}}_{n-1})\tilde{\mathfrak{y}}_1 \cdots \tilde{\mathfrak{y}}_n] \qquad 31.1$$

holds, where $\tilde{\mathfrak{y}}$ is got from $\mathfrak{y}$ by substituting the functions g for f.

The method of section III is not directly applicable because of the rather special character of functionals of the predicate calculus;[9] but the idea is the

---

sive A). From the enumeration we can *find* to each such function an argument $\mathfrak{n}$ where its value differs from $\tau(\mathfrak{n})$.

This result is rather stronger than the one derived directly from Gödel's second undecidability theorem (for a suitable system, say Z): this implies that there can be no term with variable p of the system for which it can be *proved in* the system that it is a bound for the numbers that have to be substituted for the ϵ-terms of a proof with number p to reduce the proof to a sequence of true formulae. But this leaves the possibility that for some function which is computable in the system it is merely undecided in the system whether it furnishes such a bound. And with Z (cf. [3]) this is just the case with a certain computable term (but not computable *in* Z), namely the μ-symbol definition of the function τ of [3].

[9] Note that the class of functionals that may be added to a formalism to which the ϵ-substitution method is to be applied, must be restricted, that is, not all functionals which are defined for every computable function may be added. This is in sharp contrast to the addition of symbols for computable functions. We consider functionals suggested by appendix I.

Let us fix on some unique system of definitions $\gamma_n(a)$ (terms of Z_μ) for functions $g_n(a)$ which differ from zero for a finite number of arguments only; let $\mathfrak{x}_n$ be the number of the formula $\neg$ $Prov[\gamma_n(b), s(b, b)]$ where b is a free variable. Then $\mathfrak{x}_n$ is uniquely defined by $g_n(a)$. Let the functional $\mathfrak{x}_x[g(x)]$ be defined for our functions by the rule $\mathfrak{x}_x[g_n(x)] = \mathfrak{x}_n$.

For other functions $g(a)$, $\mathfrak{x}_x[g(x)] = \mathfrak{z}_x[g(x)] = 0$, also $\mathfrak{z}_x[g_n(x)]$ has the value of μ_y $Prov[y, s(\mathfrak{x}_n, \mathfrak{x}_n)]$.

Now, consider the critical ϵ-formula $Prov[\mathfrak{z}, s(\mathfrak{x}, \mathfrak{x})] \rightarrow Prov\{\epsilon_y\, Prov[y, s(\mathfrak{x}, \mathfrak{x})], s(\mathfrak{x}, \mathfrak{x})\}$

same: we suppose that the functionals $\mathfrak{y}_1 \cdots \mathfrak{y}_n$ can be represented by terms of the formalism when the f are terms of the formalism (i.e. they can be proved in the formalism to satisfy the same recursive relations as $\mathfrak{y}$); we consider then the terms $\mathfrak{e}_1, \mathfrak{e}_2(a), \cdots, \mathfrak{e}_n(a_1 \cdots a_{n-1})$ of the formalism,

$$\mathfrak{e}_1 = \mu_{x_1}(y_1) \cdots (Ex_n)(y_n) \neg A(x_1 \cdots x_n y_1 \cdots y_n)$$

$$\mathfrak{e}_{r+1}(a_1 \cdots a_r) = \mu_{x_{r+1}}(y_{r+1}) \cdots (Ex_n)(y_n)$$

$$\neg A[\mathfrak{e}_1 \cdots \mathfrak{e}_r(a_1 \cdots a_{r-1}) x_{r+1} \cdots x_n a_1 \cdots a_r y_{r+1} \cdots y_n]$$

and we denote by $\mathfrak{Y}$ the term representing $\mathfrak{y}$ when the $\mathfrak{e}$ are substituted for f. Then, from a proof of $\neg \mathfrak{A}$ we get a proof of

$$\neg A[\mathfrak{e}_1 \cdots \mathfrak{e}_n(\mathfrak{Y}_1 \cdots \mathfrak{Y}_{n-1})\mathfrak{Y}_1 \cdots \mathfrak{Y}_n]. \qquad 31.2$$

If we could find substitutions for the ϵ-matrices of this proof so that, when $\mathfrak{e}_r(a_1 \cdots a_{r-1})$ is replaced by the computable function $e_r(a_1 \cdots a_{r-1})$, $\mathfrak{Y}_j$ is replaced by the value of $\mathfrak{y}_j$ for the arguments e in place of f, we should have checked (γ). When the $\mathfrak{y}$ are functionals of the predicate calculus this is automatically achieved by applying the Hilbert substitution method to the proof of 31.2. According to *remark 2*, in general, special artifices have to be employed even if one deals only with *functions*; here we deal with functionals. Below, we explain *how to replace the recursive functionals considered by recursive functions with ϵ-terms as parameters.*

32. We consider number theory and start with an outline of the argument. *We assume*:

(i) Given functionals $\mathfrak{y}_1 \cdots \mathfrak{y}_n$ represented in the formalism, so that $\mathfrak{Y}$ represents the functional $\mathfrak{y}$ with arguments $\mathfrak{e}$, $\mathfrak{y}^*$ the functional $\mathfrak{y}$ with arguments

---

where $Prov(m, n)$ is the 'proof' predicate of $Z\mu$, $\mathfrak{x}$ the term $\mathfrak{x}_x\{\epsilon_y\ Prov[y, s(x, x]\}$ (supposed to occur in the system considered), and $\mathfrak{z}$ is the term $\mathfrak{z}_x\{\epsilon_y\ Prov[y, s(x, x)]\}$.

We show that the analogue to the usual substitution method does not terminate. Substitute for the ϵ-matrix $\epsilon_x Prov\ [x, s(a, a)]$ the functions $g_i(a)$ according to the rule: $g_0(a) = 0$ for all a; $g_1(a) = 0$ unless $a = \mathfrak{x}_0$, and $g_1(\mathfrak{x}_0) = \mathfrak{z}_0$; generally $g_{i+1}(a) = 0$ unless $a = \mathfrak{x}_0, \cdots \mathfrak{x}_i$; $g_{i+1}(\mathfrak{x}_j) = \mathfrak{z}_j, j \leq i$. The substitutions do not come to an end and $\mathfrak{x}_r \neq \mathfrak{x}_s$ for $r \neq s$.

Note in passing that the definitions of these functionals $\mathfrak{x}$ and $\mathfrak{z}$ can be so extended to computable functions that they are negations *II* of the proved formula of the predicate calculus

$$(x)(Ey)(z)\{Prov[y, s(x, x)] \vee \neg\ Prov[z, s(x, x)]\}$$

(cf. appendix I). Let $\mathfrak{t}(n)$ be a term of Z_μ which represents the computable function $f(n)$ in the sense that for any numeral $\mathfrak{n}$, $\mathfrak{t}(\mathfrak{n}) = \mathfrak{m}$ is proved in Z_μ when $f(\mathfrak{n}) = \mathfrak{m}$. Let $\mathfrak{x}_x[f(x)]$ be the number of one of the formulae $\neg\ Prov[\mathfrak{t}(b), s(b, b)]$, so that $\mathfrak{x}_x[f(x)]$ is many-valued depending on the representation of $f(n)$ considered; then, as shown in appendix I, for any such $\mathfrak{t}(n)$, $\neg\ Prov[\mathfrak{t}(\mathfrak{x}), s(\mathfrak{x}, \mathfrak{x})]$ is true and can be proved in Z_μ. Let $\mathfrak{z}_x[f(x)]$ be the number of its shortest proof $(\mu_y Prov[y, s(\mathfrak{x}, \mathfrak{x})])$. Then any such numbers $\mathfrak{x}$, $\mathfrak{z}$ are a counter-example to the proposed Erfüllung $f(a)$.

Note that if $f(x) \neq \mathfrak{g}(x)$ for only *one* x, terms $\mathfrak{t}(x)$ and $\mathfrak{g}(x)$ of Z_μ representing $f(x)$, $g(x)$ respectively, necessarily differ, and $\mathfrak{x}_x[f(x)] \neq \mathfrak{x}_x[g(x)]$. With the recursive functionals $\omega_x[f(x)]$ used in the text, if $f(x) = g(x)$ for a suitable *finite* numbers of arguments, then $\omega_x[f(x)] = \omega_x[g(x)]$.

$e_2(a) \cdots e_n(a_1 \cdots a_{n-1})$: then we can find terms $\mathfrak{n}_1 \cdots \mathfrak{n}_{n-1}$ and prove in the formalism

$$(x_1) \cdots (x_{n-1})[x_1 \leq \mathfrak{n}_1 \;\&\; \cdots\; x_{n-1} \leq \mathfrak{n}_{n-1} \cdot \rightarrow \cdot$$
$$\mathfrak{e}_2(x_1) = e_2(x_1) \;\&\; \cdots\; \mathfrak{e}_n(x_1 \cdots x_{n-1}) = e_n(x_1 \cdots x_{n-1})] \rightarrow$$
$$\mathfrak{H}_1 = \mathfrak{y}_1^* \;\&\; \cdots\; \mathfrak{H}_n = \mathfrak{y}_n^*. \qquad 32.1$$

In other words, if the terms e and $\mathfrak{e}$ are equal over a certain range of their variables, the values of the functionals are equal (*cf. remark 3* of paragraph 29). Also we assume that the $\mathfrak{n}$ can be defined from $\mathfrak{e}$ only.

(ii) When recursive functions are substituted for f in $\mathfrak{y}$, the resulting term is a recursive function; in particular, if $e_2(a) \cdots e_n(a_1 \cdots a_{n-1})$ are recursive so is $\mathfrak{y}^*$. The e may contain parameters.

We shall define below primitive recursive functions $e_r(a_1 \cdots a_{r-1} ; p)$ with parameters p so that to any terms $\mathfrak{n}_1 \cdots \mathfrak{n}_{n-1}$ in Z_μ one can find certain ϵ-terms $\mathfrak{p}_r$ so that

$$a_1 \leq \mathfrak{n}_1 \;\&\; \cdots\; a_{r-1} \leq \mathfrak{n}_{r-1} \cdot \rightarrow \cdot\; \mathfrak{e}_r(a_1 \cdots a_{r-1}) = e_r(a_1 \cdots a_{r-1} ; \mathfrak{p}_r) \qquad 32.2$$

can be proved in Z_μ. By (ii), $\mathfrak{y}^*$ in 32.1 is a recursive function with $\mathfrak{p}_r$ as parameters, and by *remark 2* of paragraph 29 we can find numbers p_r so that the functions $e_r(a_1 \cdots a_{r-1} ; p_r)$ are the $g_r(a_1 \cdots a_{r-1})$ required.

Remark. The problem could be put as follows: if we had at our disposal a formalism with *bound function* variables, from a proof of $\neg\mathfrak{A}$ we get

$$(Ex_1)(E\dot{x}_2) \cdots (E\dot{x}_n)(y_1) \cdots (y_n) \neg A[x_1 \dot{x}_2(y_1) \cdots \dot{x}_n(y_1 \cdots y_{n-1}) y_1 \cdots y_n]$$

and hence

$$(Ex_1)(E\dot{x}_2) \cdots (E\dot{x}_n) \neg A[x_1 \dot{x}_2(\mathfrak{H}_1) \cdots \dot{x}_n(\mathfrak{H}_1 \cdots \mathfrak{H}_{n-1}) \mathfrak{H}_1 \cdots \mathfrak{H}_n] \qquad 32.3$$

where $\mathfrak{H}$ is the representation of the functional $\mathfrak{y}$ with $\dot{x}$ as arguments. From a proof of 32.2 one might expect to get computable functions for $\dot{x}$. But we have no general method of doing this: the attempts to apply the Hilbert substitution method to formal systems with bound function variables have so far failed.

We define a sequence of (primitive) recursive functions $g_2(a; p) \cdots g_n(a_1 \cdots a_{n-1}; p)$ with p as parameter, and show that the $\dot{x}$ required are among this sequence; i.e. we prove in Z_μ

$$(Ex_1)(Ep) \neg A[x_1\, g_2(\tilde{\mathfrak{H}}_1^p ; p) \cdots g_n(\tilde{\mathfrak{H}}_1^p \cdots \tilde{\mathfrak{H}}_{n-1}^p) \tilde{\mathfrak{H}}_1^p \cdots \tilde{\mathfrak{H}}_n^p] \qquad 32.4$$

where $\tilde{\mathfrak{H}}^p$ is the recursive *function* got by substituting $g_r(a_1 \cdots a_{r-1} ; p)$ for $f_r(a_1 \cdots a_{r-1})$ in $\mathfrak{y}$.

32.4 is a formula of Z_μ, proved in Z_μ, to which the Hilbert substitution method can be applied to find $\mathfrak{x}$ and $\mathfrak{p}$.

33. Lemma. Let $e_1'(a_1^1 \cdots a_n^1 b_1^1 \cdots b_n^1)$ denote the primitive recursive term

$$\mu_{x_1}(y_1) \cdots (Ex_n)(y_n)[y_1 \leq b_1^1 \,\&\, \cdots y_n \leq b_n^1 \cdot \rightarrow \cdot x_1 \leq a_1^1 \,\&\, \cdots x_n \leq a_n^1 \,\&\, \neg A(x_1 \cdots x_n\, y_1 \cdots y_n)]$$

where a_i^1, b_i^1 are free variables. Let

$$e_{r+1}(y_1 \cdots y_r\,;\, c_1 \cdots c_r\,;\, a_{r+1}^{r+1} \cdots a_n^{r+1}\, b_{r+1}^{r+1} \cdots b_n^{r+1})$$

denote the term

$$\mu_{x_{r+1}}(y_{r+1}) \cdots (Ex_n)(y_n)[y_{r+1} \leq b_{r+1}^{r+1} \,\&\, \cdots y_n \leq b_n^{r+1} \cdot \rightarrow \cdot x_{r+1} \leq a_{r+1}^{r+1} \,\&\, \cdots x_n \leq a_n^{r+1} \,\&\, \neg A(c_1 \cdots c_r\, x_{r+1} \cdots x_n\, y_1 \cdots y_n)]$$

where the index is not a power, but a second name for the variable. Let

$$e'_{r+1}(y_1 \cdots y_r\,;\, \cdots a_i^j \cdots\,;\, \cdots b_i^j \cdots)$$
$$= e_{r+1}[y_1 \cdots y_r\,;\, e'_1(a_1^1 \cdots a_n^1\, b_1^1 \cdots b_n^1) \cdots e'_r(y_1 \cdots y_{r-1}\,;\, \cdots a_p^q \cdots\,;\, \cdots b_p^q \cdots)\; a_{r+1}^{r+1} \cdots a_n^{r+1} b_{r+1}^{r+1} \cdots b_n^{r+1}]$$

where

$$1 \leq j \leq r + 1, \quad j \leq i \leq n, \quad 1 \leq q \leq r, \quad q \leq p \leq n.$$

Then we show how to write down terms $\mathfrak{n}_1^1 \cdots \mathfrak{n}_n^1$, $\mathfrak{m}_1^1 \cdots \mathfrak{m}_n^1$ without variables, $\mathfrak{n}_i^{j+1}$, $\mathfrak{m}_i^{j+1}$ with free variable N so that the formulae

$$\mathfrak{e}_1 = e'_1(\mathfrak{n}_1^1 \cdots \mathfrak{n}_n^1\, \mathfrak{m}_1^1 \cdots \mathfrak{m}_n^1) \tag{33.1}$$

and

$$y_1 \leq N \,\&\, \cdots y_r \leq N \cdot \rightarrow \mathfrak{e}_{r+1}(y_1 \cdots y_r) = e'_{r+1}(y_1 \cdots y_r;\, \cdots \mathfrak{n}_i^j \cdots;\, \cdots \mathfrak{m}_i^j \cdots) \tag{33.2}$$

can be proved in Z_μ (this means that the $\mathfrak{n}$ and $\mathfrak{m}$ are terms of Z_μ).

Remark. The argument is metamathematical, explaining how the proof is to be carried out for given $\mathfrak{A}$.

(a) Consider the formula

$$(Ex)(y)\mathfrak{B}(a;\, x,\, y)$$

where a is a free variable, and $\mathfrak{B}$ may contain bound variables other than x and y. Denote

$$\mu_z(v)[v \leq N \rightarrow \mu_x(y)\mathfrak{B}(v;\, x,\, y) \leq z] \text{ by } \mathfrak{n} \text{ or } \mathfrak{n}(N), \tag{33.3}$$

$$\mu_z(v)(w)[v \leq N \,\&\, w \leq \mathfrak{n}. \rightarrow \mu_y\overline{\mathfrak{B}(v;\, w,\, y)} \leq z] \text{ by } \mathfrak{m}. \tag{33.4}$$

Remark. $\mathfrak{n}$ is the maximum of the least numbers x for which $(y)\mathfrak{B}(a;\, x,\, y)$ when $a \leq N$; and $\mathfrak{m}$ is the maximum of the least numbers y for which $\neg\mathfrak{B}(a;\, c,\, y)$ when $a \leq N$, $c \leq \mathfrak{n}$.

Then we show that the formulae

$$a \leq N \to \mu_x(y)\mathfrak{B}(a; x, y) = \mu_x(y)[y \leq \mathfrak{m} \to\cdot x \leq \mathfrak{n} \,\&\, \mathfrak{B}(a; x, y)] \qquad 33.5$$

$$a \leq N \to\cdot (Ex)(y)\mathfrak{B}(a; x, y) \sim (Ex)(y)[y \leq \mathfrak{m} \to\cdot x \leq \mathfrak{n} \,\&\, \mathfrak{B}(a; x, y)] \qquad 33.6$$

can be proved in Z_μ.

Remark. The idea of the proof is quite obvious. Let $a \leq \mathfrak{n}$, denote

$$\mu_x(y)\mathfrak{B}(a; x, y) \text{ by } \lambda(a), \qquad \mu_x(y)[y \leq \mathfrak{m} \to\cdot x \leq \mathfrak{n} \,\&\, \mathfrak{B}(a; x, y)]$$

by $\sigma(a)$, and $\mu_y \neg \mathfrak{B}(a; c, y)$ by $\rho(a, c)$. Take

(i) The case when $(Ex)(y)\mathfrak{B}(a; x, y)$ holds. Then also $(y)\mathfrak{B}[a; \lambda(a); y]$ holds, and there is certainly an $x \leq \mathfrak{n}$, the maximum of $\lambda(a)$, $a \leq N$, for which $(y)\mathfrak{B}(a; x, y)$ holds. Also, if $c < \lambda(a)$, there is a $y < \max \rho(a, c)$, $a \leq N$, $c \leq \lambda(a)$, so that $\neg\mathfrak{B}(a; c, y)$: thus $\mathfrak{B}(a; c, y)$ cannot hold for all y less than $\max \rho(a, c)$, $a \leq N$, $c \leq \lambda(a)$, and, *a fortiori*, for all $y \leq \mathfrak{m}$ unless it holds for all y. Therefore, if there is an x at all for which $(y)\mathfrak{B}(a; x, y)$ holds, 33.5 and 33.6 hold, too.

(ii) Consider an a, still $\leq N$, where $(x)(Ey)\overline{\mathfrak{B}(a; x, y)}$. Then, for any x, $\neg\mathfrak{B}[a; x, \rho(a, x)]$, and, for $x \leq \mathfrak{n}$ $\rho(a, x) \leq \mathfrak{m}$ since $\mathfrak{m}$ is the maximum of $\rho(a, x)$ for $x \leq \mathfrak{n}$ *and all* $a \leq N$. Thus, there can be no $x \leq \mathfrak{n}$ so that for all $y \leq \mathfrak{m}$ $\mathfrak{B}(a; x, y)$ holds. Again 33.5 and 33.6 hold.

We give more details because, in applications, one has to apply the Hilbert substitution method to proofs of formulae of the form 33.5 and 33.6 so that it is essential that these proofs can be written out in a system for which the method has been justified (e.g. Z_μ). The reader may skip the formal details of part (a) of the arguments, and continue with part (b).

Description of the proofs of 33.5 and 33.6. By μ_1, [2] *I*, 396,

$$(Ex)(y)\mathfrak{B}(a; x, y) \to (y)\mathfrak{B}[a; \lambda(a), y].$$

By 33.3 $a \leq N \to \lambda(a) \leq \mathfrak{n}$ so that

$$a \leq N \,\&\, (Ex)(y)\mathfrak{B}(a; x, y)\cdot\!\to (Ex)(y)[x \leq \mathfrak{n} \,\&\, \mathfrak{B}(a; x, y)].$$

Since $A \to\cdot B \to A$, taking $x \leq \mathfrak{n} \,\&\, \mathfrak{B}(a; x, y)$ for A, $y \leq \mathfrak{m}$ for B

$$a \leq N \,\&\, (Ex)(y)\mathfrak{B}(a; x, y)\cdot \to (Ex)(y)[y \leq \mathfrak{m} \to\cdot x \leq \mathfrak{n} \,\&\, \mathfrak{B}(a; x, y)], \qquad 33.7$$

one half of 33.6.

Also, by μ_2, [2], I, 396,

$$(y)\{y \leq \mathfrak{m} \to\cdot \lambda(a) \leq \mathfrak{n} \,\&\, \mathfrak{B}[a; \lambda(a), y]\} \to \sigma(a) \leq \lambda(a),$$

and $\quad a \leq N \,\&\, (Ex)(y)\mathfrak{B}(a; x, y)\cdot\!\to\cdot \lambda(a) \leq \mathfrak{n} \,\&$

$$(y)\{y \leq \mathfrak{m} \to \mathfrak{B}[a; \lambda(a), y]\} \quad \text{so that}$$

$$a \leq N \,\&\, (Ex)(y)\mathfrak{B}(a; x, y) \to \sigma(a) \leq \lambda(a). \qquad 33.8$$

Since, by μ_2, $(y)\mathfrak{B}(a; c, y) \to \lambda(a) \leq c$, $c < \lambda(a) \to (Ey)\overline{\mathfrak{B}(a; c, y)}$. Also, by μ_1,

$(Ey)\overline{\mathfrak{B}(a;\, c,\, y)} \rightarrow \neg\mathfrak{B}[a;\, c,\, \rho(a,\, c)]$. By 33.3 $a \leq N \rightarrow \lambda(a) \leq \mathfrak{n}$, and hence $a \leq N \;\&\; c < \lambda(a) \cdot\!\rightarrow c \leq \mathfrak{n}$; also, by 33.4, $a \leq N \;\&\; c \leq \mathfrak{n} \cdot\!\rightarrow \rho(a,\, c) \leq \mathfrak{m}$, so that

$$c < \lambda(a) \cdot\!\rightarrow\!\cdot\, a \leq N \rightarrow (Ey)[y \leq \mathfrak{m} \;\&\; \neg\, \mathfrak{B}(a;\, c,\, y)].$$

Since $A \rightarrow\!\cdot\, A \vee B$, taking $\neg\mathfrak{B}(a;\, c,\, y)$ for A, $\mathfrak{n} < c$ for B

$$c < \lambda(a) \rightarrow \{a \leq N \rightarrow (Ey)[y \leq \mathfrak{m} \;\&\!\cdot \neg\, \mathfrak{B}(a;\, c,\, y) \vee \mathfrak{n} < c]\}.$$

Substituting $\sigma(a)$ for c

$$\sigma(a) < \lambda(a) \rightarrow \{a \leq N \rightarrow (Ey)\{y \leq \mathfrak{m} \;\&\!\cdot \neg\, \mathfrak{B}[a;\, \sigma(a),\, y] \vee \mathfrak{n} < \sigma(a)\}\}.$$

By 33.7 and μ_1

$$a \leq N \;\&\; (Ex)(y)\mathfrak{B}(a;\, x,\, y) \cdot \rightarrow (y)\{y \leq \mathfrak{m} \rightarrow\!\cdot\, \sigma(a) \leq \mathfrak{n} \;\&\; \mathfrak{B}[a;\, \sigma(a),\, y]\}$$

so that

$$a \leq N \;\&\; (Ex)(y)\mathfrak{B}(a;\, x,\, y) \cdot\!\rightarrow \neg\, \sigma(a) < \lambda(a).$$

By the trichotomy of pairs of integers and 33.8

$$a \leq N \;\&\; (Ex)(y)\mathfrak{B}(a;\, x,\, y) \cdot\!\rightarrow \sigma(a) = \lambda(a).$$

This proves one half of 33.5.

(ii) By μ_1, [2], I, 396, $(x)(Ey)\overline{\mathfrak{B}(a;\, c,\, y)} \rightarrow \overline{\mathfrak{B}[a;\, c,\, \rho(a,\, c)]}$.

By 33.4 $a \leq N \;\&\; c \leq \mathfrak{n} \cdot\!\rightarrow \rho(a,\, c) \leq \mathfrak{m}$ so that

$$a \leq N \;\&\; (x)(Ey)\overline{\mathfrak{B}(a;\, x,\, y)} \cdot \rightarrow (Ey)\{y \leq \mathfrak{m} \;\&\; [c \leq \mathfrak{n} \rightarrow \neg\mathfrak{B}(a;\, c,\, y)]\}.$$

Since c does not occur in the premiss

$$a \leq N \;\&\; (x)(Ey)\overline{\mathfrak{B}(a;\, x,\, y)} \cdot \rightarrow (x)(Ey)\{y \leq \mathfrak{m} \;\&\; [x \leq \mathfrak{n} \rightarrow \neg\mathfrak{B}(a;\, x,\, y)]\}.$$

But $[A \;\&\; (B \rightarrow \neg\, C)] \sim [A \;\&\; (\neg B \vee \neg C)] \sim \neg[A \rightarrow (B \;\&\; C)]$; taking $y \leq \mathfrak{m}$ for A, $x \leq \mathfrak{n}$ for B, $\mathfrak{B}(a;\, x,\, y)$ for C

$$a \leq N \;\&\; (Ex)(y)\overline{\mathfrak{B}(a;\, x,\, y)} \cdot \rightarrow (x)(Ey)\{\neg[y \leq \mathfrak{m} \rightarrow\!\cdot\, x \leq \mathfrak{n} \;\&\; \mathfrak{B}(a;\, x,\, y)]\}.$$

Hence, together with μ_3, [2], I, 396

$$a \leq N \;\&\; (Ex)(y)\overline{\mathfrak{B}(a;\, x,\, y)} \cdot\!\rightarrow \sigma(a) = 0 \;\&\; \lambda(a) = 0, \qquad \text{i.e. } \sigma(a) = \lambda(a).$$

This establishes the second halves of 33.5 and 33.6.

(*b*) To define the terms $\mathfrak{n}_1^1$, $\mathfrak{m}_1^1$, the argument of (*a*) is applied to the formula

$$(Ex_2)(y_2) \cdots (Ex_n)(y_n) \neg\, A(x_1 \cdots x_n y_1 \cdots y_n)$$

in place of $\mathfrak{B}(a;\, x,\, y)$ without free variable a; $\mathfrak{n}$ and $\mathfrak{m}$ of 33.3 and 33.4 being denoted by $\mathfrak{n}_1^1$, $\mathfrak{m}_1^1$. Then 33.5 reduces to

$$\mu_{x_1}(y_1) \cdots (Ex_n)(y_n) \neg\, A(x_1 \cdots x_n y_1 \cdots y_n)$$
$$= \mu_{x_1}(y_1)\{y_1 \leq \mathfrak{m}_1^1 \rightarrow [x_1 \leq \mathfrak{n}_1^1 \;\&\; (Ex_2) \cdots (y_n) \neg\, A(x_1 \cdots x_n y_1 \cdots y_n)]\}.$$

Now, by [2], I, 397, to prove $\mu_x A(x) = \mu_x B(x)$ it is sufficient to prove $A(x) \sim B(x)$. Consider then the formula

$$(Ex_2)(y_2)\{y_1 \le \mathfrak{m}_1^1 \rightarrow [x_1 \le \mathfrak{n}_1^1 \;\&\; (Ex_3)(y_3) \cdots (Ex_n)(y_n) \neg A(x_1 \cdots x_n y_1 \cdots y_n)]\} \qquad 33.9$$

with free variables x_1 and y_1, and apply 33.3, 33.4 to determine $\mathfrak{n}$ and $\mathfrak{m}$ *when* $x_1 \le \mathfrak{n}_1^1$, *and* $y_1 \le \mathfrak{m}_1^1$, denoting these terms by $\mathfrak{n}_2^1$, $\mathfrak{m}_2^1$. Then the equivalence

$$33.9 \quad \sim (Ex_2)(y_2)\{(y_1 \le \mathfrak{m}_1^1 \;\&\; y_2 \le \mathfrak{m}_2^1) \rightarrow [x_1 \le \mathfrak{n}_1^1 \;\&\; x_2 \le \mathfrak{n}_2^1 \;\&\; (Ex_3)(y_3) \cdots (Ex_n)(y_n) \neg A(x_1 \cdots x_n y_1 \cdots y_n)]\}$$

can be proved in Z_μ as follows:

(i) if $y_1 > \mathfrak{m}_1^1$ both formulae are true;

(ii) if $y_1 \le \mathfrak{m}_1^1$ and $x_1 > \mathfrak{n}_1^1$, 33.9 is false, and so is the second formula, e.g. if we take $y_2 = 0$;

(iii) But if $x_1 \le \mathfrak{n}_1^1$, $y_1 \le \mathfrak{m}_1^1$, the equivalence reduces to 33.6.

Repeating the argument, bounds $\mathfrak{n}_i^1$, $\mathfrak{m}_i^1$ are found for all the variables.

(c) Recall that $\mathfrak{e}_{r+1}(y_1 \cdots y_r)$ was defined to be the term

$$\mu_{x_{r+1}}(y_{r+1}) \cdots (Ex_n)(y_n) \neg A[\mathfrak{e}_1 \cdots \mathfrak{e}_r(y_1 \cdots y_{r-1})x_{r+1} \cdots x_n y_1 \cdots y_n].$$

Suppose that the terms $\mathfrak{n}_i^j$, $\mathfrak{m}_i^j$, $j \le r$, have already been found so that one can prove in Z_μ, $s \le r$

$$(y_1 \le N \;\&\; \cdots y_{r-1} \le N) \rightarrow \mathfrak{e}_s(y_1 \cdots y_{s-1}) = e_s'(y_1 \cdots y_{s-1}; \cdots \mathfrak{n}_i^j \cdots; \cdots \mathfrak{m}_i^j \cdots).$$

Now consider the term

$$\mu_{x_{r+1}}(y_{r+1}) \cdots (Ex_n)(y_n) \neg A(c_1 \cdots c_r x_{r+1} \cdots x_n y_1 \cdots y_n),\; = \lambda_1,$$

say, with free variables $c_1 \cdots c_r$, $y_1 \cdots y_r$. By (*b*), we find terms $\mathfrak{n}_{r+1} \cdots \mathfrak{n}_n$, $\mathfrak{m}_{r+1} \cdots \mathfrak{m}_n$, containing the free variables $c_1 \cdots c_r$, N so that one can prove in Z_μ

$$(y_1 \le N \;\&\; \cdots y_r \le N) \rightarrow \lambda_1 = \mu_{x_{r+1}}(y_{r+1}) \cdots (Ex_n)(y_n)\{(y_{r+1} \le \mathfrak{m}_{r+1} \;\&\; \cdots y_n \le \mathfrak{m}_n) \rightarrow [x_{r+1} \le \mathfrak{n}_{r+1} \;\&\; \cdots x_n \le \mathfrak{n}_n \;\&\; \neg A(x_1 \cdots x_n y_1 \cdots y_n)]\}.$$

Now substitute the terms

$$e_1'(\cdots \mathfrak{n}_i^1 \cdots; \cdots \mathfrak{m}_i^1 \cdots), \cdots e_r'(y_1 \cdots y_{r-1}; \cdots \mathfrak{n}_i^j \cdots; \cdots \mathfrak{m}_i^j \cdots)$$

for $c_1 \cdots c_r$ in $\mathfrak{n}$ and $\mathfrak{m}$ and denote the resulting terms by $\mathfrak{n}_{r+1}^{r+1}, \cdots, \mathfrak{n}_n^{r+1}$, $\mathfrak{m}_{r+1}^{r+1}, \cdots, \mathfrak{m}_n^{r+1}$.

Remark. In the applications below we shall need the result only when

$$(Ex_1)(y_1) \cdots (Ex_n)(y_n) \neg A(x_1 \cdots x_n y_1 \cdots y_n)$$

has been proved.

REFERENCES

[1] G. Kreisel, *On the concepts of completeness and interpretation of formal systems* (forthcoming).

[2] D. Hilbert and P. Bernays, ***Grundlagen der Mathematik***, vol. I (1934) and vol. II (1939), Berlin.

[3] W. Ackermann, *Zur Widerspruchsfreiheit der reinen Zahlentheorie*, ***Mathematische Annalen***, vol. 117 (1940), pp. 162–194.

[4] A. E. Ingham, ***The distribution of prime numbers***, Cambridge, 1932.

[5] J. E. Littlewood, *Large numbers*, ***Mathematical gazette***, vol. 32 (1948), pp. 163–171.

[6] W. Ackermann, *Begründung des "tertium non datur" mittels der Hilbertschen Theorie der Widerspruchsfreiheit*, ***Mathematische Annalen***, vol. 93 (1925), pp. 1–36.

[7] J. von Neumann, *Zur Hilbertschen Beweistheorie*, ***Mathematische Zeitschrift***, vol. 26 (1927), pp. 1–46.

[8] R. Péter, *Über die mehrfache Rekursion*, ***Mathematische Annalen***, vol. 113 (1936), pp. 489–527.

[9] E. C. Titchmarsh, ***The zeta-function of Riemann***, Cambridge, 1930.

[10] G. Gentzen, *Die Widerspruchsfreiheit der reinen Zahlentheorie*, ***Mathematische Annalen***, vol. 112 (1936), pp. 493–565.

[11] K. Gödel, *Zur intuitionistischen Arithmetik und Zahlentheorie*, ***Ergebnisse eines mathematischen Kolloquiums***, Heft 4 (1933), pp. 34–38.

[12] G. Gentzen, *Neue Fassung des Widerspruchsfreiheitsbeweises für die reine Zahlentheorie*, ***Forschungen zur Logik und zur Grundlegung der exakten Wissenschaften***, Neue Folge, Heft 4 (1938), pp. 19–44.

THE UNIVERSITY, READING, ENGLAND

THE JOURNAL OF SYMBOLIC LOGIC
Volume 17, Number 1, March 1952

ON THE INTERPRETATION OF NON-FINITIST PROOFS

PART II. INTERPRETATION OF NUMBER THEORY. APPLICATIONS.

G. KREISEL

V. 34. Since the Hilbert substitution method has been applied by Ackermann to the formalism Z of number theory in [3], we shall give the no-counter-example interpretation for extensions (c) of Z, e.g. Z_μ. The work required on top of Ackermann's investigations to check conditions (α) and (β), consists only in a suitable definition of a class of recursive functionals. The verification of (γ) and (δ) is much more difficult, and requires the ideas of paras. 32 and 33 in the preceding section.

Ordinal recursive functions of finite order. 35. We call the primitive recursive functions of [2], *I*, 287, of *order zero*. We take for granted that the function $\omega(p, m, n)$ is of order zero with respect to its three variables, where $\omega(p, m, n) = 0$ if and only if $m \underset{p}{<} n$ [3], p. 179. The proof of this is only indicated by Ackermann, but, as he says, is easily written out in full.

Def. 1: Primitive recursive functions of finite order.

(a) Functions of order zero have already been defined.

(b) $\phi(n)$ is of order $p+1$, either if it is of order p, or

$$\phi(0) = \mathfrak{a}, \qquad \phi(n) = \psi\{n, \phi[\mathfrak{t}^*(n)]\},$$

where $\mathfrak{a}$, $\psi(n, m)$, $\mathfrak{t}(n)$ have already been introduced, $\mathfrak{t}^*(n) = \mathfrak{t}(n)\omega^*[p+1, \mathfrak{t}(n), n]$, and

$$\omega^*(p+1, m, n) = 1 \quad \text{if} \quad \omega(p+1, m, n) = 0,$$

0 otherwise.

Observe that this scheme is a finite scheme in the sense of para. 17.

Def. 2: k-tuple recursive functions of finite order.

(a) Any multiple recursive function of order p is k-tuple of order $p+1$.

(b) $\phi(n_1 \cdots n_k)$ is k-tuple of order $p+1$ if
$\phi(0 \cdots 0) = \mathfrak{a}$ and
$\phi(n_1 \cdots n_k) = \psi\{n_1 \cdots n_k \cdots \phi[\phi_1^j(n_1 \cdots n_k) \cdots \phi_k^j(n_1 \cdots n_k)] \cdots\}$
$\phi_i^j(n_1 \cdots n_k) = n_i$ for $i < i_j$, $\phi_{i_j}^j(n_1 \cdots n_k) \underset{p+1}{<} n_{i_j}$,
(lexicographic ordering), and $\mathfrak{a}$, ψ, ϕ_i^j are k-tuple of order $p+1$.

Such functions are used in the definition of counter-examples.

Remark. k-tuple recursive functions of order p can be defined by primitive recursive definitions of order $p+2$, so that we may restrict ourselves to primitive recursive ones of finite order without 'losing' those of def. 2.

The principle of the proof is this: we map k-tuples in lexicographic order $\underset{p}{<}$ on

Received December 9, 1949. Appendix II was added in April 1951.

Part I appeared in this JOURNAL, vol. 16 (1951), pp. 241–267. References are on page 267 of Part I.

a section $v(p + 1; k)$ of the integers in order $\underset{p+1}{<}$: the mapping is 0-primitive recursive. We define a primitive recursive function $\Phi(n)$ for $n \underset{p+2}{<} 2^{v(p+1,k)}$ so that

$$\Phi(2^{c_1} + \cdots 2^{c_r} - 1) = 2^{\phi(c_1)} \cdots \mathfrak{p}_r^{\phi(c_r)}$$

where $c_r \underset{p+1}{<} c_{r-1} \cdots \underset{p+1}{<} c_1 \underset{p+1}{<} v(p + 1; k)$, $\mathfrak{p}_r$ is the r^{th} prime number, c is the number of the k-tuple $(n_1 \cdots n_k)$ in our mapping, and $\phi(c)$ denotes the value of $\phi(n_1 \cdots n_k)$. $\Phi(n)$ can actually be defined by a $p + 2$ *-primitive* recursive definition, and then the values of $\phi(n_1 \cdots n_k)$ are got from $\Phi(n)$ by means of the inverse mapping which is also 0-primitive recursive.

36. We must now show that these functions are computable *in* Z_μ, actually also in Z, in sense (1) of remark 2 to para. 29. This result is needed for the following applications:

(i) for introducing explicit definitions for terms representing them in extensions (c) of Z; that is, if $\mathfrak{e}(n)$ is an ϵ-term of Theorem I representing the ordinal recursive function $f(n)$ of finite order, by adding the recursive relations for $f(n)$ to the formalism Z we can prove *in* Z

$$(x)[f(x) = \mathfrak{e}(x)]$$

which is used in verifying (γ) and (δ);

(ii) it follows from footnote 8 that the general definition of the counter-example cannot be recursive of finite order;

(iii) Theorem II will show that any function computable in Z_μ is ordinal recursive of finite order so that we have a simple characterization of functions computable in Z_μ. Also it will follow that our class of functionals is the smallest class which is needed for the interpretation of extensions of Z.

Theorem I. *If $\phi(n)$ is ordinal recursive of order p, we find a term $\mathfrak{e}(n)$ of Z so that $\mathfrak{e}(n)$ can be proved in Z to satisfy the recursive relations by which $\phi(n)$ is defined.*

Suppose

$$\phi(n) = \psi\{n, \phi[\mathfrak{t}^*(n)]\}, \tag{36.1}$$

and suppose terms ψ_1, $\mathfrak{t}_1^*(n)$ of Z have been found for ψ and $\mathfrak{t}^*$. Adopt some numbering of quasi-recursive definitions, e.g. that of [2], II, 407–410. We take $\mathfrak{l}$ to be the number of 36.1 together with the (quasi-recursive) definitions of ψ and $\mathfrak{t}^*$. Using the notation of [2], II, 410, $\mathfrak{h}(\mathfrak{l}, n, r) = 0$ if and only if $r = 2^k 3^m$, $k = \nu(r, 0)$, $m = \nu(r, 1)$, and m is the number of a proof from the definition with number $\mathfrak{l}$ of the formula $\phi(n) = k$. We shall show that

$$\phi_1(n) = \nu\{\mu_x[\mathfrak{h}(\mathfrak{l}, n, x) = 0], 0\} \tag{36.2}$$

can be proved in Z to satisfy the relation

$$\phi_1(n) = \psi_1\{n, \phi_1[\mathfrak{t}_1^*(n)]\}.$$

Note that since ν and $\mathfrak{h}$ are primitive recursive, by [2], I, 410 the term $\phi_1(n)$ can be represented by a term of Z.

The crucial part of the proof is to establish in Z

$$(n)(Er)(v)[\mathfrak{h}(\mathfrak{l}, n, r) = 0 \,\&\cdot \mathfrak{h}(\mathfrak{l}, n, v) = 0 \rightarrow \nu(v, 0) = \nu(r, 0)]. \qquad 36.3$$

Assume that the relevant results for ψ_1, and $\mathfrak{t}_\mathfrak{l}^*(n)$ have been established. We call a proof from the definition with number $\mathfrak{l}$ (by repeated substitution) an $\mathfrak{l}$-proof.

(i) $\mathfrak{h}(\mathfrak{l}, n, s) = 0 \rightarrow (Er)\{\mathfrak{h}[\mathfrak{l}, \mathfrak{t}_\mathfrak{l}^*(n), r] = 0 \,\&\, \nu(s, 0) = \psi_1[n, \nu(r, 0)]\}$ 36.4

$$\mathfrak{h}[\mathfrak{l}, \mathfrak{t}_\mathfrak{l}^*(n), r] = 0 \rightarrow (Es)\{\mathfrak{h}(\mathfrak{l}, n, s) = 0 \,\&\, \nu(s, 0) = \psi_1[n, \nu(r, 0)]\}. \qquad 36.5$$

(i.e. if $\phi(n) = \nu(s, 0)$ can be proved, then also $\phi[\mathfrak{t}_\mathfrak{l}^*(n)] = \nu(r, 0)$, and $\phi(n) = \nu(s, 0) = \psi_1\{n, \phi[\mathfrak{t}_\mathfrak{l}^*(n)]\}$, and conversely). From the definition of $\mathfrak{h}(\mathfrak{l}, n, r)$, if $\nu(s, 1)$ is the number of an $\mathfrak{l}$-proof $\mathfrak{P}$ of $\phi(n) = \nu(s, 0)$, $\mathfrak{P}$ must contain an $\mathfrak{l}$-proof of $\phi[\mathfrak{t}_\mathfrak{l}^*(n)] = k$, and of $\psi_1(n, k) = \nu(s, 0)$, hence 36.4. Conversely, if we have already an $\mathfrak{l}$-proof $\mathfrak{P}_1$ of $\phi[\mathfrak{t}_\mathfrak{l}^*(n)] = \nu(r, 0)$, by the assumption that 36.3 holds for ψ_1, there is an $\mathfrak{l}$-proof $\mathfrak{P}_2$ of $\psi[n, \nu(r, 0)] = k_1$ for a unique integer k_1. Writing $\mathfrak{P}_2$ after $\mathfrak{P}_1$, we get an $\mathfrak{l}$-proof of $\phi(n) = k_1$, hence 36.5.

(ii) $\mathfrak{t}_\mathfrak{l}^*(n) \underset{p}{<} n$. Trivial.

(iii) To prove 36.3 we throw 36.4 and 36.5 into a form suitable for induction of order p.

By 36.5, there is an s so that $\mathfrak{h}(\mathfrak{l}, n, s) = 0$ provided $\mathfrak{h}[\mathfrak{l}, \mathfrak{t}_\mathfrak{l}^*(n), r] = 0$ for some r; all the more if for *all* $m \underset{p}{<} n$ $(Er)[\mathfrak{h}(\mathfrak{l}, m, r) = 0]$, by (ii).

By 36.4, $\nu(s, 0)$ is uniquely determined by $\mathfrak{h}(\mathfrak{l}, n, s) = 0$ provided $\nu(r, 0)$ is uniquely determined by $\mathfrak{h}[\mathfrak{l}, \mathfrak{t}_\mathfrak{l}^*(n), r] = 0$; again, all the more, if for *all* $m \underset{p}{<} n$, $\nu(r, 0)$ is uniquely determined by $\mathfrak{h}(\mathfrak{l}, m, r) = 0$. Hence

$$(m)\{m \underset{p}{<} n \rightarrow (Er)(v)[\mathfrak{h}(\mathfrak{l}, m, r) = 0 \,\&\cdot \mathfrak{h}(\mathfrak{l}, m, v) = 0$$
$$\rightarrow \nu(v, 0) = \nu(r, 0)]\} \rightarrow (Er)(v)[\mathfrak{h}(\mathfrak{l}, n, r) = 0$$
$$\&\cdot \mathfrak{h}(\mathfrak{l}, n, v) = 0 \rightarrow \nu(v, 0) = \nu(r, 0)]. \qquad 36.6$$

Applying ordinal induction of order p to 36.6

$$(n)(Er)(v)[\mathfrak{h}(\mathfrak{l}, n, r) = 0 \,\&\cdot \mathfrak{h}(\mathfrak{l}, n, v) = 0 \rightarrow \nu(v, 0) = \nu(r, 0)]. \qquad 36.7$$

By [2], II, 366, for any given p, 36.7 can be proved from 36.6 by repeated application of *ordinary* induction, hence in Z.

Observe, from uniqueness in 36.7

$$\mathfrak{h}(\mathfrak{l}, n, r) = 0 \rightarrow \nu(r, 0) = \nu\{\mu_x[\mathfrak{h}(\mathfrak{l}, n, x) = 0], 0\}. \qquad 36.8$$

(iv) To prove the recursive relations, substitute $\mu_x[\mathfrak{h}(\mathfrak{l}, n, x) = 0]$ for s in 36.4. Since from 36.7 follows $(n)(Er)[\mathfrak{h}(\mathfrak{l}, n, r) = 0]$, the premiss reduces to a provable formula, and hence

$$(Er)\{\mathfrak{h}[\mathfrak{l}, \mathfrak{t}_\mathfrak{l}^*(n), r] = 0 \,\&\, \nu\{\mu_x[\mathfrak{h}(\mathfrak{l}, n, x) = 0], 0\} = \psi_1[n, \nu(r, 0)]\}.$$

Using 36.2, and 36.8 with $t_1^*(n)$ instead of n,

$$\phi_1(n) = \psi_1\{n, \phi_1[t_1^*(n)]\},$$

as required.

Remark. The principle is to use a numbering of $\mathfrak{l}$-proofs, and establish some elementary properties of them, in our case 36.4 and 36.5, essentially free variable formulae. Then induction is applied to these results. The idea is due to von Neumann, and was used by Péter in [8] to represent interlocked (*verschränkt*) recursive functions by terms of Z.

Concerning the formalization of this type of proof see the third footnote on p. 49 of [2], II.

Ordinal recursive functionals. 37. Observe first that the function $\delta(a, b)$, $=$ $a - b$ for $b \leq a$, 0 for $a \leq b$, is 0-primitive recursive, [2], I, 303, and of a given set of terms $t_1 \cdots t_n$, the *maximum* is defined as follows: if M_i denotes the maximum of $t_1 \cdots t_i$, $M_{i+1} = M_i + \delta(t_{i+1}, M_i)$. Further the maximum value ϕ of functions $f_1(a) \cdots f_m(a)$ for $a \leq n$, and a term t, satisfies the recursive relations

$$\phi_x[f_1(x) \cdots f_m(x)\ t, 0] = \max [f_1(0) \cdots f_m(0)\ t]$$

$$\phi_x[f_1(x) \cdots f_m(x)\ t, n + 1]$$
$$= \max \{f_1(n + 1) \cdots f_m(n + 1)\, \phi_x[f_1(x) \cdots f_m(x)\ t, n]\}.$$

If t and f are terms of Z these recursive relations can be proved for the term

$$\mu_x(y)[y \leq n \rightarrow \cdot t \leq x \,\&\, f_1(y) \leq x \cdots \&\, f_m(y) \leq x].$$

The maximum of the n^{th} iterations of functions $f_1(a) \cdots f_m(a)$, and t, satisfies:

$$\omega_x[f_1(x) \cdots f_m(x)\ t, 0] = \phi_x[f_1(x) \cdots f_m(x)\ t, 0]$$

$$\omega_x[f_1(x) \cdots f_m(x)\ t, n + 1] = \phi_x\{f_1(x) \cdots f_m(x)\ t, \omega_x[f_1(x) \cdots f_m(x)\ t, n]\}.$$

By Gödel's argument, [2], I 414, ω, also, can be represented by a term of Z when the f and t are terms of Z. Both ϕ and ω are increasing functions of n.

Remark: By the usual mapping of n-tuples we can reduce our functionals to functionals with a *single* function variable which simplifies the definitions below: Define

$$F(x) = 2^{f_1(x)} \cdots \mathfrak{p}_m^{f_m(x)}$$

$\phi_x[F(x), t; 0]$ = maximum of exponents of $F(0)$ and t, see [2], I, 319,
$\phi_x[F(x), t; n + 1]$ = maximum of exponents of $F(n + 1)$, and of $\phi_x[F(x), t; n]$.
Similarly
$\omega_x[F(x), t; 0]$ = $\phi_x[F(x), t; 0]$
$\omega_x[F(x), t; n + 1]$ = maximum of the exponents of $F(t)$ for $t \leq \omega_x[F(x), t; n]$.

By the use of [2], II, 168, we can replace a function of r variables by one of one variable.

Def: *An ordinal recursive functional of finite order, containing the free variable a and the function variable $g(b)$, is either*

(a) an ordinal recursive function with the variable a,

(b) a term $g(\mathfrak{x})$ where $\mathfrak{x}$ is an ordinal recursive functional containing a and g,

(c) a term $f(a, \mathfrak{x})$ where $f(a, b)$ is an ordinal recursive function of finite order, or

(d) $\omega_x[g(x), \mathfrak{x}_1 ; \mathfrak{x}_2]$ where ω is the iteration functional defined above, and $\mathfrak{x}_1$, $\mathfrak{x}_2$ are ordinal recursive functionals containing a and g.

LEMMA. *Let* $\mathrm{t}(c)$ *be a term of* Z_μ *with the single free variable* c, *and let an ordinal recursive functional of finite order with the variables* a *and* g, *be represented by the term* $\mathfrak{x}(a)$, *when* t *is substituted for* g. *Then we find a term* $\mathfrak{N}(a)$ *of* Z_μ *so that for any term* $\mathrm{t}_1(c)$ *of* Z_μ

$$(x)[x \leq \mathfrak{N}(a) \rightarrow \mathrm{t}(x) = \mathrm{t}_1(x)] \rightarrow \mathfrak{x}(a) = \mathfrak{x}_1(a)$$

can be proved in Z_μ.

Note that $\mathfrak{N}(a)$ *is independent of* t_1 : *e.g.* $\mathrm{t}_1(c)$ *may contain free variables other than* c *and these will not occur in* $\mathfrak{N}$. *Both* t *and* t_1 *may contain bound variables.*

The result is proved by induction. For expressions (a) in the definition it is trivial. For (b), the functional is $\mathrm{t}(\mathfrak{x}_0)$, say $\mathfrak{x}$, so that for $\mathfrak{N}$ we may take the maximum of $\mathfrak{N}_0(a)$ and $\mathfrak{x}_0$. For (c), if $f(a, \mathfrak{x}_0)$ is again denoted by $\mathfrak{x}$, $\mathfrak{N}(a) = \mathfrak{N}_0(a)$. Lastly for (d), suppose the terms $\mathfrak{N}_0$ and $\mathfrak{N}_0'$ have already been found, and $\omega_x[\mathrm{t}(x)\mathfrak{x}_0, \mathfrak{x}_0']$ is represented by $\mathfrak{x}$. Then, for $\mathfrak{N}(a)$ take max $[\mathfrak{x}, \mathfrak{N}_0, \mathfrak{N}_0']$.

Remark. This is the analogue of the theorem at the end of para. 23, and justifies the first condition of para. 32 for ordinal recursive functionals of finite order.

Note that when recursive terms are substituted for the f in ϕ and ω the resulting term is a 0-primitive recursive function. By the definition of the functionals, it follows that any ordinal recursive functional of finite order reduces to an ordinal recursive function of finite order, if a function of this class is substituted for the function variable. Note that the function may contain parameters. This shows that the second condition of para. 32 also applies to our functionals.

A free variable formalism. 38. With regard to a development of arithmetic in a free variable formalism we mention for reference the system F_1 which consists of the elementary calculus of number theory with free variables, free function variables, the schema for primitive recursive definitions of finite order, the recursive relations for ϕ and ω, and finally the schema for ordinal induction of finite order:
if $\mathrm{t}(n)$ is a term of F_1, $A(0)$ and $A(n) \vee \neg A[\mathrm{t}_p^*(n)]$ have been proved in F_1, then $A(a)$, where $\mathrm{t}_p^*(n) = \mathrm{t}(n)\omega^*[p, \mathrm{t}(n), n]$.

Note that by our formulation we avoid bound variables in the induction schema.—(The formulation is quite obvious; actually in the proof of 36.7 the introduction of the bound variable m was artificial, and only used to bring the induction into the form discussed in [2].)

Construction of counter-examples. 39. By a straightforward application of the remarks of para. 29 to [3], we show first that a no-counter-example interpretation can be given to extensions (*c*) of Z which satisfies conditions (α) and (β). The only point to be checked is that ordinal recursive functionals provide the functionals required.

Note in passing that the concepts of *rank* (= Rang, [2], II, 25) and *matrix* (= Grundtyp, [2], II, 57) of ϵ-expressions have been defined not only for terms of Z but for all terms of the predicate calculus including function symbols.

THEOREM II. *If the formula*

$$(x_1)(Ey_1) \cdots (x_n)(Ey_n)A(x_1 \cdots x_n y_1 \cdots y_n) \qquad 39.1$$

has been proved in an extension (c) *of* Z, *we find an ordinal recursive functional* q *of finite order so that*

$$A[af_2(q_1) \cdots f_n(q_1 \cdots q_{n-1})q_1 \cdots q_n]$$

is verifiable (*for any* $a, f_2 \cdots f_n$), $q_i = \eta_i^{(n)}(q)$ *and* $\eta_1^{(n)}(m) \cdots \eta_n^{(n)}(m)$ *is the* m^{th} *n-tuple of integers in some* (*primitive recursive*) *ordering of n-tuples.*

Recall that any formula 2.3 can be reduced to the form 39.1 by the use of $\eta_i^{(j)}$.

LEMMA. *Given a proof in an extension* (c) *of* Z *of a formula*

$$(Ex_1) \cdots (Ex_n)A(x_1 \cdots x_n)$$

where $a_1 \cdots a_n$ *are the only variables of* $A(a_1 \cdots a_n)$, *we can find numbers* $\mathfrak{n}_1 \cdots \mathfrak{n}_n$ *so that* $A(\mathfrak{n}_1 \cdots \mathfrak{n}_n)$ *is true.* Also, if

(i) $\phi(n)$ is a function so that when numbers $\leq n$ are substituted for the ϵ-terms of the given proof, the terms occurring in the resulting sequence of numerical formulae $\leq \phi(n)$,

(ii) $$\omega(0) = \phi(0), \qquad \omega(n+1) = \phi[\omega(n)],$$

(iii) $\lambda(a, b)$, $\rho(a, b)$ are as defined in [3], 182, 194,

(iv) $\mathfrak{e}$ is the number of ϵ-terms, $\mathfrak{g}$ the number of critical matrices of the given proof,

(v) $\tau^*(c, \mathfrak{p}, \mathfrak{n}, \mathfrak{a})$, for fixed $\mathfrak{p}$, is defined together with auxiliary functions $\theta(\mathfrak{a})$, $\nu(\mathfrak{a})$, $\chi^*(c, \mathfrak{p}, \mathfrak{n}, \mathfrak{a})$ as follows:

$$a = 2^{\nu(a)}[2\theta(a) + 1] - 1$$

$$\tau^*(c, 0, \mathfrak{n}, \mathfrak{a}) = \tau^*(c, \mathfrak{p}, \mathfrak{n}, 0) = \chi^*(c, \mathfrak{p}, \mathfrak{n}, 0) = \chi^*(c, 0, \mathfrak{n}, \mathfrak{a}) = 0,$$

$$\theta(\mathfrak{a}) \neq 0 : \chi^*(c, 1, \mathfrak{n}, \mathfrak{a}) = 2^{\nu(\mathfrak{a})}[2\theta(\mathfrak{a}) - 1] - 1$$

$$\theta(\mathfrak{a}) = 0, \qquad \nu(\mathfrak{a}) \neq 0 : \chi^*(c, 1, \mathfrak{n}, \mathfrak{a}) = 2^{\nu(\mathfrak{a})-1}(2c+1) - 1$$

$$\begin{aligned}\chi^*(c, \mathfrak{p}, \mathfrak{n}, \mathfrak{a}) &= \mathfrak{a} - 1 \quad \text{if} \quad \mathfrak{p} > 1, \quad \mathfrak{a} \text{ even}\\ &= \tau^*(c, \mathfrak{p} - 1, \mathfrak{n}, \mathfrak{a}) \quad \text{if} \quad \mathfrak{a} = 2^{\mathfrak{a}_1} - 1\\ &= 2^{\mathfrak{a}_1} + \chi^*\{c, \mathfrak{p}, \mathfrak{n} + \lambda(\mathfrak{a}, \mathfrak{p}), \chi^*[c, \mathfrak{p}, \mathfrak{n} + \lambda(\mathfrak{a}, \mathfrak{p}), \mathfrak{a}]\}\\ &\text{if } \mathfrak{a} = 2^{\mathfrak{a}_1} + \cdots 2^{\mathfrak{a}_r} - 1 \qquad \mathfrak{a}_{i+1} \underset{\mathfrak{p}-1}{<} \mathfrak{a}_i\end{aligned}$$

$$\tau^*(c, \mathfrak{p}, \mathfrak{n}, \mathfrak{a}) = 2^{\mathfrak{a}} + \tau^*\{c, \mathfrak{p}, \mathfrak{n} + \lambda(\mathfrak{a}, \mathfrak{p}), \chi^*[c, \mathfrak{p}, \mathfrak{n} + \lambda(\mathfrak{a}, \mathfrak{p}), \mathfrak{a}]\},$$

then

$$\mathfrak{n}_i \leq \omega\{\lambda\{\tau^*[2^{\mathfrak{e}[\omega(1)+1]}, \mathfrak{g}, 1, \rho(\mathfrak{g}, \mathfrak{n}), \mathfrak{g} + 1\}\}. \qquad 39.2$$

The definitions (v) are those of [3], 182 bottom, if one uses instead of Ackermann's function $\omega(\mathfrak{m}, \mathfrak{n})$ our $\omega(n)$.

The difference between ϵ-terms of Z and those of extensions (c) of Z is the occurrence of function symbols; more precisely of symbols other then the four symbols for the successor, predecessor, sum, and product, which occur in Z. Since the *rank* of terms of extensions (c) is defined, Ackermann's rule for forming successive substitutions, p. 171, and the definition of indices of $\mathfrak{p}$-sequences on p. 176 apply. So does the proof of the crucial theorems I–V where it is only required that the function symbols should be computable.

Further the proofs of theorems 4–17 remain correct if throughout $\tau(\mathfrak{m}, \mathfrak{e}, \mathfrak{p}, \mathfrak{n}, \mathfrak{a})$ and $\chi(\mathfrak{m}, \mathfrak{e}, \mathfrak{p}, \mathfrak{n}, \mathfrak{a})$ are replaced by $\tau^*[\psi(\mathfrak{m}, \mathfrak{n}, \mathfrak{e}), \mathfrak{p}, \mathfrak{n}, \mathfrak{a}]$ and $\chi^*[\psi(\mathfrak{m}, \mathfrak{n}, \mathfrak{e}), \mathfrak{p}, \mathfrak{n}, \mathfrak{a}]$, $\psi(\mathfrak{m}, \mathfrak{n}, \mathfrak{e})$ defined on p. 182, with similar modifications in the predicate $\mathfrak{J}$. The only fact about $\psi(\mathfrak{m}, \mathfrak{n}, \mathfrak{e})$ used in the proofs of Theorems VI and VIII is that $\psi(\mathfrak{m}, \mathfrak{n}, \mathfrak{e}) = 2^{\mathfrak{e}[\omega(\mathfrak{m},\mathfrak{n})+1]}$ where $\omega(\mathfrak{m}, \mathfrak{n})$ is a bound for the terms occurring in the proof after $\mathfrak{n}$ substitutions. So is our $\omega(\mathfrak{n})$.

This gives the bound $\lambda\{\tau^*[2^{\mathfrak{e}[\omega(1)+1]}, \mathfrak{g}, 1, \rho(\mathfrak{g}, \mathfrak{e})], \mathfrak{g} + 1\}$ for the number of steps in the substitution, and hence 39.2.

To prove the theorem it is sufficient to verify that $\omega(n)$ is a functional of finite order. Terms occurring in an extension (c) of Z fall into three groups: terms without variables, ϵ-terms, and function symbols whose arguments are ϵ-terms where we adopt the convention that $f[g(\mathfrak{t})]$ is written as $h(\mathfrak{t})$ if $f(a)$, $g(a)$ are symbols for computable functions, and $\mathfrak{t}$ is an ϵ-term.

Let $f_1(a) \cdots f_r(a)$ be the function symbols occurring in the proof, functions of *one* variable being used throughout. Denote the maximum of the terms without variables by $\mathfrak{k}$, then the $\phi(n)$ of the lemma is simply the functional

$$\phi_x[F(x), \mathfrak{k}, n] \qquad \text{where} \qquad F(a) = 2^{f_1(a)} \cdots \mathfrak{p}_r^{f_r(a)}$$

of para. 37. By the definition of ordinal recursive functionals of finite order, for any proof in an extension (c) of Z, $\omega(n)$ is seen to be one of them.

Example. Consider the theorem that a bounded, monotone increasing sequence of reals a_n converges. To simplify notation let them lie between 0 and 1, and let us use binary scale, in which rationals $n/2^m$ terminate.

$$\frac{a(n, m)}{2^m} \leq a_n < \frac{a(n, m) + 1}{2^m}.$$

Then convergence means:

$$(m)(En_0)(n)[n > n_0 \rightarrow a(n, m) = a(n_0, m)] \cdot$$
$$\vee\ (Er)(Es)[a(r + 1, s) < a(r, s) \vee a(r, 0) > 1 \vee a(r, 0) < 0].$$

A counter-example would be a number $\mathfrak{m}$, and a function $N(n_0)$ so that for all n_0

$$N(n_0) > n_0 \quad \text{and} \quad a[N(n_0), \mathfrak{m}] \neq a(n_0, \mathfrak{m}), \text{ also}$$
$$a(r, \mathfrak{m}) \leq a(r + 1, \mathfrak{m}) \quad \text{and} \quad 0 \leq a(r, 0) \leq 1. \qquad 39.3$$

This is impossible: take $n_0 = 0$, $n_1 (= N(n_0))$, $\cdots$, $n_{i+1}(= N(n_i))$; then if for all r, $a(r, \mathfrak{m}) \leq a(r + 1, \mathfrak{m})$,

$$a(0, \mathfrak{m}) + 1 \leq a(n_1, \mathfrak{m})$$
$$a(n_1, \mathfrak{m}) + 1 \leq a(n_2, \mathfrak{m})$$
$$\vdots$$
$$a(n_{2\mathfrak{m}}, \mathfrak{m}) + 1 \leq a(n_{2\mathfrak{m}+1}\ \mathfrak{m})$$

so that $a(n_{2\mathfrak{m}+1}, \mathfrak{m}) > 2^{\mathfrak{m}} + a(0, \mathfrak{m})$, hence $a(n_{2\mathfrak{m}+1}, 0) > 1$. In our notation, for some n_0, $0 \leq n_0 \leq \omega_x[N(x), 0, 2^{\mathfrak{m}} + 1]$, 39.3 breaks down.

40. Condition (γ) is checked in THEOREM III. *Suppose the formula* $\neg\mathfrak{A}$ *has been proved in an extension (c) of* Z, *and* q *is an ordinal recursive functional containing the function variable* $F(a) = 2^{f_2(a)} \cdots \mathfrak{p}_{n-1}^{f_n(a)}$. *Then we determine a number* $\mathfrak{x}$, *and functions* $g_2(a_1) \cdots g_n(a_1 \cdots a_{n-1})$, *zero except for a finite number of arguments, so that*

$$\neg A[\mathfrak{x} g_2(\tilde{q}_1) \cdots g_n(\tilde{q}_1 \cdots \tilde{q}_{n-1})\tilde{q}_1 \cdots \tilde{q}_n] \tag{40.1}$$

where $\tilde{q}$ *is got from* q *by substituting* $g_{i+1}[\eta_1^{(n)}(a) \cdots \eta_i^{(n)}(a)]$ *for* $f_{i+1}(a)$, *and* $\tilde{q}_i = \eta_i^{(n)}(\tilde{q})$.

(i) To the proof of $\neg\mathfrak{A}$ add the proof of the implication $\neg\mathfrak{A} \rightarrow \neg A[\mathfrak{e}_1\mathfrak{e}_2(a_1) \cdots \mathfrak{e}_n(a_1 \cdots a_{n-1})a_1 \cdots a_n]$ where the $\mathfrak{e}$ were defined in para. 33.

(ii) By para. 37, find the term $\mathfrak{q}$ of Z_μ which represents the functional q with $\mathfrak{e}$ as arguments, so that

$$\neg\mathfrak{A} \rightarrow \neg A[\mathfrak{e}_1\mathfrak{e}_2(\mathfrak{q}_1) \cdots \mathfrak{e}_n(\mathfrak{q}_1 \cdots \mathfrak{q}_{n-1})\mathfrak{q}_1 \cdots \mathfrak{q}_n] \tag{40.2}$$

can also be proved in Z_μ.

(iii) By the lemma of para. 33, find a term $\mathfrak{N}$ without variables $a_i^j \cdots b_i^j$ so that

$$(x_1) \cdots (x_{n-1})[(x_1 \leq \mathfrak{N} \ \& \cdots x_{n-1} \leq \mathfrak{N}) \rightarrow \mathfrak{e}_{r+1}(x_1 \cdots x_r) = e_{r+1}(x_1 \cdots x_r; \cdots a_i^j \cdots ; \cdots b_i^j \cdots)] \rightarrow \mathfrak{q} = q'(\cdots a_i^j \cdots ; \cdots b_i^j \cdots)$$
$$1 \leq r < n,$$

can be proved in Z_μ, where the primitive recursive terms e are defined in the lemma of para. 33, and q' represents the functional q with $e(\cdots ; \cdots a_i^j \cdots ; \cdots b_i^j \cdots)$ as arguments: i.e. an ordinal recursive function, by the note at the end of para. 38.

(iv) Let $\mathfrak{N}_1$ be the maximum of $\mathfrak{N}, \mathfrak{q}_1 \cdots \mathfrak{q}_n$.

(v) By the lemma of para. 33, define terms $\mathfrak{n}_i^j\ \mathfrak{m}_i^j$ and write out the proofs in Z_μ of the formulae

$$(y_1 \leq \mathfrak{N}_1 \& \cdots y_{n-1} \leq \mathfrak{N}_1) \rightarrow \mathfrak{e}_r(y_1 \cdots y_{r-1}) = e_r(y_1 \cdots y_{r-1}; \cdots \mathfrak{n}_i^j \cdots ; \cdots \mathfrak{m}_i^j \cdots).$$

(vi) From (v), (iv) and (iii) we get proofs of

$$\begin{aligned} \mathfrak{q} &= q'(\cdots \mathfrak{n}_i^j \cdots ; \cdots \mathfrak{m}_i^j \cdots) \\ \mathfrak{e}_1 &= e_1(\cdots \mathfrak{n}_i^1 \cdots ; \cdots \mathfrak{m}_i' \cdots) \\ \mathfrak{e}_{r+1}(\mathfrak{q}_1 \cdots \mathfrak{q}_r) &= e_{r+1}(q_1' \cdots q_r'; \cdots \mathfrak{n}_i^j \cdots ; \cdots \mathfrak{m}_i^j \cdots) \text{ where} \\ q_i' &= \eta_i^{(n)}[q'(\cdots \mathfrak{n}_i^j \cdots ; \cdots \mathfrak{m}_i^j \cdots)]. \end{aligned} \tag{40.3}$$

Now apply the Hilbert substitution method to the proofs in (i), (ii), and (vi). It provides substitutions $n_i^j\ m_i^j$ for the ϵ-terms $\mathfrak{n}_i^j\ \mathfrak{m}_i^j$ which reduce the given proof to a sequence of true formulae. In particular, $\neg\mathfrak{A}$ is reduced to a true formula, so is the conclusion in 40.2, and all the formulae 40.3. It follows that functions $e_1 \cdots e_n$ with the parameters $n_i^j\ m_i^j$ are the y and g required.

41. To check (δ), we have to show that if an A_n of the form

$$A[af_2(q_1) \cdots f_n(q_1 \cdots q_{n-1})q_1 \cdots q_n] \qquad 41.1$$

is verifiable for an ordinal recursive functional q, and $\mathfrak{B}$ can be proved from $\mathfrak{A}$ in an extension (c) of Z, then we find a functional q^*, so that

$$B[af_2(q_1^*) \cdots f_r(q_1^* \cdots q_{r-1}^*)q_1^* \cdots q_r^*]$$

is verifiable, $\mathfrak{B}$ being the formula

$$(x_1)(Ey_1) \cdots (x_r)(Ey_r)B(x_1 \cdots x_r y_1 \cdots y_r) \quad \text{and} \quad q_i^* = \eta_i^{(r)}(q^*).$$

(i) Again

$$A[\mathfrak{e}_1\mathfrak{e}_2(a_1) \cdots \mathfrak{e}_n(a_1 \cdots a_{n-1})a_1 \cdots a_n] \to \mathfrak{A}$$

can be proved in Z_μ.

(ii)

$$\mathfrak{B} \to (Ey)B[af_2(y_1) \cdots f_r(y_1 \cdots y_{r-1})y_1 \cdots y_r], \qquad y_1 = \eta_i^{(r)}(y)$$

can also be proved, i.e. in ϵ-notation

$$\mathfrak{B} \to B(af_2(\mathfrak{y}_1) \cdots f_r(\mathfrak{y}_1 \cdots \mathfrak{y}_{r-1})\mathfrak{y}_1 \cdots \mathfrak{y}_r], \qquad \mathfrak{y}_i = \eta_i^{(r)}(\mathfrak{y})$$

for an ϵ-term $\mathfrak{y}$.

(iii) Substitute in (i) $\eta_i^{(n)}(\mathfrak{q})$ for a_i, where $\mathfrak{q}$ is defined in para. 40 (ii), and add also the proofs of para. 40 (vi), which ensure that if $\mathfrak{e}_r(a_1 \cdots a_{r-1})$ is replaced by the function $e_r(a_1 \cdots a_{r-1})$, $\mathfrak{q}$ is replaced by the *value* of q for the arguments e.

Now, apply the substitution method to the formulae consisting of the proofs of $\mathfrak{B}$ from $\mathfrak{A}$, (ii) and (iii). By the final substitution all formulae will be reduced to true ones. Since 41.1 is verifiable and $\mathfrak{q}$ gets its 'true' value,

$$A[\mathfrak{e}_1\mathfrak{e}_2(\mathfrak{q}_1) \cdots \mathfrak{e}_n(\mathfrak{q}_1 \cdots \mathfrak{q}_{n-1})\mathfrak{q}_1 \cdots \mathfrak{q}_n]$$

is replaced by a true formula, hence also $\mathfrak{A}$, by the proof of $\mathfrak{B}$ from $\mathfrak{A}$ also $\mathfrak{B}$, and for any a and functions f we get a number $\mathfrak{h}$ so that

$$B[af_2(\mathfrak{h}_1) \cdots f_r(\mathfrak{h}_1 \cdots \mathfrak{h}_{r-1})\mathfrak{h}_1 \cdots \mathfrak{h}_r]$$

holds. This $\mathfrak{h}$ is easily checked to be bounded by an ordinal recursive functional of finite order containing the variables $a, f_2, \cdots, f_n$.

VI. 42. We shall apply the ideas of the present paper to a theorem of analytic number theory whose interpretation has given trouble. Littlewood proved in 1914 that for any ϵ and n there are numbers n_1 and n_2 so that $n_1 > n$, $n_2 > n$

$$\psi(n_1) - n_1 > (\tfrac{1}{4} - \epsilon)\sqrt{n_1} \log \log \log n_1$$

$$n_2 - \psi(n_2) > (\tfrac{1}{4} - \epsilon)\sqrt{n_2} \log \log \log n_2$$

where $\psi(n)$ is the logarithm of the l.c.m. of numbers $\leq n$, and also a similar result for the difference between $\pi(n)$ and $li(n)$. It was believed that the proof gave no indication on how the n_1 and n_2 were to be found. In the thirties Skewes took up this problem and Littlewood reports on Skewes' work in [5]. It seems worth while to give a systematic discussion of all this based on the present paper. The discussion will avoid the somewhat mysterious difficulties in stepping from $\psi(m)$ to $li(m)$, which are referred to in Littlewood's report.

43. By theorem II of section V we could be sure of finding the n_1 and n_2 if 42.1 were written out in one of the extensions (c) of Z which we have discussed. I have discussed elsewhere how proofs in large parts of the theory of functions of a complex variable can be presented in Z_μ. Here we shall only give a method of constructing rational approximations to zeros of computable regular functions. This will enable us to deal with the present problem if we remember our familiar principle of para. 26. If $(x)A(x) \to B$, and B is false, we find an $\mathfrak{n}$ so that $\neg A(\mathfrak{n})$; or in other words if $(x)A(x) \to B$, then B holds provided $A(n)$ is true for sufficiently many n. (A is assumed to be decidable.)

Computable regular functions.

43. From now onwards we consider only computable regular functions.

LEMMA A. *Suppose that*

(i) *$f(z)$ is regular, and $|f(z)/f(z_0)| < e^M$ in $|z - z_0| < r_0$,*

(ii) *$f(z)$ has at most N zeros in $|z - z_0| < \frac{1}{2} r_0$,*

(iii) *z' is at distance $\geq \delta$ from the nearest zero of $f(z)$.*

Then
$$|f(z')| > |f(z_0)|\, e^{-4M}(2\delta/r_0)^N.$$

By Landau's lemma, p. 15 of [9],

$$\left|\log \frac{f(z')}{f(z_0)} \Pi \left(\frac{z_0 - \rho}{z' - \rho}\right)\right| < 4M \qquad \text{if } |z' - z_0| < \tfrac{3}{8} r_0$$

where ρ are those zeros of $f(z)$ which satisfy $|\rho - z_0| < \frac{1}{2} r_0$. Thus

$$|f(z')| > |f(z_0)|\, e^{-4M}\, \Pi \left|\frac{z' - p}{z_0 - p}\right|.$$

Since $|z' - \rho| > \delta$, $|\rho - z_0| < \frac{1}{2} r_0$, and since the product contains no more than N factors, the lemma follows.

Application. Suppose D is a domain which can be covered by (a finite number of) circles C_i, $|z - z_i| < r_i$, and $|f(z_i)| \geq c$; further suppose that $f(z)$ is bounded in $\Sigma C_i'$ where C_i' is the circle $|z - z_i| < 2r_i$; lastly suppose that there are no more than N zeros of $f(z)$ in C_i'. Then, if z' in D is at a distance $\geq \delta$ from the nearest zero of $f(z)$

$$|f(z')| > c e^{-4M}(\delta/\max r_i)^N.$$

Observe in passing that these conditions are satisfied by the ζ-function in the rectangle $\mathfrak{R}$, $0 \leq \sigma_1 < \sigma < \sigma_2$, $0 < t_1 < t < T$, say, since $\mathfrak{R}$ can be covered by circles with centres on $\sigma = 1 + \eta$, $\eta > 0$ where $|\zeta| > c$, and $|\zeta|$ is bounded in these circles.

LEMMA B. *Suppose that*

(i) *$f(z)$ is regular in D, and satisfies the conditions of lemma A in D,*

(ii) *$|f'(z)| < F$ in D.*

Then we can locate the zeros of $f(z)$ in D in arbitrarily small squares of side X so that $f(z)$ exceeds some constant ϵ_0 on the sides of the X squares. Also for any Y we *can separate the zeros in the sense that there are no zeros of* $f(z)$ *at a distance* YX *from any square which contains a zero.*

(a) Without real loss of generality we suppose D to be rectangular, and we throw a grating of gauge X over D whose lines are horizontal and vertical. By shifting the grating arbitrarily little we can make $|f(z)| > \epsilon_0 > 0$ on the grating. To get a rough lower bound for ϵ_0, we argue:

Given $X > \delta > 0$ we shift the grating bodily through a distance $\frac{1}{3}\delta/N$ upwards and to the right, where N is a bound for the number of zeros of $f(z)$ in D. Repeat this $3N$ times. Now, the region swept out by the lines of the grating in each step is a grating $\mathfrak{G}$ of *strips*, each of width $\frac{1}{3}\delta/N$. There must be one grating $\mathfrak{G}_0$ which is at a distance $>\frac{1}{3}\delta/N$ from all the zeros of $f(z)$; since $f(z)$ satisfies lemma A, $|f(z)| > 2\epsilon_0$, say, on $\mathfrak{G}_0$. This grating can be found by trial: let $\delta_1 < \frac{1}{2}\epsilon_0/F$ and cover the total region swept out by the grating in the $3N$ steps by circles with rational centres and radius $<\delta_1$; this can be done by a *finite* number of circles. Calculate $f(z)$ at their centres to accuracy $\frac{1}{2}\epsilon_0$, which is possible since $f(z)$ is computable. There will be one grating ($\mathfrak{G}_0$) which is covered by circles at whose centres $|f(z)| > 2\epsilon_0$, and therefore $|f(z)| > \epsilon_0$ in $\mathfrak{G}_0$.

Now the integral of $-if'(z)/f(z)$ round the squares formed by the mesh $\mathfrak{G}_0$ can be evaluated: it need only be evaluated to accuracy π since it is a multiple of 2π. Thus the singularities of $f'(z)/f(z)$, and therefore the zeros of $f(z)$, are located in squares of side X.

(b) To separate the zeros we choose X so small that a generalization of the *Schubfachprinzip* can be used: if X_0 is the shorter side of D, take $X < X_0Y^{-N}$.

LEMMA C. *If* $f(z)$ *satisfies the conditions of lemma* A, *and if we have found a (rational)* z_1 *so that*

(i) $$|z_1 - z_0| < \tfrac{1}{4}r_0,$$

(ii) *there is a zero* ρ *of* $f(z)$ *so that* $|\rho - z_1| < \frac{1}{3}\chi$, $|\rho - z_0| < \frac{1}{4}r_0$,

(iii) *the zeros of* $f(z)$ *at a distance* $<(N-1)\chi$ *from* z_1 *lie in one quadrant,*

(iv) $$\chi < \tfrac{1}{4}r_0/M,$$

then $$|f'(z_1)/f(z_1)| > \frac{1}{\chi}.$$

By Landau's lemma

$$\left|\frac{f'(z_1)}{f(z_1)} - \sum_{|\rho-z_0|<\frac{1}{2}r_0}\frac{1}{z_1-\rho}\right| < \frac{4M}{r_0} \qquad \text{if} \qquad |z_1 - z_0| < \tfrac{1}{4}r_0,$$

$$\left|\sum_{|\rho-z_1|<(N-1)\chi}\frac{1}{z_1-\rho}\right| \geq \max\frac{1}{|\rho - z_1|} > \frac{3}{\chi}$$

$$\text{and} \quad \left|\sum_{(N-1)\chi<|\rho-z_1|,|\rho-z_0|<\frac{1}{2}r_0}\frac{1}{z_1-\rho}\right| < \frac{1}{\chi}$$

since in the latter sum $|\rho - z_1| > (N-1)\chi$, and there are at most $N-1$ such ρ.

Hence $$\left|\frac{f'(z_1)}{f(z_1)}\right| > \frac{3}{\chi} - \frac{1}{\chi} - \frac{4M}{r_0} > \frac{1}{\chi}.$$

COROLLARY. *If the zeros are located closely* χ *is small, and therefore* $|f'(z_1)/f(z_1)|$ *is large.*

43. We now use these lemmas to discuss 42.1. We base our remarks on theorem 34 of [4], and theorem H of [4].

Note that $\zeta(s)$ is computable since

$$\left| \zeta(s) - \sum_1^N \frac{1}{n^s} - \frac{1}{(s-1)N^{s-1}} \right| < \frac{1}{N^\sigma}$$

(actually, for rational s, $\neq 1$, the rational approximations of $\zeta(s)$ can be defined by a computable term of Z).

Step I. On the Riemann hypothesis we can find n_1 and n_2 of 42.1: from lemma 3 on p. 99 of [4], lemma 4, and a trivial restatement of the Phragmen-Lindelöf theorem on p. 95, we define two recursive functions $g_1(\epsilon, n)$, $g_2(\epsilon, n)$ so that for any ϵ and n either $g_1(\epsilon, n) > n$, $g_2(\epsilon, n) > n$

$$\begin{aligned} &\psi[g_1(\epsilon, n)] - g_1(\epsilon, n) > (\tfrac{1}{4} - \epsilon)\sqrt{g_1(\epsilon, n)} \log\log\log g_1(\epsilon, n), \\ &g_2(\epsilon, n) - \psi[g_2(\epsilon, n)] > (\tfrac{1}{4} - \epsilon)\sqrt{g_2(\epsilon, n)} \log\log\log g_2(\epsilon, n) \end{aligned} \qquad 43.1$$

holds, or there is a zero of $\zeta(s)$ in the rectangle $t < g^2$, $\frac{1}{2} + \eta < \sigma < 1$, $\eta > 0$; a lower bound for η can be given (p. 101). All this is also done in Littlewood's report.

Concerning the bound for η, note that it is to be expected from our principle, since if the conclusion 43.1 holds when the Riemann hypothesis is true, it should also hold when the Riemann hypothesis is nearly true: not *all* zeros need lie on $\sigma = \frac{1}{2}$, but only those whose imaginary part lies below a certain bound depending on ϵ and n, and they need not lie *on* the line $\sigma = \frac{1}{2}$, but near it.

Step II. Suppose then we have found an ϵ and an n so that 43.1 is false. Then the integral of $-i\zeta'(s)/\zeta(s)$ along the contour of

$$\mathfrak{R}: \tfrac{1}{2} + \eta < \sigma < 1, \qquad 300 < t < g^2$$

is $\geq 2\pi$, 300 being chosen since it is known that for $t < 300$ the zeros of the ζ-function lie on $\sigma = \frac{1}{2}$. By lemmas B and C the zeros in $\mathfrak{R}$ can be located arbitrarily closely, and separated.

44. The real problem is to show that if we have found a counter-example to 43.1, we can construct two more functions $g_{12}(\epsilon, n)$, $g_{22}(\epsilon, n)$ so that 42.1 holds. This and more is asserted by theorem H of p. 88 in [4]. But the theorem *mentions* an upper bound for the real parts of the zeros of ζ, which we do not necessarily find if 43.1 breaks down.

Actually, this upper bound is a mere flourish. To see this we observe (suppose now we have a counter-example to 43.1):

(1) Since a bound for the number of zeros of $\zeta(s)$, $|t| < T$ is given by theorem 25 of [4], we can separate the zeros in $\mathfrak{R}$.

(2) If $\sigma + it_0$ is the bottom right hand corner of any X square in $\mathfrak{R}$ which contains a zero, we can find the squares of side X of our mesh nearest to $1 + \eta + it_0$, $\eta > 0$, which contain a zero. We choose one of them, whose nearest

corner to $1 + \eta + it_0$ is $\sigma_1^* + it_1^*$. By choosing X sufficiently small at the start, we can ensure that $| \zeta'(s_1^*)/\zeta(s_1^*) |$ is as large as we please.

(3) If we consider the difference between $\psi(n)$ and n, we take

$$f(s) = \int_1^\infty \frac{c(x)}{x^{1+s}}\, dx \qquad \text{for } \sigma > 1, \qquad \text{where } c(x) = s[\psi(x) - x \pm c_0 x^{\frac{1}{2}(1+\eta)}]$$

and

$$f(s) = -\zeta'(s)/\zeta(s) - s/(s-1) \pm c_0 s[s - \tfrac{1}{2}(1+\eta)]^{-1}$$

(theorem 33 of [4]).

To study the difference between $\pi(n)$ and $li(n)$ we take the analogous Dirichlet integral.

(4) Since, by theorem 23 of [4], $| \psi(x) - x | < Axe^{-a\sqrt{\log x}}$, for $\sigma \geq 1 + \eta$, $\eta > 0$, a bound for

$$\int_X^\infty \frac{| c(x) |}{| x^{1+s} |}\, dx$$

can be given which tends to zero as X tends to infinity.

(5) $f(s)$ is regular in $| s - 1 - \eta | < \frac{1}{2}$.

(6) $f(s)$ is regular in $| s - 1 - \eta - it_0 | < | s_1^* - 1 - \eta - it_0 | + \epsilon_1$ for some ϵ_1 since by the construction of the squares in lemma B, and the maximum modulus principle, $| \zeta(s) | > \epsilon_0$ outside the squares which contain a zero.

45. Now we use Landau's theorem H; it says: if $f(s) = \int_1^\infty c(x)x^{-1-s}\, dx$ for $\sigma > 1$, and if for all x $c(x) \geq 0$, then $| f^{(n)}(1 + \eta) | \geq | f^{(n)}(1 + \eta + it_0) |$. But if for all n $| f^{(n)}(1 + \eta) | \geq | f^{(n)}(1 + \eta + it_0) |$, $f(s)$ can be continued as far from $1 + \eta + it_0$ as from $1 + \eta$. If it can't there must be an x so that $c(x) < 0$. By our principle we should find the x. In detail:

THEOREM H*.

(1) *If* $\int_1^\infty c(x)x^{-1-s}\, dx$ *converges for* $\Re(s) \geq 1 + \eta > 1$ *to* $f(s)$,

(2) $f(s)$ *is regular and* $<M$ *in* $| s - 1 - 2\eta | < c + \epsilon$, $\epsilon > 0$,

(3) *for any* $N > 0$ *we can find an* s_1^* *so that* $| f(s_1^*) | > 2M(c + \epsilon)\epsilon^{-1} + N$, $| s_1^* - 1 - 2\eta - it_0 | \leq c - \epsilon' - \epsilon''$, $\epsilon' > 0$, $\epsilon'' > 0$ (*where* s_1^* *depends on* N),

(4) $f(s)$ *is regular and* $<M'$ *in* $| s - 1 - 2\eta - it_0 | < c - \epsilon'$,

then $c(x)$ *changes sign infinitely often, and we can give a recursive bound to the* n^{th} *change of sign.*

Observe that it is sufficient to give a bound for one change of sign. Since $\int_1^n c(x)x^{-1-s}\, dx$ is an integral function of s, conditions 45.1 apply to $\int_n^\infty c(x)x^{-1-s}\, dx$ if they apply to $\int_1^\infty c(x)x^{-1-s}\, dx$, so that we get a change of sign for $x > n$.

Next note that if for some n

$$\left| \frac{d^n f}{ds^n}(1 + 2\eta + it_0) \right| \geq \left| \frac{d^n f}{ds^n}(1 + 2\eta) \right| + \gamma, \qquad \gamma > 0, \qquad 45.2$$

$c(x)$ changes sign for some $x < X$ where

$$\int_X^\infty \frac{|\,c(x)\,|\,(\log x)^n}{x^{1+2\eta}}\,dx \leq \tfrac{1}{2}\gamma$$

and this inequality has a solution by (1).

Lastly we show that 45.2 is implied by the conditions 45.1 for some

$$n < \log \frac{M\epsilon''(c+\epsilon)}{M'\epsilon'(c-\epsilon')} \Big/ \log \frac{c-\epsilon'}{c-\epsilon'-\epsilon''}.$$

For, by Cauchy's inequality and (2), $|\,f^{(r)}(1+2\eta)\,| < r!M(c+\epsilon)^{-r}$ and by (4) $|\,f^{(r)}(1+2\eta+it_0)\,| < r!M'(c-\epsilon')^{-r}$.

By Taylor's expansion

$$\left| f(s_1^*) - \sum_0^n \frac{(s_1^*-1-2\eta-it_0)^r}{r!} f^{(r)}(1+2\eta+it_0) \right|$$

$$\leq \sum_{n+1}^\infty \frac{|\,c-\epsilon'-\epsilon''\,|^r}{r!}\,|\,f^{(r)}(1+2\eta+it_0)\,| < M'\left(\frac{c-\epsilon'-\epsilon''}{c-\epsilon'}\right)^{n+1}\frac{c-\epsilon'}{\epsilon''}.$$

Thus by (3)

$$\left| \sum_0^n \frac{(s_1^*-1-2\eta-it_0)^r}{r!} f^{(r)}(1+2\eta+it_0) \right| > N + \frac{M(c+\epsilon)}{\epsilon}.$$

Since $$\sum_0^\infty \frac{|\,c-\epsilon'-\epsilon''\,|^r}{r!}\,|\,f^{(r)}(1+2\eta)\,| < \frac{M(c+\epsilon)}{\epsilon+\epsilon'+\epsilon''}$$

$$|\,f^{(r)}(1+2\eta)\,| < |\,f^{(r)}(1+2\eta+it_0)\,| - \frac{N}{n} \quad \text{for some } r \leq n.$$

Step III. If 43.1 is false, we find by para. 44, s_1^*, c, ϵ, ϵ', ϵ'' so that $c(x)$ and $f(s)$ of 44(3) satisfy the conditions 45.1.

Appendix I. The naive interpretation of bound variable formulae of the predicate calculus of first order to which suitable free variable formulae have been added as axioms, is false.

The proof is nowadays trivial. Let $Prov(a, b)$ be the primitive recursive formula which holds if and only if a is the number of a proof in Z_μ of the formula with number b ([2], II, 294–306). Let $s(a, b)$ be the number of the formula which is got by substituting the numeral a for the free variable in the formula with the number b (ibid. 307).

Let $A(a, b)$ denote the formula $Prov[a, s(b, b)]$. By the predicate calculus of first order

$$(Ey)A(yx) \vee (z)\overline{A(zx)}$$

i.e.

$$(x)(Ey)(z)[A(yx) \vee \neg A(zx)],$$

but for any computable function $\phi(n)$ we determine numbers $\mathfrak{x}$ and $\mathfrak{z}$ so that

$$A[\phi(\mathfrak{x}), \mathfrak{x}] \vee \neg A[\mathfrak{z}, \mathfrak{x}]$$

is false.

PROOF. *Let $\phi(b)$ be a computable term of Z_μ with free variable b, and let the number of the formula $\neg A[\phi(b), b]$, i.e. of $\neg Prov[\phi(b), s(b, b)]$ be $\mathfrak{x}$; then the number of $\neg A[\phi(\mathfrak{x}), \mathfrak{x}]$ is $s(\mathfrak{x}, \mathfrak{x})$, and $\neg A[\phi(\mathfrak{x}), \mathfrak{x}]$ is true: for, if it were false, $\phi(\mathfrak{x})$ would be the number of a proof of the formula with number $s(\mathfrak{x}, \mathfrak{x})$, i.e. of $\neg A[\phi(\mathfrak{x}), \mathfrak{x}]$; since Z_μ is consistent, this may be excluded. On the other hand, since $\phi(\mathfrak{x})$ is computable, there is a proof in Z_μ of $\neg A[\phi(\mathfrak{x}), \mathfrak{x}]$, the proof having the number $\mathfrak{z}$ say. Then $A[\mathfrak{z}, \mathfrak{x}]$ is true.*

Since by [2], II, 392–403, a function which is computable in any of the general senses, can be expressed by a term of Z_μ, our result is established.

Note in passing that by the same argument the monotone decreasing sequence of positive rationals

$$a_n = \sum_{m=1}^{n} a(n, m)2^{-m},$$

where $prov(a, b) = 0$ or $= 1$,

$$a(n, m) = \prod_{r=1}^{n} prov(r, m) \qquad \text{and} \qquad Prov(a, b) \sim prov(a, b) = 0,$$

have no computable limit.

Appendix II. We have to show that for systems (b) of para. 18 which include all primitive recursive formulae, there is no interpretation in which every formula of the system is associated with a *single* free variable formula. (We prove a little more, namely that there is no interpretation which associates a conjunction of numerical formulae with every formula of the system.) By the usual Gödel method we get the following theorem:

THEOREM. *Let the system considered be an extension* (b) *of* Σ, *a system of type* 18 (a), *which contains all 0-primitive recursive terms. Let $\mathfrak{h}_1(l, n, r) = \mathfrak{h}(l, n, r)$ (of* [2], II, 410) *if for $s < r$ we have $\mathfrak{h}(l, n, r) \neq 0$, and let $\mathfrak{h}_1(l, n, r) = 1$ otherwise; let the quasi-recursive function $f(n, m)$ be represented by $\nu\{\mu_r[\mathfrak{h}_1(l, 2^n 3^m, r) = 0], 0\}$, where $f(n, m)$ is the number of a formula decidable in Z_μ when m is the number of a formula of Σ (in some suitable numbering). Then we find a formula Q of the system Σ which either can be proved in the system although there is an $\mathfrak{n}$ so that $f(\mathfrak{n}, \mathfrak{q})$ is the number of a false formula, or else can be disproved in an extension* (b) of Σ *although $f(n, \mathfrak{q})$ is, for all n, the number of a true formula.*

Note. Both $\mathfrak{h}_1$ and ν are 0-primitive recursive functions.

Let $s(a, b)$ be the primitive recursive substitution function for Σ, and $\mathfrak{e}(z)$ the number of the negation of the formula of Σ with number z, and $Prov(a, b)$ the 'proof predicate' for Z_μ.

Consider the formula Q_0,

$$(Ex)(Ey)(Ez)(Eu)\{\mathfrak{h}_1(l, 2^x 3^{s(a,a)}, y) = 0 \ \& \ \nu(y, 0) = z \ \& \ Prov[u, \mathfrak{e}(z)]\},$$

with number $\mathfrak{q}_0$, substitute $\mathfrak{q}_0$ for a so that the resulting formula Q has the number $\mathfrak{q}$.

If we can find numbers x, y, z, u to satisfy Q, then $f(x, \mathfrak{q})$ is the number of a false formula, $\mathfrak{q}$ of a true one.

If there are no such numbers, we add the verifiable free variable formula

$$\mathfrak{h}_1(l, 2^x 3^{\mathfrak{q}}, y) = 0 \;\&\; \nu(y, 0) = z \cdot \rightarrow \neg\, Prov[u, \mathfrak{e}(z)]$$

to the system Σ. Then Q can be disproved in this extension (b) of Σ, but no formula with the number $f(n, \mathfrak{q})$ can be disproved in Z_μ. Since these formulae are decidable in Z_μ, they are true.

THE UNIVERSITY, READING, ENGLAND