Introductory note to **1938a*

The text that follows consists of notes for a lecture Gödel delivered in Vienna on 29 January 1938 to a seminar organized by Edgar Zilsel. The lecture presents an overview of possibilities for continuing Hilbert's program in a revised form. It is an altogether remarkable document: biographically, it provides, together with **1933o* and **1941*, significant information on the development of Gödel's foundational views; substantively, it presents a hierarchy of constructive theories that are suitable for giving (relative) consistency proofs of parts of classical mathematics (see §§2–4 of the present note); and, mathematically, it analyzes Gentzen's 1936 proof of the consistency of classical arithmetic in a most striking way (see §7). A surprising general conclusion from the three documents just mentioned is that Gödel in those years was intellectually much closer to the ideas and goals pursued in the Hilbert school than has been generally assumed (or than can be inferred from his own published accounts).

1. The setting.

Edgar Zilsel (1891–1944) had been connected with the Vienna Circle. By 1938 his main interest was history and sociology of science.[a] The concrete stimulus for Gödel's preparing the lecture was Zilsel's question whether anything new had happened in the foundations of mathematics and his request that Gödel should describe the "status of the consistency question" to the seminar. This is reported in Gödel's notes[b] concerning

[a]Cf. *Zilsel 1976*, which collects in German translation articles published in English after Zilsel's emigration. The English versions are collected in *Zilsel 199?*. Zilsel had taught physics and philosophy at the *Volkshochschule* in Vienna, but as a result of the Dollfuss coup in 1934, he was dismissed from or eased out of this position (see *Behrmann 1976* and *Dvořak 1981*, pp. 23–25) and thereafter worked as a *Gymnasium* teacher. He emigrated in 1938 to England and moved on to America the next year. In 1944, while teaching at Mills College in California, he committed suicide.

Testimony differs as to whether Zilsel was a member of the Vienna Circle or merely someone sympathetic to their views who attended some of the sessions; see *Dvořak 1981*, pp. 30–31.

[b]The notes (*Zusammenkunft bei Zilsel*) are in Gödel's *Nachlass* (document no. 030114) and were transcribed from Gabelsberger by Cheryl Dawson. In editing the text and preparing this note, we have also used a document (no. 040147) entitled *Konzept* (i.e., draft), evidently an earlier draft of Gödel's notes for this lecture. On our use of this document, see the textual notes.

the organizational meeting of the seminar on 2 October 1937 at Zilsel's home; it was at this organizational meeting that Zilsel made his request and that Gödel, after some reflection, agreed to speak on the consistency question.

As Zilsel did not have a university position, the seminar was probably an informal, private affair and continued to meet at Zilsel's home.[c] Gödel's notes do not give a clear view of the seminar's intended theme, except that it was rather general. The list of names of those present or mentioned as possible participants confirms this; they spread over several fields and include no one, except for Gödel, who was much involved in mathematical logic.[d] In his immediate response to Zilsel's request Gödel suggested presenting a German version of the lecture **1933o*, given in Cambridge, Mass. (published in this volume); but on reflection he added that that talk was "zu prinzipiell", which can be translated roughly as "too general".

The Zilsel lecture gives, as we remarked, an overview of possibilities for a revised Hilbert program. The central element of that program was to prove the consistency of formalized mathematical theories by finitist means. Gödel's 1931 incompleteness theorems have been taken to imply that for theories as strong as first-order arithmetic this is impossible, and indeed, so far as Gödel ventures to interpret Hilbert's finitism, that is Gödel's view in the present text as well as earlier in **1933o* (though not in *1931*) and later in **1941*, *1958* and *1972*. The crucial questions then are what extensions of finitist methods will yield consistency proofs, and what epistemological value such proofs will have.

Two developments after *Gödel 1931* are especially relevant to these questions. The first was the consistency proof for classical first-order arithmetic relative to intuitionistic arithmetic obtained by Gödel (*1933e*). The proof made clear that intuitionistic methods went beyond finitist ones (cf. footnote j below). Some of the issues involved had been discussed in Gödel's lecture **1933o*, but also in print, for example in *Bernays 1935* and *Gentzen 1936*. Most important is Bernays' emphasis on the "abstract element" in intuitionistic considerations.[e] The second development was Gentzen's consistency proof for first-order arithmetic using as the additional principle—justified from an intuitionistic

[c]Karl Popper writes (*1976*, p. 84) of having given a paper to a gathering there several years earlier.

It is likely that the seminar did not continue long after Gödel's talk, since the *Anschluß* took place only six weeks later, and not long after that Zilsel emigrated (cf. fn. a).

[d]We are indebted to Katalin Makkai for researching this matter.

[e]Significantly, Gödel refers to that earlier discussion in *1958*.

standpoint—transfinite induction up to ϵ_0. Already in *1933o* (page 31) Gödel had speculated about a revised version of Hilbert's program using constructive means that extend the limited finitist ones without being as wide and problematic as the intuitionistic ones:

> But there remains the hope that in future one may find other and more satisfactory methods of construction beyond the limits of the system A ⟦capturing finitist methods⟧, which may enable us to found classical arithmetic and analysis upon them. This question promises to be a fruitful field for further investigations.

The Cambridge lecture does not suggest any intermediate methods of construction; by contrast, Gödel presents in the Zilsel lecture two "more satisfactory methods" that provide bases to which not only classical arithmetic but also parts of analysis might be reducible: quantifier-free theories for higher-type functionals and transfinite induction along constructive ordinals. Before looking at these possibilities, we sketch the pertinent features of the Cambridge talk, because they give a very clear view not only of the philosophical and mathematical issues Gödel addresses, but also of the continuity of his development.[f]

2. Relative consistency.

Understanding by mathematics "the totality of the methods of proof actually used by mathematicians", Gödel sees the problem of providing a foundation for these methods as falling into two distinct parts (page 1):

> At first these methods of proof have to be reduced to a minimum number of axioms and primitive rules of inference, which have to be stated as precisely as possible, and then secondly a justification in some sense or other has to be sought for these axioms, i.e., a theoretical foundation of the fact that they lead to results agreeing with each other and with empirical facts.

The first part of the problem is solved satisfactorily through type theory and axiomatic set theory, but with respect to the second part Gödel considers the situation to be extremely unsatisfactory. "Our formalism", he contends, "works perfectly well and is perfectly unobjectionable as long as we consider it as a mere game with symbols, but as soon as

[f]Cf. Solomon Feferman's detailed introductory note to *1933o* in this volume, p. 36.

we come to attach a meaning to our symbols serious difficulties arise" (page 15). Two aspects of classical mathematical theories (the non-constructive notion of existence and impredicative definitions) are seen as problematic because of a necessary Platonist presupposition "which cannot satisfy any critical mind and which does not even produce the conviction that they are consistent" (page 19). This analysis conforms with that given in the Hilbert school, for example in *Hilbert and Bernays 1934*, *Bernays 1935* and *Gentzen 1936*.[g] Gödel expresses the belief, again as the members of the Hilbert school did, that the inconsistency of the axioms is most unlikely and that it might be possible "to prove their freedom from contradiction by unobjectionable methods".

Clearly, the methods whose justification is being sought cannot be used in consistency proofs, and one is led to the consideration of parts of mathematics that are free of such methods. Intuitionistic mathematics is a candidate, but Gödel emphasizes (page 22) that

> the domain of this intuitionistic mathematics is by no means so uniquely determined as it may seem at first sight. For it is certainly true that there are different notions of constructivity and, accordingly, different layers of intuitionistic or constructive mathematics. As we ascend in the series of these layers, we are drawing nearer to ordinary non-constructive mathematics, and at the same time the methods of proof and construction which we admit are becoming less satisfactory and less convincing.

The strictest constructivity requirements are expressed by Gödel (pages 23–25) in a system A that is based "exclusively on the method of complete induction in its definitions as well as in its proofs". That implies that the system A satisfies three general characteristics:[h] (**A1**) Universal quantification is restricted to "infinite totalities for which we can give a finite procedure for generating all their elements"; (**A2**) Existential statements (and negations of universal ones) are used only as abbreviations, indicating that a particular (counter-)example has been found without—for brevity's sake—explicitly indicating it; (**A3**) Only decidable notions and calculable functions can be introduced. As the method of complete induction possesses for Gödel a particularly high degree of evidence, "it would be the most desirable thing if the freedom from contradiction of ordinary non-constructive mathematics could be proved by methods allowable in this system A" (page 25).

[g]Of these writings *Bernays 1935* is more ready to defend Platonism, with certain qualifications.

[h]The designations (**A1**)–(**A3**) are introduced by us for ease of reference.

Gödel infers that Hilbert's original program is unattainable from two claims: first, *all* attempts for finitist consistency proofs actually undertaken in the Hilbert school operate within system A; second, *all* possible finitist arguments can be carried out in analysis and even classical arithmetic. The latter claim implies jointly with the second incompleteness theorem that finitist consistency proofs cannot be given for arithmetic, let alone analysis. Gödel puts this conclusion here quite strongly: "... unfortunately the hope of succeeding along these lines ⟦using only the methods of system A⟧ has vanished entirely in view of some recently discovered facts" (page 25). But he points to interesting partial results and states the most far-reaching one, due to *Herbrand 1931* in a beautiful and informative way (page 26):

> If we take a theory which is constructive in the sense that each existence assertion made in the axioms is covered by a construction, and if we add to this theory the non-constructive notion of existence and all the logical rules concerning it, e.g., the law of excluded middle, we shall never get into any contradiction.

Gödel conjectures that Herbrand's method might be generalized to treat Russell's "ramified type theory", i.e., we assume, the theory obtained from system A by adding ramified type theory instead of classical first-order logic.[i]

There *are*, however, more extended constructive methods than those formalized in system A; this follows from the observation that system A is too weak to prove the consistency of classical arithmetic together with the fact that the consistency of classical arithmetic can be established relative to intuitionistic arithmetic.[j] The relative consistency proof is made possible by the intuitionistic notion of absurdity, for which "exactly the same propositions hold as do for negation in ordinary mathematics—at least, this is true within the domain of arithmetic"

[i]In *Konzept*, p. 0.1, Gödel mentions Herbrand's results again and also the conjecture concerning ramified type theory. The obstacle for an extension of Herbrand's proof is the principle of induction for "transfinite" statements, i.e., formulae containing quantifiers. Interestingly, as discovered in *Parsons 1970*, and independently by Mints (*1971*) and Takeuti (*1975*, p. 175), the induction axiom schema for purely existential statements leads to a conservative extension of A, or rather its arithmetic version, primitive recursive arithmetic. How Herbrand's central considerations can be extended (by techniques developed in the tradition of Gentzen) to obtain this result is shown in *Sieg 1991*.

[j]In his introductory note to *1933e* (these *Works*, Vol. I, p. 284), Troelstra mentions relevant work also of Kolmogorov, Gentzen and Bernays. Indeed, as reported in *Gentzen 1936*, p. 532, Gentzen and Bernays discovered essentially the same relative consistency proof independently of Gödel. According to Bernays (*1967*, p. 502), the above considerations made the Hilbert school distinguish intuitionistic from finitist methods. Hilbert and Bernays (*1934*, p. 43) make the distinction without referring to the result discussed here.

(page 29). This foundation for classical arithmetic is, however, "of doubtful value": the principles for absurdity and similar notions (as formulated by Heyting) employ operations over *all* possible proofs, and the totality of all intuitionistic proofs cannot be generated by a finite procedure; thus, these principles violate the constructivity requirement (**A1**).

Despite his critical attitude towards Hilbert and Brouwer, Gödel dismisses neither in **1933o* when trying to make sense out of Hilbert's program in a more general setting,[k] namely, as a challenge to find consistency proofs for systems of "transfinite mathematics" relative to "constructive" theories. And he expresses his belief that epistemologically significant reductions may be obtained.

3. Layers of constructivity.

In his lecture at Zilsel's, Gödel explores options for such relative consistency proofs and goes beyond **1933o* by considering—in detail—three different and, as he points out, *known*[l] ways of extending the arithmetic version of system A: the first one uses higher-type functionals, the second introduces absurdity and a concept of consequence, and the third adds the principle of transfinite induction for concretely defined ordinals of the second number class. The first way is related to the hierarchy of functionals in *Hilbert 1926* and is a precursor of the *Dialectica* interpretation; the very brief section IV of the text is devoted to it. Of course, the second way is based on intuitionistic proposals, whereas the last way is due to Gentzen. The second and third way are discussed extensively in sections V and VI.

The broad themes of the preceding discussion are taken up in sections I through III, putting the notion of reducibility first; a theory T is called *reducible to a theory* S if and only if either

$$S \text{ is a subsystem of } T \text{ and } S \text{ proves } Wid\ S \rightarrow Wid\ T$$

or

$$S \text{ proves } Wid\ T.$$

As an example for the first type of reducibility Gödel alludes to his own relative consistency proof for the axiom of choice; as an example

[k]This is in contrast to the lecture at Zilsel's, where he makes some (uncharacteristically polemical) remarks against the members of the Hilbert school and against Gentzen in particular; cf. fn. ss below.

[l]When introducing the three ways of extending the basic system, Gödel starts out by saying: Drei Wege sind bisher bekannt (i.e., Three ways are known up to now).

for the second type of reducibility he mentions the reduction of analysis to logic (meaning, we assume, simple type theory). Concerning the epistemological side of the problem, Gödel emphasizes that a proof is satisfactory in the first case only if S is a proper part of T, and in the second case only if S is more evident, more reliable than T. Though he admits that the latter criterion is subjective, he points to the fact that there is general agreement that constructive theories are better than non-constructive ones, i.e., those that incorporate "transfinite" existential quantification.

Acknowledging the vagueness of the notion of constructivity, Gödel formulates in section II what he calls a *Rahmendefinition* that incorporates the requirements (**A1**) through (**A3**) which, in **1933o*, had motivated the system A. Now there are four conditions:[m] (**R1**) restates (**A3**), namely, that the primitive operations must be computable and that the basic relations must be decidable; (**R2**) combines aspects of (**A1**) and (**A2**) to restrict appropriately the application of universal and existential quantifiers; (**R3**) is an open list of inference rules and axioms that includes defining equations for primitive recursive functions, axioms and rules of the classical propositional calculus, the rule of substitution, and ordinary complete induction (for quantifier-free formulae); (**R4**) indicates what was in **1933o* the positive motivation for restricting universal quantification, namely, the finite generation of objects: "Objects should be surveyable (that is, denumerable)."

This point of view is modified in **1941*, where we find (on pages 2 and 3) versions of the first two conditions. The list of basic axioms and rules is introduced later in a very similar way, though not as part of the *Rahmendefinition*, but rather as part of the specification of the system Σ for finite-type functionals. There is no analogue to (**R4**). These changes (in exposition) are most interesting, as both (**R3**) and (**R4**) are viewed as "problematic" in the lecture at Zilsel's. As to (**R3**), the restriction to induction for just natural numbers is viewed as problematic, because induction for certain transfinite ordinals is also evident; (**R4**) is problematic "because of the concept of function". Here, it seems, is the germ of Gödel's analysis in *1958* of the distinction between (strictly) finitist and intuitionistic considerations.

Gödel focuses in the present text immediately on number theory. The theory that corresponds to system A is obviously *primitive recursive arithmetic*, PRA, and it is viewed as the fundamental system in the hierarchy of constructive systems described briefly in section III.

[m]For convenience in our further discussions, we use (**Ri**) to refer to these conditions, which were simply numbered by Gödel.

4. Higher-type functionals.

The first kind of extension of PRA consists in the introduction of (defining equations for) functionals of finite type. Gödel suggests continuing, as Hilbert does in *1926*, the introduction of types into the transfinite, "if it is demanded that types only ⟦be admitted⟧ for those ordinal numbers which have been defined in an earlier system". The mathematical details for the basic finite type theory, as well as its transfinite extension, are sketchy. Nevertheless, this extension is most interesting for at least three reasons: It is the only extension that, according to Gödel, satisfies *all* requirements of the "Rahmendefinition"; it is connected to Hilbert and Ackermann's hierarchies of recursive functionals; and, finally, it is the first known articulated, even if very rudimentary, step in the evolution of Gödel's *Dialectica* interpretation.

Gödel indicates only by an example how to extend PRA by a recursion schema for functionals; the example he gives is as follows:

$$\Phi(f,1,k) = f(k)$$
$$\Phi(f,n+1,k) = f(\Phi(f,n,k)).$$

This definition of the iteration functional is almost identical to the ones given in *Hilbert 1926*, page 186, and *Ackermann 1928*, page 118 (*van Heijenoort 1967*, pages 388 and 495 respectively), where the iteration functional is used for defining the Ackermann function.[n] The latter function was, as will be recalled, the first example of a calculable function that cannot be defined by ordinary primitive recursion. However, it can be defined by primitive recursion if higher-type objects are allowed. Hilbert (*1926*, page 186; *van Heijenoort 1967*, page 389) formulated the full definition schema as follows:

$$\rho(\mathfrak{g},\mathfrak{a},0) = \mathfrak{a}$$
$$\rho(\mathfrak{g},\mathfrak{a},n+1) = \mathfrak{g}(\rho(\mathfrak{g},\mathfrak{a},n),n);$$

$\mathfrak{a}$ is of arbitrary type (and ρ and $\mathfrak{g}$ must satisfy the obvious type restrictions). It is not entirely clear, but certainly most plausible, that Gödel envisioned using some form of the full schema (for what he calls here "closed systems").

Gödel claims that all his requirements are satisfied. Equality of

[n]This "iterator" suffices to define the recursor functional, which directly yields the schema of definition by primitive recursion. See *Diller and Schütte 1971* or *Troelstra 1973*, theorem 1.7.11 (p. 56).

higher-type objects could pose a problem for (**R1**); on this point see below. The claim that (**R4**) is satisfied would seem to mean that Gödel thought of the higher-type variables as ranging over the primitive recursive functionals themselves, which are given by terms. But then it is puzzling that Gödel earlier says that (**R4**) is "problematic ... because of the concept of function" (page 3).

Various claims and conjectures are formulated in subsections IV, 4 and 5. We shall make some frankly speculative remarks, which, together with (the introductory note to) **1941*, may help the reader to speculate further on Gödel's statements. However, before discussing the assertions in subsection 4, we try to clarify the character of the extension procedure that is claimed to "contain" certain additions. The procedure is described more directly in a remark of *Konzept* (quoted in full below—see editorial note n to the text); the recursion schema mentioned there is the schema depicted above for defining the iteration functional:

> In general this recursion schema ⟦is used⟧ for the introduction of Φ_i ⟦with the help of⟧ earlier f_i and the functions obtained by substituting in one another. ... This hierarchy can be continued by introducing functions whose arguments are such Φ and admitting once again recursive definitions according to a numerical parameter. And that can even be extended into the transfinite.

Already in *Ackermann 1928*, pages 118–119 (*van Heijenoort 1967*, page 495), there is a discussion of a hierarchy A_i of classes of functionals. The elements of A_i are, in van Heijenoort's terminology (*ibid.*, page 494), functionals of level i, i.e., their arguments are at most of level $(i-1)$ and their values are natural numbers. Gödel does not explicitly envisage equations of terms of higher type; if he indeed does not, then his hierarchy corresponds to Ackermann's. Otherwise, one is led quite directly to a hierarchy of classes H_i, restricted to functionals of type level $\leq i$, where numbers are of type level 0, and functionals with arguments of type σ and values of type τ are of type level $\max[\text{level}(\sigma)+1, \text{level}(\tau)]$.[o]

Gödel began section III with the statement, "The finitary systems form a hierarchy." In section VI he uses S_1 and S_ω, evidently for systems in this hierarchy. A reasonable conjecture is that they are the systems for functionals of higher type in either the hierarchy A_i or H_i.

[o]This reading would, on the conjecture stated immediately below, make the systems S_i correspond roughly to the subsystems $\mathbf{T}_i$ of the $\mathbf{T}$ of *Gödel 1958* described in *Parsons 1972*, section 4. Parsons uses the term "rank".

It is, however, more likely that Gödel envisaged systems with equations only of numbers, thus closer to that of *Spector 1962* than to the **T** of *Gödel 1958*.[p] Otherwise one would expect some mention of the problems involved in interpreting equations of higher type so as to be decidable. (Hilbert also did not discuss the interpretation of such equations.) Such equations do occur in **1941*, pages 15 and 17, but from the notes of Gödel's lectures on intuitionistic logic at Princeton at the time it is clear that this use was informal, and the Σ of **1941* had only equality of numbers as a primitive.[q]

Although it would be natural today to take S_ω as the union of the S_i and thus roughly the Σ of **1941* (or perhaps the **T** of *1958*), this is hard to reconcile with Gödel's later statement (page 13) that recursion on ϵ_0 may be obtainable in S_ω. We will understand S_ω as a system containing functionals of lowest transfinite type (what on the other understanding would be $\mathrm{S}_{\omega+1}$).

The "addition of recursion on several variables", as 4.1 asserts, is contained in the procedure since the schema allows nested recursion. The claim in 4.2, "addition of the statement *Wid*", seems to be clear, as S_{i+1} proves the consistency of S_i; that also fits well with the discussion of the provability predicates B_i on page 10. Finally, in 4.3 the "addition of Hilbert's rule of inference" is claimed to be contained in the procedure. In *1931c* Gödel reviews Hilbert's paper in which the infinitary rule was introduced, and in *1933e* he considers Herbrand's formulation of classical arithmetic, which includes that rule in the following form (*Herbrand 1931*, page 5; *Herbrand 1971*, page 291):

> Let $A(x)$ be a proposition without apparent variables ⟦i.e., a quantifier-free formula⟧; if it can be proved by intuitionistic procedures that this proposition, intuitionistically considered, is true for every x, then we add $(x)A(x)$ to the hypotheses.

As, for Herbrand, intuitionistic and finitistic considerations are coextensive, that is exactly Hilbert's formulation. A mathematically precise version of Gödel's claim follows from *Rosser 1937*; cf. also Feferman's introductory note to *Gödel 1931c* (these *Works*, Volume I, page 208).

[p]In his introductory note to *1958* and *1972*, Troelstra points out that a subsystem $\mathbf{T}_0$ of **T** is sufficient for Gödel's interpretation of first-order arithmetic; this subsystem is common to **T** and the extensional variant. Thus, to obtain the Gödel interpretation in the latter, extensionality is needed only to derive Troelstra's "replacement schemas" (3) (these *Works*, Vol. II, p. 224).

On the problems of higher-type equality, see further the informative remarks in the same note, pp. 227–229.

[q]See the introductory note to **1941* in this volume, p. 186.

There is no indication in this section as to how the system for functionals (of finite type) is to be used in consistency proofs of classical or, respectively, intuitionistic number theory; there is not even an explicit claim that the consistency of number theory can be established relative to $\cup \mathrm{S}_n$ (or in S_ω; see above). In subsection 5 of the present section, Gödel formulates only negative claims and conjectures. In 5.1, he says that "with finite types one cannot prove the consistency of number theory". We assume that he intended to say what he formulated at the end of the Yale lecture, namely, the system that goes up only "to a given finite type" is not sufficient, instead the system for all finite types is needed. This can be rephrased in our terminology: No system S_i, $i < \omega$, will suffice; their union is needed for relative consistency. The negative conjecture in 5.2 parallels that for transfinite induction in VI, 13: Even when the extension procedure is iterated transfinitely, along ordinals satisfying the restrictive condition mentioned above, one will not be able to prove the consistency of analysis. This conjecture is in stark contrast to that for the "modal-logical" route, which, according to Gödel in section V, 11, "leads furthest" and by means of which the consistency of analysis is "probably obtainable".

5. The modal-logical route.

In section V, Gödel turns to the detailed examination of the "modal-logical" route, that is, giving consistency proofs relative to intuitionistic systems of the sort that had been first formalized by Heyting. Intuitionistic mathematics, as a framework for such proofs, does not satisfy the conditions Gödel laid down at the outset, because of the free application of negation and the conditional. Thus the now well-known translation of classical into intuitionistic arithmetic (from *Gödel 1933e*) gives a simple relative consistency proof. And Gödel was quite correct in conjecturing the possibility of extending this route to stronger classical theories: Both analysis and a (carefully formulated) set theory can be shown to be consistent relative to their versions with intuitionistic logic.[r]

Gödel reformulates his proof from *1933e*, to show that full intuitionistic logic is not used. He proves (page 7) that for every formula of number theory containing only the conditional and universal quantifica-

[r]For references and brief discussion, see A. S. Troelstra's introductory note to *1933e*, these *Works*, Vol. I, pp. 284–285. To the work relevant to his introductory note to *Gödel 1933f* (these *Works*, Vol. I, pp. 296–298), Troelstra has suggested adding *Flagg 1986*, *Flagg and Friedman 1986* and *Shapiro 1985*.

tion, $\neg\neg A \supset A$ is provable in an intuitionistic system. The details are not entirely clear, but Gödel wants to emphasize that one does not use the principle $\neg A \supset (A \supset B)$. The logical axioms he states are all principles of positive implicational logic, except for the rule of generalization (C, page 7) and axiom B7, which states for elementary formulae, also with free variables, $p \supset q . \supset\subset . \neg p \vee q$.[s] The role of this axiom in Gödel's intended argument is not clear.

The proof of $\neg\neg A \supset A$ proceeds by induction on the construction of A, and the atomic case is simply said to be "clear". Nothing beyond minimal logic is used in the induction step. Thus it would be for the atomic case that B7 is used.[t] The likely interpretation of Gödel's intention is that he assumes (as part of finitist arithmetic, in line with (**R3**)) some of classical truth-functional logic applied to quantifier-free formulae. One might reason as follows for atomic A: $\neg A \vee A$ follows from B5 and B7. Then we can reason by dilemma: Assuming $\neg A$, we infer $\neg\neg\neg A$ and therefore $\neg\neg\neg A \vee A$; assuming A, we infer $\neg\neg\neg A \vee A$; by disjunction-elimination we have $\neg\neg\neg A \vee A$ and by B7 $\neg\neg A \supset A$. Thus B7 is the only assumption used beyond minimal logic, but this argument does use the introduction and elimination rules for disjunction not mentioned by Gödel.[u]

In fact minimal logic is sufficient, and this use of B7 redundant, given that $\neg A$ has been defined as $A \supset 0 = 1$. For $0 = 1 \supset A$ is derivable (in fact for all A) by induction on the construction of A. In the atomic case, where A is an equation $s = t$, we use primitive recursion to define a function ϕ such that $\phi(0) = s$ and $\phi(Sx) = t$, so that $\phi(1) = t$, and then $0 = 1 \supset s = t$ follows.[v] It is doubtful that Gödel had this argument in mind, since if so there would be no reason for the presence of B7.

[s]We use Gödel's symbols for connectives, including the unusual '$\supset\subset$' for the biconditional.

[t]If one sets out to prove $\neg\neg(x = y) \supset x = y$ in intuitionistic arithmetic, one will normally use $\neg A \supset (A \supset B)$ or some equivalent logical principle, and this seems to be essential. Consider $\neg\neg(Sx = 0) \supset Sx = 0$. This could not be proved by minimal logic from the axioms of intuitionistic arithmetic, because minimal logic is sound if one interprets the connectives other than negation classically, and $\neg A$ as true for any truth-value of A. But on that interpretation $\neg\neg(Sx = 0) \supset Sx = 0$ is false. But note that the equivalence of $\neg A$ and $A \supset 0 = 1$ fails on this interpretation.

[u]Why does Gödel assume B7 instead of directly assuming $\neg\neg A \supset A$ for atomic A? Possibly he thought it more evident when $\neg A$ is defined as $A \supset 0 = 1$.

Gödel does remark that conjunction and disjunction are definable from negation and the conditional (p. 6), but the context indicates that this would be after classical logic has been derived for the restricted language. However, if in the atomic case we define $p \vee q$ as $\neg p \supset q$, then B7 reduces to $p \supset q \supset\subset \neg\neg p \supset q$, which yields $\neg\neg p \supset p$ by putting p for q and applying axiom 5. Possibly that is what Gödel had in mind by calling the atomic case "clear".

[v]We owe this observation to A. S. Troelstra.

If A contains only negation, the conditional and universal quantification, and if every atomic formula is negated, then the provability of $\neg\neg A \supset A$ in minimal logic follows from theorem 3.5 of *Troelstra and van Dalen 1988*.[w] Thus for a formula satisfying Gödel's conditions, this will be true for the formula A^g obtained by replacing each atomic subformula P by $\neg\neg P$. Since $P \supset \neg\neg P$ is provable in minimal logic, the above reasoning using axiom B7 enables us to prove $P \leftrightarrow \neg\neg P$ for atomic P, whence minimal logic suffices to prove $A^g \leftrightarrow A$. Gödel remarks further that this proof does not require any nesting of applications of the conditional to universally quantified statements. What seems to be the case is that it is not needed essentially more than is directly involved in the construction of the formula A itself.

Gödel comments that the proof goes so easily because "*Heyting's system violates all essential requirements on* constructivity" (page 8). This is an example of a somewhat disparaging attitude toward intuitionistic methods, at least as explained by Heyting and presumably Brouwer, when applied to the task at hand. But it is clear that Gödel's requirement (**R2**) is violated, as he asserts at the beginning of the section. He claims that requirement (**R3**) is violated because "certain propositions are introduced as evident" (page 4). Probably he has in mind the logical axioms applied to formulae with quantifiers, and possibly also induction applied to such formulae, since these are the non-finitary axioms of intuitionistic arithmetic. But he does not elaborate. We can understand why he thought requirement (**R4**) violated by turning to **1933o*. There he argues that intuitionistic mathematics uses methods going beyond the system A, and says (page 30), when commenting on the logical principle "$p \supset \neg\neg p$":

> So Heyting's axioms concerning absurdity and similar notions differ from the system A only by the fact that the substrate on which the constructions are carried out are proofs instead of numbers or other enumerable sets of mathematical objects. But by this very fact they do violate the principle, which I stated before, that the word "any" can be applied only to those totalities for which we have a finite procedure for generating all their elements.

The idea that intuitionistic mathematics has proofs as basic objects is central to his later analysis in *1958* of the distinction between finitist and intuitionistic mathematics.

[w] *Troelstra and van Dalen 1988*, pp. 62–68, gives a general treatment of provability by minimal logic of "negative translations", based on *Leivant 1985*.

In the present text Gödel considers an interpretation of intuitionistic logic starting from the idea that the conditional is to be understood in terms of absolute derivability, i.e., provability by arbitrary correct means, not limited to the resources of a single formal system. It is not clear how much Gödel was influenced by Heyting's early formulations of the "BHK-interpretation" of logical constants, which in *1930b* and *1931* are very sketchy and incomplete.[x] In his writings of the 1930's, Gödel does not comment on Heyting's conception of a mathematical proposition as expressing an "expectation" or "intention" whose fulfillment is given by the proof of the proposition.[y] Gödel remarks only that he used the notion of derivability to interpret intuitionistic logic already in *1933f*; but there he does not analyze it further, and here he states that the earlier work did not put any weight on constructivity. Intuitionistic propositional logic was interpreted in the result of adding the operator **B** to classical propositional logic (in effect, in a version of the modal logic S4).

In order to obtain a constructive system, Gödel proposes replacing the provability predicate **B** with a three-place relation $z\mathbf{B}p, q$, meaning "z is a derivation of q from p", which he says can "with enough good will" be regarded as decidable. He then formulates some axioms, but in a confusing way, since he sometimes uses **B** as two-place. The axioms as Gödel writes them are as follows:

(1) $z\mathbf{B}p,q \ \&\ u\mathbf{B}q,r \rightarrow f(z,u)\mathbf{B}p,r$

(2) $z\mathbf{B}\varphi(x,y) \rightarrow \varphi(x,y)$

(3) $u\mathbf{B}v \rightarrow u'\mathbf{B}(u\mathbf{B}v)$

He suggests further a rule of inference:

(4) If q has been derived with proof $\underset{\smile}{a}$, infer $a\mathbf{B}q$.

[x] *Heyting 1930* and *1930a* present his intuitionistic formal systems without discussing questions of interpretation at all. But *Heyting 1931* at least was surely known to Gödel before his work on intuitionistic logic and arithmetic. In *1933f*, fn. 1, Gödel does refer to *Kolmogorov 1932*, which presents the interpretation of intuitionistic logic as a calculus of "problems"; cf. Troelstra's comment, these *Works*, Vol. I, p. 299.

[y] *Heyting 1931*, p. 113; cf. *1930b*, pp. 958–959. In neither of these texts does Heyting directly give an explanation of the conditional, but see *1934*, p. 14, which appeared after the remarks in **1933o* but before the present text. In fact Gödel saw earlier drafts of much of *Heyting 1934*, which was the result of Heyting's work on a survey of the foundations of mathematics that was to be written jointly with Gödel. Gödel, however, never finished his part, which was to include a discussion of logicism. Heyting sent him a version of his section on intuitionism with a letter of 27 August 1932.

It seems reasonable to conjecture that this is to be a system based on classical propositional logic, with the presupposition that formulae are decidable. In particular, the $\varphi(x,y)$ of (2) is evidently not an arbitrary formula of, say, intuitionistic arithmetic, but presumably one constructed by finitistically admissible means plus $\mathbf{B}$. Gödel does not make clear how the possible second (or second and third) arguments of $\mathbf{B}$ are to be constructed, although it is clear that they can contain logical operators not admissible elsewhere, such as universal quantifiers. Let us suppose that the language contains the symbol $\top$ for an unanalyzed tautology or other trivial truth. Then we read the two-place "$z\mathbf{B}q$" as an abbreviation for "$z\mathbf{B}\ \top, q$".[z]

Gödel asserts that these axioms are sufficient to prove, for some a,

(5) $a\mathbf{B}[(u){\sim}u\mathbf{B}(0=1)]$.

Clearly from (2) we have

(6) $u\mathbf{B}(0=1) \rightarrow 0=1$,

and, assuming enough arithmetic to prove ${\sim}(0=1)$,

(7) ${\sim}u\mathbf{B}(0=1)$.

Evidently the step to (5) is to use rule (4), but Gödel gives no indication how the universal quantifier is to be introduced. Gödel may well have understood the rule, in application to a case like this where a formula with free variable has been derived, as introducing a symbol a for the general proof; in that case there would only be a notational difference between "$a\mathbf{B}[{\sim}u\mathbf{B}(0=1)]$" and "$a\mathbf{B}[(u){\sim}u\mathbf{B}(0=1)]$" as conclusion.[aa]

Gödel inquires whether the system he has sketched is constructive in the sense he has explained. His answer (page 9) is that the violation of requirement (**R2**) is avoided, since to the right of $\mathbf{B}$, where "forbidden" logical operations occur, the formula occurs "in quotes". But requirements (**R3**) and (**R4**) are still not satisfied. This defect might be removed, if one interpreted $\mathbf{B}$ as referring to proofs of the system itself. In the usual proof of Gödel's second incompleteness theorem, it is shown that formulae corresponding to axioms (1) and (3), and a version of rule (4), are derivable in the system. But then of course the conclusion has to be drawn that (2) is *not* derivable, and just for the case $0=1$ that occurs in the present argument. In view of the remark, "Essentially not

[z]A notation in *Konzept*, p. 7, indicates that Gödel thought of the two-place "$z\mathbf{B}p$" as an abbreviation for "$z\mathbf{B}Ax, p$", but he gives no explanation of what axioms the expression "Ax" refers to. In that same place (3) is stated as "$u\mathbf{B}v \rightarrow g(u)\mathbf{B}(u\mathbf{B}v)$", which makes clear that a primitive function giving the proof of $u\mathbf{B}v$ in terms of the given one of v is being assumed.

[aa]P. 7 of *Konzept* contains other formulae in which the universal quantifier occurs, some of them crossed out, but no suggestion as to how an introduction such as that we have discussed is to go.

the underlined—that is essentially the consistency of the system",[bb] it seems that Gödel saw this; so it is not clear what he thought was the value of his suggestion. By the "introduction of types" he achieves (now using $\mathbf{B}$ for "provable") that only $\mathbf{B}_{n+1} \sim \mathbf{B}_n(0=1)$ holds; this seems to be the natural direction in which the interpretation of $\mathbf{B}$ as referring to formal proofs would go.[cc]

Gödel's remarks present in a very sketchy way an idea that was pursued in subsequent work by G. Kreisel and others (apparently without knowledge of Gödel's earlier discussion), of developing formal theories of constructions and proofs, with a basic predicate like Gödel's $z\mathbf{B}p$. Gödel seems not to attack systematically the question that arises already at the beginning of this work in *Kreisel 1962a* of interpreting intuitionistic logic by giving a definition of "$z\mathbf{B}A$", where A is an arbitrary formula of first-order logic or of some intuitionistic theory. The idea that one should use the clauses of the BHK-interpretation of the intuitionistic connectives to give an inductive definition of $z\mathbf{B}A$ is a very natural one and may well have occurred to Gödel at this time. Without it, it is hard to see how a theory on the lines Gödel sketches could serve the purpose that seems to be intended for it of being a vehicle for consistency proofs. Gödel, however, seems not concerned to develop the idea very far for this purpose, but rather to exhibit where it falls short of meeting his constraints.

Kreisel in *1962a* and *1965* treated the proof relation as decidable, in agreement with Gödel's suggestion. This led to complications in the inductive definition of $z\mathbf{B}A$ for formulae of intuitionistic logic. The obvious clause for $z\mathbf{B}(A \to C)$ would be $\forall u[u\mathbf{B}A \to z(u)\mathbf{B}C]$ (thinking of z as a function), and this appears not to be decidable. Kreisel thus altered the definition to:

(*) $z\mathbf{B}(A \to C)$ iff z is a pair $\langle z_1, z_2\rangle$ and $z_2\mathbf{B}\forall u[u\mathbf{B}A \to z_1(u)\mathbf{B}C]$.[dd]

The same needed to be done for universal quantification. With an axiom like Gödel's (2), (*) and $z\mathbf{B}(A \to C)$ imply $u\mathbf{B}A \to z_1(u)\mathbf{B}C$. Thus Kreisel's definiens implies the above-mentioned obvious one.

There were considerable difficulties in developing a theory along these lines. Kreisel's ideas were naturally developed in the framework of the type-free lambda-calculus, but then the resulting theory is inconsistent;

[bb] "The underlined" seems to be the formula expressing consistency; see editorial note y to the text below.

[cc] Possibly $\mathbf{B}_n$ is intended to mean provability in S_n; see §4 above.

[dd] *Kreisel 1962a*, p. 205; *1965*, p. 128.

see *Goodman 1970*, §9.[ee] Development of a theory along these lines modified so as to avoid paradox was never carried much beyond the interpretation of arithmetic; see *Goodman 1970* and *1973*. A lucid treatment of the basic issues is *Weinstein 1983*. An alternative approach to a theory of constructions involved abandoning the requirement of the decidability of **B**. This led to various typed theories of which that of Per Martin-Löf is best known; see for example *Martin-Löf 1975*, *1984* and the accounts in *Beeson 1985* and *Troelstra and van Dalen 1988*. *Sundholm 1983* criticizes from this point of view the motivation of Kreisel's approach, and thus indirectly Gödel's suggestion of decidability.

6. Transfinite induction and recursion.

Gödel mentions repeatedly that the modal-logical route we just discussed was the heuristic viewpoint guiding Gentzen's 1936 consistency proof for classical arithmetic. However, the characteristic principle by means of which Gentzen went beyond finitist mathematics (if that is thought of as being formalized in the system S_1) is the principle of transfinite induction for both proofs and definitions of functions. And it is to the number-theoretic formulation of these principles that Gödel turns immediately—a task, incidentally, that was not taken up explicitly by Gentzen. First of all it has to be clarified how to grasp (in a mathematically expressible way) specific countable ordinals α from a finitist standpoint. That can be achieved in the system S_1 with function parameters, by considering definable linear orderings $\prec$ of the natural numbers such that, for a definable functional Φ and for all functions f, S_1 proves:

$$\sim \{f(\Phi(f)+1) \prec f(\Phi(f))\};$$

i.e., $\prec$ is provably well-founded. Such an ordering is said to represent α if it is order-isomorphic to α. Two points should be noted. First, there is obviously no function quantification in S_1; universal statements concerning functions (as well as numbers) are expressed just using free-variable statements. Second, the connection between α and $\prec$ is not formulated within finitist mathematics (and is also not needed for the further systematic considerations).

[ee]Although Goodman's statement (p. 109) is more cautious, his argument seems to prove the inconsistency of the "starred theory" of *Kreisel 1962a* as it stands. *Kreisel 1965* envisages the typed λ-calculus, but then it is not clear what type can be assigned to z_2 in (*). Goodman's own solution involves a stratification of constructions into levels; see *1970*, §§10, 13, and the criticism in *Weinstein 1983*, pp. 265–266.

Gödel observes[ff] that S_1 proves the transfinite induction principles for such $\prec$. The *principle of proof* by transfinite induction is formulated as follows: If one can prove $E(a)$ from the assumption $(x)(x \prec a \rightarrow E(x))$, then one is allowed to infer $(x)E(x)$. This inference is represented by the rule

$$\frac{(x)(x \prec a \rightarrow E(x)) \rightarrow E(a)}{(x)E(x).}$$

The corresponding *principle of definition* by transfinite induction (what would nowadays be called a schema of transfinite *recursion*) for $\prec$ is formulated in this way:[gg] if g_i, $1 \leq i \leq n$, are functions with $g_i(x) \prec x$ when x is not minimal, and A is any term (in the language of a definitional extension of S_1), then there is a unique solution to the functional equation

$$\varphi(x) = A(\varphi(g_1(x)), \ldots, \varphi(g_n(x))).$$

That can be done for (orderings representing) ordinals like $\omega + \omega$, ω^2, ω^3, etc. Gödel points out that ω^ω is a precise limit for S_1, since S_1 proves the proof principle of transfinite induction for quantifier-free formulae up to any ordinal $\alpha < \omega^\omega$, but not up to ω^ω. It seems, amazingly, that this result was rediscovered and established in detail only more than twenty years later in *Church 1960* and *Guard 1961*.

In current terminology ω^ω is the *proof-theoretic ordinal* of S_1. The analysis of formal theories in terms of their proof-theoretic ordinals has been a major topic in proof theory ever since Gentzen's consistency proof and his subsequent analysis of the (un-)provability of transfinite induction in number theory. Returning to the present text, Gödel asserts that what can be done in the system S_1, namely, prove induction inferences from other axioms, can also be done in number theory for even larger ordinals. And, as in the case of S_1, there are ordinals for which this cannot be done in number theory. One such ordinal is the first epsilon-number ϵ_0. That ordinal is definable as the limit of the sequence

$$\alpha_1 = 2^{\omega+1} \text{ and } \alpha_{n+1} = 2^{\alpha_n},$$

where exponentiation[hh] is given by

[ff]Treatments of transfinite induction in PRA (i.e., S_1 without function parameters) are given in *Kreisel 1959c* and *Rose 1984*, pp. 165 ff.

[gg]A systematic treatment of so-called ordinal recursive functions, with references to the literature, in particular to the work of Péter, Kreisel and Tait, is found in *Rose 1984*.

[hh]Cf. editorial note bb to the text.

$$2^1 = 2, 2^\beta = \sum_{x<\beta} 2^x.$$

I.e., we associate with each element the sum of the preceding ones; Gödel views this as a "very intuitive construction procedure".[ii] ϵ_0 is obtained by countably iterating the transition from α to 2^α, and if this transition were given, Gödel states in section VI, 9, ϵ_0 would be given. How could one obtain this transition formally in arithmetic? One would assume that α is already represented in the sense above; then one would have to *define* an ordering that represents 2^α and *prove* that that ordering is a well-founded relation (using a suitable functional). Clearly, one may use both the proof and definition principle of transfinite induction for α.

In section VI, 6, Gödel gives a straightforward constructive *definition* of an ordering $\prec$ representing ϵ_0 and explains in the next subsection the obstacle against carrying out the *usual proof* of its well-foundedness.[jj] The proof proceeds by transfinite induction and uses the impredicative induction property "being an ordinal". This property can't be formulated in the language of arithmetic since it requires genuine universal quantification over functions (to be used in the induction schema). Gentzen's consistency proof together with Gödel's second incompleteness theorem is needed to show more than the failure of the usual argument, namely, that there is *no* proof in number theory at all. Gödel does not remark, as he did for ω^ω and S_1, that ϵ_0 is the proof-theoretic ordinal of arithmetic. That for any arithmetic statement transfinite induction up to any ordinal less than ϵ_0 is indeed provable in arithmetic is shown in *Hilbert and Bernays 1939* (page 366); it is also shown in *Gentzen 1943*, where the unprovability of transfinite induction up to ϵ_0 is established without appeal to the second incompleteness theorem. For a modern and very beautiful presentation of these mathematical considerations (incorporating advances due mostly to Schütte and Tait) see *Schwichtenberg 1977*.

In spite of the fact that the transition from α to 2^α (and thus ϵ_0) is not given in arithmetic, Gödel emphasizes (page 12) with respect to the epistemological side:

> ... one will not deny a high degree of intuitiveness to the inference by induction on ϵ_0 thus defined, as in general to the procedure of *defining an ordinal by induction on ordinals* (even though this is an impredicative procedure).

[ii]He emphasized that also in *Konzept*, p. 10, where he added "... und eine Reduktion darauf erscheint mir als wertvoll" ("... and a reduction ⟦to that construction⟧ seems to me to be valuable").

[jj]This (usual) argument is carefully presented in Supplement V of *Hilbert and Bernays 1970*, pp. 534–5.

Consequently, it is natural for Gödel to consider the system obtained from S_1 by adding the principle of transfinite induction up to ϵ_0 (for quantifier-free E). He asks as the first important, philosophical question, whether this theory is still constructive.[kk] Indeed, all requirements are satisfied except for (**R3**); that condition demands the *exclusive* use of ordinary induction. But in a sense transfinite induction is just a generalization of ordinary induction, and Gödel thinks that therefore "... the deviation from the requirement 3 ⟦our (**R3**)⟧ is perhaps not such a drastic one" (page 13). Gödel's reason for accepting induction up to ϵ_0 is not the special combinatorial character of ϵ_0, but rather the fact that ϵ_0 falls into a broader class of ordinals *definable by recursion on already defined ordinals.* And to this procedure, though impredicative, Gödel "will not deny a high degree of intuitiveness". Such broadened views of "constructivity" underlie developments in proof theory described, e.g., in *Feferman 1981* and *1988a*, but also in *Feferman and Sieg 1981*, where the use of generalized inductive definitions is emphasized. As to the use of large constructive ordinals in current proof theory, we refer to *Buchholz 1986*, *Pohlers 1989* and *Rathjen 1991.*

The second important, mathematical question is: How far does one get with extensions of S_1 by adding the inference rule for induction on ordinals that are obtained by ordinal recursive procedures along already obtained ordinals? As to ϵ_0, Gentzen did prove the consistency of number theory and, Gödel adds, "probably also of Weyl's *Kontinuum*".[ll] As the precise theory underlying the development of analysis in *Weyl 1918* is (in one interpretation) a conservative extension of number theory, Gödel is indeed right; cf. *Feferman 1988.* He even thinks that with sufficiently large ordinals one can establish the consistency of analysis and of parts of set theory, as Gentzen had hoped. But he doubts that ordinals satisfying his principle of definability will be sufficiently large. Gödel conjectures at the end of subsection 13 and in 14 that the addition of transfinite induction for such ordinals may not lead to stronger theories than the S_i, and that transfinite induction up to ϵ_0 is already provable in S_ω (see §4 above).[mm]

[kk]A detailed analysis of such quantifier-free theories is given in *Rose 1984*, chapters 6 and 7.

[ll]Gödel presumably meant that the method of Gentzen's proof would yield a proof of the consistency of Weyl's *Kontinuum*, not that Gentzen had literally proved this.

[mm]In *Tait 1968* it is shown that the **T** of *Gödel 1958* is closed under recursion on standard orderings of type less than ϵ_0; from *Kreisel 1959c* it then follows that induction on these orderings is also derivable. This would confirm Gödel's conjecture for a suitable formulation of S_ω (as interpreted in §4 above).

7. Interpreting Gentzen's consistency proof.

Subsections 16 through 19 are a remarkable *tour de force*: on less than two pages Gödel analyzes, with a surprising twist, the essence of Gentzen's consistency proof for classical arithmetic and indicates precisely where in the proof the ordinal exponentiation step occurs that forces the use of all ordinals below ϵ_0. As we mentioned above, Gödel repeatedly points out that the modal-logical route, as a way of assigning a finitist meaning to transfinite statements, was the heuristic viewpoint guiding Gentzen's proof. The latter point was made already in *Konzept*[nn] and is in complete accord with Gentzen's intentions; in section 13 (page 536) of his *1936* (see also *Gentzen 1969*, page 173) Gentzen writes:

> The concept of the "*statability of a reduction rule* (*Reduziervorschrift*)" for a sequent, to be defined below, serves as a formal substitute for the contentual *concept of correctness*; it gives a special *finitist interpretation* of statements, which replaces the *in-itself conception* of them.[oo]

At the very end of his paper (page 564; *1969*, page 201) Gentzen points back to the definition of *Reduziervorschrift* in §13 and claims that the *most crucial part* of his consistency proof consists in providing a finitist meaning to the theorems of classical arithmetic:

> For every arbitrary statement, so long as it has been proved, a *reduction rule* according to 13.6 *can be stated*, and this fact represents the finitist sense of the statement in question, which is gained precisely through the consistency proof.

Gödel claims that the finished proof is only remotely connected to the modal-logical one and maintains that Gentzen proves of each theorem a double-negation translation "... in a different sense from the modal-logical". Gödel formulates the different sense in a mathematically and conceptually perspicuous way: it turns out to be the sense provided by the "no-counterexample interpretation" introduced by Kreisel in *1951*!

[nn]In *Konzept*, page iii, we read: "Die transfiniten Aussagen erhalten einen finiten Sinn." (I.e., the transfinite statements obtain a finitist meaning.)

[oo]In this and in the succeeding quotation, the translation in *Gentzen 1969* is revised. The "in-itself conception" (*an-sich Auffassung*, §9.2) is what we would call realist or Platonist.

These matters are formulated paradigmatically in subsection 16 by considering a formula

(1) $(x)(Ey)(z)(Eu)A(x,y,z,u)$,

in prenex normal form with a decidable matrix A. Proving the negation of this formula constructively means presenting a number c, a unary function f, and a proof of

(2) $\sim A(c,y,f(y),u)$.

A proof of the double negation of (1) consists then in a proof that such a c and f cannot exist: for each f and c one can find functionals $y_{f,c}$ and $u_{f,c}$ such that

(3) $A(c,y_{f,c},f(y_{f,c}),u_{f,c})$;

thus, there cannot be counterexamples c and f.[pp] The functionals y and u are called by Gödel a *reduction*. In subsections 17 through 19 Gödel sketches how to find reductions for theorems in number theory from their formal proofs. (And it is for the treatment of modus ponens that ordinal exponentiation comes in.) Here is not the place to show that Gödel captures the mathematical essence of Gentzen's proof, as that would require a somewhat detailed description of that proof.

Gödel's analysis and presentation are surprising indeed. What accounts partially for the dramatic difference between Gentzen's and Gödel's presentations is the latter's free use of functionals and, to be sure, neglect of all formal details. Functionals do occur in Gentzen's presentation and also in Bernays' description of Gentzen's unpublished consistency proof in *Hilbert and Bernays 1970*, but only in a cautious way to express that the *Reduziervorschriften* are independent of arbitrary choices (see, e.g., pages 536 and 537 in *Gentzen 1936*). The difficult and cumbersome presentation (and what is perceived as an unmotivated manner of associating ordinals to derivations) resulted in a quite general dismissal of Gentzen's first consistency proof; it is the second consistency proof in *Gentzen 1938a* that has been at the center of proof-theoretic research. A widely shared attitude of logicians towards (the "most crucial part" of) Gentzen's first proof can be gleaned from a remark in *Kreisel 1971*. With respect to that most crucial part, i.e., the finitist sense given to logically complex theorems by the *Reduziervorschriften*, Kreisel writes (page 252):

> He ⟦Gentzen⟧ has reservations about his own proposal of expressing this [finitist] sense in terms of the reductions used in

[pp]The reasoning behind the necessary shift of quantifiers is made explicit in **1941*, p. 9.

> his proof because the proposed sense is only "loosely connected" with the form of the theorem considered (and, it might be added, the connection is so tortuous that one couldn't possibly remember it).[qq]

That is particularly striking when contrasted with Gödel's uncovering of the no-counterexample interpretation in his (clearly more sympathetic) reading in late 1937 and early 1938.

8. Concluding remarks.

Gödel commented to Zilsel, as reported in his notes on the organizational meeting, that Gentzen's result is of only mathematical interest, "... ist nur mathematisch interessant"; and this judgment was not uninformed: Gödel had read (at least the unpublished version of) the consistency proof carefully and had discussed it extensively with Bernays.[rr] On this point he obviously modified his views when preparing the lecture for the Zilsel seminar. When discussing the finitist character of the system obtained from PRA by the principle of transfinite induction, Gödel points out, as the reader may recall, that it violates only (**R3**). He continues: "This ⟦new⟧ inference can be considered as a generalization of ordinary induction, and in this respect the deviation from the requirement 3 is perhaps not such a drastic one." In his concluding section VII Gödel evaluates the epistemological significance of consistency proofs relative to the systems he considered; with respect to Gentzen's proof he states, "one will not be able to deny of Gentzen's proof that it reduces operating with the transfinite E to something more evident (the first ϵ-number)".

Gentzen's consistency proof meets, consequently, the general condition Gödel formulated for a "satisfying" relative consistency proof, namely, that such a proof should reduce to something that is more evident. In comparison with a reduction to the basic finitist system Gödel considers the epistemological significance to be "very much diminished".

[qq]It should be mentioned briefly that Kreisel misrepresents Gentzen's remark on p. 564 (to which he alludes in this quotation): according to Gentzen, the loose connection to the form of the theorem is *not* due to the reductions, but rather to the initial (standard) double negation translation, so that, for example, an existential statement does not have its strong finitist meaning, but only the weaker one of its translation.

[rr]As to this episode, see *Kreisel 1987*, pp. 173–174.

But then we have to realize that Gödel expressed in this lecture a rather high regard for Hilbert's original program; if that could have been carried out, "that would have been without any doubt of enormous epistemological value". When comparing Gödel's philosophical remarks with those of Bernays (e.g., in *1935*) or with the reflective considerations of Gentzen (e.g., in *1936* and *1938*) one still finds a marked affinity of their general views.[ss] It is the absolutely unencumbered mathematical analysis that most distinguishes Gödel's presentation from theirs.

Wilfried Sieg and Charles Parsons[tt]

The text of **1938a* was transcribed from Gabelsberger shorthand by Cheryl Dawson and edited by Cheryl Dawson, Charles Parsons and Wilfried Sieg. The translation is by Charles Parsons, revised using suggestions of John and Cheryl Dawson and Wilfried Sieg.

[ss]Despite Gödel's rather critical remarks concerning Gentzen, e.g., on p. 13, "But here again the drive of Hilbert's pupils to derive something from nothing stands out." It is difficult to justify this remark either narrowly, as applying to Gentzen's paper, or more broadly, as applying, e.g., also to Bernays.

[tt]We are grateful to Robert S. Cohen, Cheryl Dawson, John Dawson, Warren Goldfarb, Nicolas Goodman and especially Solomon Feferman and A. S. Troelstra for information, suggestions, and/or comments.

Vortrag bei Zilsel
(*1938a)

I. Problemstellung
II. Allgemeine Charakterisierung der finiten Systeme
III. Finite Zahlentheorie: Tragweite und Erweiterung im allgemeinen
IV. Höhere Typen: Tragweite und Erweiterung
V. Modalitätslogik: Tragweite und Erweiterung
VI. ϵ_o Zahl: Tragweite und Erweiterung
VII. Konklusion

I

⟦Von Widerspruchsfreiheit kann man nur reden mit bezug auf Teilsysteme der Mathematik. Deswegen ist die Frage aber nicht weniger wichtig, weil es Systeme gibt, die so umfassend sind.—Es handelt sich um eine mathematische Frage: jeder Widerspruchsfreiheitsbeweis ist selbst mathematisch und ist daher in bestimmten mathematischen Systemen mit bestimmten Schlußweisen durchführbar.⟧[a] Ich glaube, man muß bei dieser Frage zunächst feststellen:

1. Der Widerspruchsfreiheitsbeweis hat Sinn nur im Sinne einer Reduktion T ⟦ist⟧ reduzierbar ⟦auf⟧ S:[b]

   1. *Wid* $S \rightarrow$ *Wid* T beweisbar ⟦in⟧ S *Teilsystem* ⟦von T⟧
   2. *Wid* T beweisbar ⟦in⟧ S

2 ist vernünftig. Beispiele: Geometrie, Analysis—Logik, Auswahlaxiom *Teilsystem*!

2. Triviale Bemerkung, aber nicht überflüssig. Hilbert ⟦*1928*, Seite 85⟧:

---

[a]This passage is taken—almost literally—from *Konzept*, p. 1. See the textual notes.

[b]"Reduzierbar" is used in the text in both senses indicated here; cf. I 4 A and B below, but also the argument in III 3. As to the examples, one finds a more detailed statement in *Konzept*, p. 1: "In diesem Sinne kann man sagen, die Widerspruchsfreiheit

Lecture at Zilsel's
(**1938a*)

I

⟦One can only talk of consistency in relation to partial systems of mathematics. The question is no less important for that reason, since there are systems that are so encompassing. We have to do with a mathematical question: every consistency proof is itself mathematical and can therefore be carried out in definite mathematical systems with definite modes of inference.⟧[a] I believe that in treating this question one must first note:
1. The consistency proof is meaningful only in the sense of a reduction: T is reducible to S:[b]

1. $Wid\, S \rightarrow Wid\, T$ provable ⟦in⟧ S (a *subsystem* of T)
2. $Wid\, T$ provable ⟦in⟧ S

2 is reasonable. Examples: geometry, analysis—logic, axiom of choice *subsystem*!
2. A trivial observation ⟦that is⟧ not superfluous. Hilbert says ⟦*1928*, page 85; translation from *van Heijenoort 1967*, page 479⟧:

Widerspruchsfreiheit der Geometrie sei auf die der Analysis, die der Analysis auf die der Logik reduziert (nicht im anderen Sinne)." ["In this sense one can say that the consistency of geometry is reduced to that of analysis, that of analysis to that of logic (not in the other sense)."] We assume that Gödel means by "Logik" here the simple theory of types, so that the sense he has in mind is sense 2; clearly, "Auswahlaxiom, *Teilsystem*" is meant to give an example of reducibility in sense 1. On these matters see also section 3 of the introductory note.

> Schon jetzt möchte ich als Schlußergebnis die Behauptung aussprechen: die Mathematik ist eine voraussetzungslose Wissenschaft. Zu ihrer Begründung brauche ich weder den lieben Gott ... noch die Annahme einer besonderen auf das Prinzip der vollständigen Induktion abgestimmten Fähigkeit unseres Verstandes, wie Poincaré, noch die Brouwersche Urintuition und endlich auch nicht, wie Russell und Whitehead, Axiome der Unendlichkeit [und] Reduzierbarkeit ....

Es wird also so getan, als könne man die Widerspruchsfreiheit aus gar keinen Voraussetzungen ableiten, was nicht einmal dann richtig wäre, wenn
2 | sich das ursprüngliche Hilbertsche Programm durchführen ließe.

3. Die Frage der Widerspruchsfreiheit zerfällt ⟦für die gesamte Mathematik⟧ also in Teilfragen⟦: Kann man das System T mittels des Systems S als widerspruchsfrei erweisen?—Und hier gibt es teils positive, teils negative Antworten.⟧[c] In der Aufzählung der Teilantworten erschöpft sich die mathematische Seite der Frage.

4. Die Frage hat aber auch eine erkenntnistheoretische Seite. *Man will ja einen Widerspruchsfreiheitsbeweis zum Zwecke der besseren Fundierung der Mathematik* (*Tieferlegung der Fundamente*) ⟦führen⟧, und es kann mathematisch sehr interessante Beweise geben, die das nicht leisten (⟦wie etwa⟧ Tarski⟦s für die⟧ Analysis[d]). Befriedigend ⟦ist⟧ ein Beweis ⟦nur⟧, wenn er

> A. *auf einen echten Teil reduziert* ⟦oder⟧
> B. *auf etwas zurückführt, was zwar nicht Teil, aber was evidenter, zuverlässiger, etc. ist, so daß dadurch die Überzeugung gestärkt wird.*

A bedeutet zweifellos einen objektiven Fortschritt (Überflüssigmachung von Voraussetzungen ⟦ist⟧ fast dasselbe wie ⟦ein⟧ Beweis). B ⟦ist⟧ zunächst problematisch, weil subjektiv verschieden, aber *de facto* nicht so schlimm; denn es besteht im allgemeinen Übereinstimmung, daß die konstruktiven Systeme besser sind als die, welche mit dem Existential "es gibt" arbeiten. Und auch historisch handelt es sich darum, die nichtkonstruktive Mathematik auf die konstruktive zurückzuführen.

II

Die Schwierigkeit, diese Frage (ob das geht oder nicht) in positivem oder negativem Sinn zu entscheiden, liegt an der Verschwommenheit des *Begriffs*

---

[c]These sentences, as well as the addition "für die gesamte Mathematik", are taken from *Konzept*, p. 1, #3.

> Already at this time I should like to assert what the final outcome will be: mathematics is a presuppositionless science. To found it I do not need God ... or the assumption of a special faculty of our understanding attuned to the principle of mathematical induction, as does Poincaré, or the primal intuition of Brouwer, or, finally, as do Russell and Whitehead, axioms of infinity ⟦or⟧ reducibility ...

Thus one acts as if consistency could be derived from no presuppositions at all, which would not even be correct if the original Hilbert program could be carried out.

3. The question of consistency ⟦for the whole of mathematics⟧ thus divides into partial questions. ⟦Can one show the system T to be consistent by means of the system S?—And here there are partly positive, partly negative answers.⟧[c] The enumeration of the partial answers exhausts the mathematical side of the question.

4. The question has, however, also an epistemological side. *After all we want a consistency proof for the purpose of a better foundation of mathematics (laying the foundations more securely)*, and there can be mathematically very interesting proofs that do not accomplish that (as, for example, Tarski's for analysis[d]). A proof is only satisfying if it either

> A. *reduces to a proper part* or
> B. *reduces to something which, while not a part, is more evident, reliable, etc., so that one's conviction is thereby strengthened.*

A signifies without doubt an objective step forward (making assumptions superfluous is almost the same as a proof). B is at the outset problematic because subjectively different—but de facto not so bad, since there exists general agreement that constructive systems are better than those that work with the existential "there is". And also historically the task has been to reduce non-constructive to constructive mathematics.

II

The difficulty in deciding this question (whether or not that works) positively or negatively is due to the haziness of the *concept "constructive"*.

---

[d] Gödel refers presumably to truth definitions for languages of finite type, as presented, e.g., in section 4 of *Tarski 1935*, and their use in consistency proofs.

"konstruktiv". Ich beginne mit einer Rahmendefinition,[e] welche wenigstens notwendige, wenn nicht schon hinreichende Bedingungen angibt (⟦die⟧ auch nicht ganz scharf, aber immerhin recht brauchbar zu sein scheinen).
1. Die Grundoperationen und Relationen müssen berechenbar und entscheidbar sein. (D. h., ⟦man benötigt ein⟧ Verfahren; hier ⟦ist⟧ ein Punkt, ⟦der⟧ wohl nicht scharf ⟦ist⟧.[f]) *Daraus folgt die Anwendbarkeit des Aussagenkalküls* ⟦und⟧ insbesondere *rekursive Definitionen.*
3 |2. Einschränkung in der Verwendung von ( ) und (E). Ich formuliere zunächst möglichst scharf. (Die Lesarten ergeben sich dann von selbst.) E soll überhaupt nicht unter den Grundzeichen vorkommen, und die Operationen des Aussagenkalküls sollen auf ( ) Aussagen nicht angewendet werden. Bei "$\neg$" ⟦ist das⟧ klar, ⟦und auch für die Implikation "$\supset$",⟧ sonst ⟦hätte man⟧ aus der Definition

$$\neg(x)F(x) \equiv (x)F(x) \supset 0 = 1.$$

Daraus folgt: ⟦man behandelt⟧ *nur Aussagen mit freien Variablen*, also von der Form: es seien irgendwelche Objekte vorgelegt etc. Ein konstruktives E ⟦ist⟧ dann als Abkürzung einführbar, indem man erlaubt zu schließen:

$$F(a) \ldots (Ex)F(x).$$

$(x)(Ey)A(x,y)$ darf aus $A(x, f(x))$ geschlossen werden, und mit solchen Regeln ⟦ist⟧ das nur dann beweisbar. Da ⟦E⟧ nur Abkürzung, ⟦al⟧so unwesentlicher Bestandteil des Systems ⟦ist⟧, welcher weggelassen werden kann, ⟦wird⟧ im Folgenden davon abgesehen. ⟦*Es*⟧ *bleibt also dabei, daß alle* ⟦*Variablen*⟧ *nur als freie Variablen auftreten.*
3. *Schlußregeln und Axiome*: jedenfalls die ⟦rekursiven⟧ Definitionen, ⟦der⟧ Aussagenkalkül, ⟦die⟧ Einsetzungsregel und vollständige Induktion—aber vielleicht noch weitere.[g] Das lasse ich offen.
4. Objekte sollen überblickbar sein (d. h. abzählbar).[h]

1 ⟦und⟧ 2 unbedingt. 3 ⟦und⟧ 4 problematisch; 4 z. B. wegen ⟦des⟧ Funktionsbegriffs; 3 wegen ⟦der⟧ Rekursion nach transfiniten Ordinalzahlen.

---

[e]The "Rahmendefinition" is provided by 1; that is supported by the text below, e.g., V 1.

[f]The interpretive emendations are based on *Konzept*, section C, p. 3 and section 2, p. 2. It seems that Gödel indicates here (as well as in the corresponding statements of *Konzept*) that he views the notion of computability as "nicht ganz scharf" when considered in the context of constructive mathematics. For a later explicit formulation

I begin with a framework definition,[e] which at least gives necessary if not sufficient conditions (which also seem to be not completely sharp, but nevertheless quite usable).
1. The primitive operations and relations must be computable and decidable (that is, one needs a procedure for computing functions and deciding relations; this is a point that is surely not sharp[f]). *From this follows the applicability of the propositional calculus* and in particular *recursive definitions.*
2. Restriction in the application of ( ) and (E). I formulate as sharply as possible at first. (Variant readings then follow of themselves.) E should not occur at all among the primitive signs, and the operations of the propositional calculus should not be applied to ( ) statements. In the case of "$\neg$" that is clear, ⟦and also for implication "$\supset$";⟧ otherwise one would have from the definition

$$\neg(x)F(x) \equiv (x)F(x) \supset 0 = 1.$$

From this follows: we allow *only statements with free variables*, thus of the form: Let certain arbitrary objects be given, etc. A constructive E can then be introduced as an abbreviation by allowing the inference:

$$F(a) \dots (Ex)F(x).$$

$(x)(Ey)A(x, y)$ may be inferred from $A(x, f(x))$, and is only provable with such rules ⟦i.e., only in that way⟧. Since E is only an abbreviation, hence an inessential part of the system, which can be left out, it will be disregarded in the sequel. *Thus all variables continue to occur only as free variables.*
3. *Rules of inference and axioms*: in any case ⟦recursive⟧ definitions, the propositional calculus, the rule of substitution, and ordinary complete induction—but perhaps still further ⟦axioms and rules⟧.[g] I leave that open.
4. Objects should be surveyable (that is, denumerable).[h]

1 and 2 definitely. 3 and 4 ⟦are⟧ problematic, 4 for example because of the concept of function, 3 because of recursion on transfinite ordinal numbers.

---

of his concerns with the correctness of Church's Thesis for intuitionistic computability, cf. *van Heijenoort 1985*, pp. 114–116, and also *Wang 1974*, pp. 81–99, in particular, pp. 87–89.

[g]The addition "rekursiven", in the sense of "primitive recursive", seems to be implied by the first comment below (section III 1).

[h]In *Konzept*, p. 2, #1, Gödel formulated: "Die Gesamtheit der Objekte (Individuen, Relationen, Funktionen), von denen im System die Rede ist, muß überblickbar sein (abzählbar sein)." ["The totality of objects (individuals, relations, functions) that are spoken of in the system must be surveyable (denumerable)."] There are complementary remarks in *Gödel *1933o*, pp. 8 and 10; but cf. also sections 2 and 3 of the introductory note.

III

1. Die finiten Systeme bilden eine Hierarchie. ⟦*Zu*⟧*unterst* ⟦*ist die*⟧ *finite Zahlentheorie*; die Objekte sind Zahlen und durch die gewöhnliche Rekursion definierte Funktionen und solche ⟦Funktionen⟧, die daraus durch Einsetzung entstehen. Schon die *Brouwersche Mathematik* geht darüber erheblich hinaus. Aber ich glaube, *Hilbert wollte mit dieser den Beweis führen.*[i]

2. *Vielleicht sogar ein noch* schwächeres, denn ⟦die⟧ *vollständige Induktion* ⟦wurde⟧ erst ⟦durch⟧ von Neumann ⟦eingeführt⟧ (Hilbert abgeschwächt),
4 | außerdem *Bernays* elementar kombinatorisch und Hilbert selbst.[j]

> 1. Einer der Gesichtspunkte für den Widerspruchsfreiheitsbeweis: Es ist nötig, durchweg dieselbe Sicherheit des Schließens herzustellen, wie sie in der gewöhnlichen niederen Zahlentheorie vorhanden ist ⟦*Hilbert 1926*, Seite 170⟧.
> 2. Es ist möglich, auf rein anschauliche und finite Weise—gerade wie die Wahrheiten der Zahlentheorie—auch diejenigen Einsichten zu gewinnen, die die Zuverlässigkeit des mathematischen Apparates gewährleisten ⟦ib., Seite 171⟧.

Schließlich *zugegeben*, daß eine Erweiterung des Standpunkts nötig ist.

3. *Wie weit kommt man mit der finiten Zahlentheorie, bzw. wie weit kommt man nicht?*

Negativ: A ⟦ist⟧ nicht reduzierbar auf ein Teilsystem T, dessen Widerspruchsfreiheit in A beweisbar ist.[k] ⟦Aus diesen Voraussetzungen hat man sowohl, daß⟧

$Wid(T)$ bew⟦eisbar in⟧ A⟦, als auch daß⟧
$Wid(T) \rightarrow Wid(A)$ beweisbar in T, also in A. ⟦Es folgt:⟧
$Wid(A)$ ⟦ist⟧ beweisbar in A.

Daraus folgt schon: die transfinite Arithmetik ⟦ist⟧ nicht mehr (Herbrand) Teil davon ⟦d. h. Teil der finiten Zahlentheorie⟧; außerdem ⟦gilt das für das System von⟧ Russell–Whitehead ohne Reduzierbarkeitsaxiom.[l]

---

[i]It seems that for Gödel "finite Zahlentheorie" coincides with Primitive Recursive Arithmetic in the sense of Hilbert and Bernays. Cf. the system A in *Gödel *1933o*, pp. 23–25. A systematic argument for this identification is given in *Tait 1981*.

[j]The additions to the text are conjectural as to the meaning. They are based on the following historical facts: in the first steps towards his program in the early twenties, Hilbert thought that the consistency problem could be solved using a system without free variables; cf. *Bernays 1967*, p. 500. But it is clearly stated already in *Bernays 1922* that free variables and metamathematical induction are needed.—"von Neumann", we think, refers to *von Neumann 1927*.

III

1. The finitary systems form a hierarchy with *finitary number theory at the lowest level*; the objects are numbers and functions defined by ordinary recursion and functions that arise from them by substitution. *Brouwer's mathematics* already goes considerably beyond it. But I believe that Hilbert *wanted to carry out the proof* ⟦of consistency⟧ *with this*.[i]
2. *Perhaps even a still* weaker ⟦system⟧, for *complete induction* was first ⟦introduced by⟧ von Neumann (Hilbert weakened), moreover *Bernays* elementary combinatorial and Hilbert himself.[j]

> 1. One of the points of view for the consistency proof: It is necessary to make inference everywhere as reliable as it is in ordinary elementary number theory ⟦*Hilbert 1926*, page 170⟧.
> 2. It is possible to obtain in a purely intuitive and finitary way, just as with the truths of number theory, those insights that guarantee the reliability of the mathematical apparatus ⟦*ibid.*, page 171, translations from *van Heijenoort 1967*, page 377, slightly revised⟧.

Finally *granted*, that an extension of the standpoint is necessary:
3. *How far do we get, or fail to get, with finitary number theory?*

Negative: A is not reducible to a subsystem T, whose consistency is provable in A.[k] ⟦Supposing that it is, we have both⟧

$Wid(T)$ provable in A, and
$Wid(T) \rightarrow Wid(A)$ provable in T, therefore in A. It follows that $Wid(A)$ is provable in A.

From this already follows: transfinite arithmetic is no longer (Herbrand) a part ⟦of finitary number theory⟧. Moreover, ⟦that holds for the system of⟧ Russell-Whitehead without the axiom of reducibility.[l]

---

[k] A is taken to be some formal system that contains finitist arithmetic T as a proper part.

[l] Gödel presumably refers to *Herbrand 1931*. There Herbrand entertains the possiblity that all finitist arguments can be carried out already in elementary (Peano) arithmetic and claims: "If this were so, the consistency of ordinary arithmetic would already be unprovable" (p. 8; *Herbrand 1971*, p. 297). Clearly, Herbrand means in this sentence "unprovable by finitist means".—Gödel's second claim seems to be that the system of *Principia mathematica* without the axiom of reducibility goes beyond finitist mathematics. That is supported by *Konzept* (p. 0.1, #2), where that system is considered to be part of transfinite mathematics.

4. *Wie also erweitern?* (Erweiterung nötig.) Drei Wege ⟦sind⟧ bisher bekannt:

1. Höhere Typen von Funktionen (Funktionen ⟦von⟧ Funktionen von Zahlen, etc.).
2. Modalitätslogischer Weg (Einführung einer Absurdität auf Allsätze angewendet und eines "Folgerns").
3. Transfinite Induktion, d. h., es wird der Schluß durch Induktion für gewisse konkret definierte Ordinalzahlen der zweiten Klasse hinzugefügt.

Bemerkung. Nur die Systeme 1 genügen sämtlichen Forderungen. 2, 3 genügen nicht der Forderung 3 (gewisse Sätze ⟦werden⟧ als evident eingeführt), ⟦Systeme⟧ 2 ⟦genügen⟧ auch nicht der Forderung 4.[m]

IV

5 |1. ⟦Die Erweiterung 1⟧ besteht wesentlich in der Einführung einer Funktionsvariablen $f(x)$, welche nicht nur über die früher definierten Funktionen, sondern auch über die mit ihrer Hilfe zu definierenden läuft, und ⟦im⟧ Definitionsschema für Funktionen ⟦von⟧ Funktionen $\Phi(f, n, k)$ durch Induktion nach n, z. B.

$$\Phi(f, 1, k) = f(k)$$
$$\Phi(f, n+1, k) = f(\Phi(f, n, k)) \text{ Iteration}$$

Beweise ⟦benutzen⟧ nur ⟦den⟧ Aussagenkalkül, ⟦die⟧ Einsetzung⟦sregel⟧ und ⟦die gewöhnliche⟧ Induktion. ⟦Alle⟧ *Forderungen* ⟦sind⟧ *erfüllt.*

2. Für f darf wieder eine mittels Φ definierte Funktion eingesetzt werden, aber ⟦das ist⟧ nicht zirkelhaft, weil Ausdrücke in eine Reihe ⟦ge⟧ordnet werden ⟦können⟧.[n]

3. Fortsetzung für die Funktionen von Funktionen von Funktionen ... und schließlich auch ins Transfinite. Trotzdem ⟦erhält man ein⟧ abgeschlossenes System, wenn verlangt wird, daß nur Typen nach solchen Ordinalzahlen ⟦zugelassen werden⟧, welche in einem früheren System definiert ⟦worden sind⟧.

---

[m]Indeed, Gödel states at the beginning of the next section that systems under 2 do not satisfy the second requirement either.

[n]In *Konzept*, p. 4, Gödel writes more explicitly: "Allgemein ⟦wird⟧ dieses Rekursionsschema zur Einführung von Φ_i ⟦mit Hilfe von⟧ früheren f_i und den durch Einsetzung ineinander gewonnenen Funktionen ⟦benutzt⟧. ... $\Phi(f, n, k) = f(f(f \ldots f(k) \ldots)$ ⟦ist⟧ berechenbar, wenn f ⟦es⟧ ist. Das kann allgemein gezeigt werden, indem man die ⟦definierenden⟧ Ausdrücke in eine bestimmte Reihenfolge bringt und zeigt, daß die

4. *How then shall we extend?* (Extension is necessary.) Three ways are known up to now:

1. Higher types of functions (functions of functions of numbers, etc.).
2. The modal-logical route (introduction of an absurdity applied to universal sentences and a ⟦notion of⟧ "consequence").
3. Transfinite induction, that is, inference by induction is added for certain concretely defined ordinal numbers of the second number class.

Remark. Only the systems 1 satisfy all requirements. 2, 3 do not satisfy requirement 3 (certain propositions are introduced as evident), ⟦the systems⟧ 2 also do not satisfy requirement 4.[m]

IV

1. ⟦Extension 1⟧ consists essentially in the introduction of a function variable $f(x)$, which ranges not only over the functions defined earlier, but also over those to be defined with its help, and in the schema of definition for functionals $\Phi(f, n, k)$ by induction on n, for example

$$\Phi(f, 1, k) = f(k)$$
$$\Phi(f, n+1, k) = f(\Phi(f, n, k)) \text{ Iteration}$$

Proofs ⟦use⟧ only the propositional calculus, ⟦the rule of⟧ substitution, and ⟦ordinary⟧ induction—All *requirements are satisfied.*
2. For f, a function defined by means of Φ can again be substituted, but that is not circular, because expressions ⟦can be⟧ ordered in a series.[n]
3. Continuation for functions of functions of functions ... and ultimately into the transfinite. Nevertheless ⟦one gets a⟧ closed system, if it is demanded that types only ⟦be admitted⟧ for those ordinal numbers which have been defined in an earlier system.

Berechnung der späteren ⟦Funktionen⟧ auf die Berechnung der vorhergehenden zurückgeführt werden kann. Diese Hierarchie kann fortgesetzt werden, indem man Funktionen einführt, deren Argumente ⟦solche⟧ Φ sind, und ⟦indem man⟧ wiederum rekursive Definitionen nach einem Zahlparameter zuläßt. Und das kann sogar ins Transfinite erweitert werden." ("In general this recursion schema ⟦is used⟧ for the introduction of Φ_i ⟦with the help of⟧ earlier f_i and the functions obtained by substituting in one another. $\ldots \Phi(f, n, k) = f(f(f \ldots f(k) \ldots)$ is computable, when f is ⟦computable⟧. That can be shown generally, by arranging the defining expressions in a certain sequence and showing that the computation of the later ⟦functions⟧ can be reduced to the computation of the earlier ones. This hierarchy can be continued by introducing functions whose arguments are such Φ and admitting once again recursive definitions according to a numerical parameter. And that can even be extended into the transfinite.")

4. In diesem Verfahren ⟦sind⟧ enthalten:
   1. Hinzufügung von Rekursionen nach mehreren Variablen,
   2. Hinzufügung der Aussage *Wid*,
   3. Hinzufügung der Hilbertschen Schlußregel.[o]
5. Wie weit kommt man damit?
   ⟦1.⟧ Negativ: Mit endlichen Typen kann man die Zahlentheorie nicht als widerspruchsfrei beweisen.
   2. Vermut⟦ung⟧: daß die Analysis bereits in keinem solchen System ⟦als widerspruchsfrei⟧ beweisbar sein wird.

6 |Das Ganze, ein interessantes offenes Problem (weil diese die einzigen Systeme sind, welche allen vier Forderungen genügen).

Nachweisen an den einzelnen Forderungen.

V

1. Ich komme jetzt zum zweiten Weg, dem modalitätslogischen. Heyting hat den Intuitionismus formalisiert. *Heytings System genügt außer der Rahmendefinition überhaupt nichts*, weil er $\neg$ ⟦und⟧ $\supset$ anwendet auf ( ) Aussagen.
2. *Alle die scheinbar schwächeren Annahmen* (*Satz vom ausgeschlossenen Dritten* ⟦*ist*⟧ *nicht* ⟦*vorausgesetzt*⟧ *und auch nicht allgemein beweisbar*) *konnte* ⟦*man*⟧ *vollkommen ersetzen, indem man beweisen kann*, daß für die aus $\neg$, $\supset$, ( ) und den arithmetischen Grundbegriffen aufgebauten Sätze alle klassisch beweisbaren Sätze gelten (und aus diesen Begriffen sind ja die übrigen &, $\vee$ definierbar).[p] Es gilt also auch der *Satz vom ausgeschlossenen Dritten*, allerdings nicht für das Heytingsche ⟦System⟧, aber für das aus diesem definierte. Aber das ⟦ist⟧ egal. (Man hat ein Modell.)
3. Dieses Resultat ⟦ist⟧ sogar schon aus geringeren Voraussetzungen als den Heytingschen ableitbar [bei Heyting problematisch: $\neg p\,.\supset p \supset q$]. ⟦Es ist⟧ *interessant mit wie geringen Voraussetzungen*, denn man ⟦hat nur⟧ einmal $\supset$ auf ( ) Aussagen angewendet.[q] Also:
   A. Als einziger neuer Grundbegriff zu der finiten Zahlentheorie wird $\supset$ hinzugefügt: $\neg p$ definiert durch $p \supset 0 = 1$ ($p \supset a$).
   B. Über das $\supset$ nur lauter anscheinend einwandfreie Annahmen, nämlich

[o]This is an altogether novel rule of inference formulated in *Hilbert 1931* (for an informative discussion see Feferman's introductory note to *Gödel 1931c*, these *Works*, Vol. I, pp. 208–213): Given a finitistic proof that each instance of the quantifier-free formula $A(x)$ is a correct numerical statement, the universal generalization $(\forall x)A(x)$ can be taken as an axiom of arithmetic. In *Herbrand 1931*, with which Gödel was thoroughly familiar and whose formalism he actually used in his *1933e*, Hilbert's rule is prominently

4. ⟦The following⟧ are contained in this procedure:
   1. Addition of recursion on several variables,
   2. Addition of the statement *Wid*,
   3. Addition of Hilbert's rule of inference.[o]
5. How far does one get with this?
   1. Negative: With finite types one cannot prove the consistency of number theory.
   2. Conjecture: that already the ⟦consistency of⟧ analysis will not be provable in any such system.

All of this ⟦constitutes⟧ an interesting open problem (because these are the only systems which satisfy all four requirements). To show for the individual requirements.

V

1. I now come to the second way—the modal-logical. Heyting has formalized intuitionism. *Heyting's system satisfies nothing at all beyond the framework definition*, because he applies $\neg$ and $\supset$ to ( )-statements.
2. *The apparently weaker assumptions (the law of the excluded middle is not assumed and is also not generally provable) could be totally replaced, in that one can prove* that, for the statements built up from $\neg$, $\supset$, ( ), and the basic arithmetical concepts, all classically provable statements hold (and the remaining ⟦connectives⟧ & and $\vee$ are of course definable from these).[p] Therefore the *law of the excluded middle* also holds, to be sure not for Heyting's ⟦system⟧, but for the system defined from it—but that doesn't matter. (One has a model.)
3. This result is already provable even from fewer presuppositions than Heyting's [in the case of Heyting, $\neg p \,.\supset p \supset q$ is problematic]. It is *interesting with how few presuppositions*, for $\supset$ has been applied to ( )-statements ⟦only⟧ once.[q] Therefore:
   A. $\supset$ is added to finitary number theory as the only new primitive concept: $\neg p$ ⟦is⟧ defined as $p \supset 0 = 1 (p \supset a)$.
   B. About $\supset$, nothing but apparently unexceptionable assumptions, namely,

---

employed. The main theorem of *Herbrand 1931* asserts the consistency of the classical system with Hilbert's rule, but with induction restricted to quantifier-free formulas. For further discussion see the introductory note.

[p]Cf. *Gödel 1933e*, in particular, p. 37.

[q]On this remark see section 5 of the introductory note.

1. Transitivität: $p \supset q \,.\supset: q \supset r \,.\supset. p \supset r$
2. Vertauschung zweier Prämissen: $p \supset. q \supset r :\supset: q \supset. p \supset r$

7 |3. Weglassung einer doppelten Prämisse: $p \supset. p \supset q : \supset. p \supset q$

4. Modus ponendo ponens: $p \supset: p \supset q \,.\supset q$
5. Identität: $p \supset p$
6. Wahrer Satz folgt aus jedem: $q \supset. p \supset q$
7. Für elementare Formeln auch mit freien Variablen: $p \supset q \,.\supset\subset. \neg p \vee q$

⟦C.⟧ Für das ( ) die Schlußregel $A \supset F(x) : A \supset (x)F(x)$.

Lauter scheinbar unverdächtige Schlüsse. Das System *wahrscheinlich abundant* und vielleicht ⟦erhält man⟧ mit & ein noch eleganteres System.[r] Jedenfalls ⟦ist⟧ keine Spur eines Satzes vom ausgeschlossenen Dritten zu bemerken. *Problematisch* ⟦ist⟧ *vielleicht 6*, aber nicht wenn $\supset$ als beweisbar gedeutet ⟦wird⟧ [Łukasiewicz $p \supset q \,.\supset p :\supset p$].[s]

4. Für jeden Satz der Zahlentheorie⟦, der mit $\supset$ und ( ) aufgebaut ist, gilt:⟧ $\neg\neg A \supset$A. ⟦Der Beweis beruht auf dem aussagenlogischen Theorem $(p \supset q) \supset (((p \supset a) \supset a) \supset ((q \supset a) \supset a))$, das mit Hilfe von Axiom 1 leicht zu beweisen ist.⟧[t]

1. Für elementare Formeln klar, auch wenn sie freie Variablen enthalten.
2. Wenn für $F(x)$ bewiesen ⟦ist⟧ $\neg\neg F(x) \supset F(x)$, so auch $\neg\neg(x)F(x) \supset (x)F(x)$:

$$\begin{aligned} (x)F(x) &\supset F(y) \\ \neg\neg(x)F(x) &\supset \neg\neg F(y) \\ &\supset F(y) \\ &\supset (y)F(y), \end{aligned}$$

⟦3.⟧ und ebenso ⟦beweist man die⟧ Übertragung von p, q auf $p \supset q$:

$$\begin{aligned} p \supset: p \supset q. &\supset q \\ p \supset. \neg\neg(p \supset q) &\supset \neg\neg q \\ &\supset q \\ \neg\neg(p \supset q) \,.\supset. p &\supset q \end{aligned}$$

8 |5. *Daß das so leicht geht, liegt eben daran, daß das Heytingsche System alle wesentlichen Forderungen an* Konstruktivität verletzt, aber vielleicht

[r]The system of pure implicational logic Gödel describes here is indeed "abundant"; compare the contemporary presentation in *Hilbert and Bernays 1934*, section 3.c)3. In Supplement III of *Hilbert and Bernays 1939* a system that contains "&" in addition to implication is investigated; see their references to the work of Hertz and Gentzen.

1. Transitivity: $p \supset q \,.\!\supset\, : q \supset r \,.\!\supset . p \supset r$
2. Interchange of two premisses: $p \supset . q \supset r : \supset : q \supset . p \supset r$
3. Leaving out a doubled premiss: $p \supset . p \supset q : \supset . p \supset q$
4. Modus ponendo ponens: $p \supset : p \supset q \,.\!\supset q$
5. Identity: $p \supset p$
6. ⟦A⟧ true proposition follows from every ⟦proposition⟧: $q \supset . p \supset q$
7. For elementary formulae, also with free variables: $p \supset q \,.\!\supset\!\subset . \neg p \vee q$

⟦C.⟧ For ( ) the rule of inference $A \supset F(x) : A \supset (x)Fx$.

Nothing but apparently innocent inferences. The system is *probably abundant* and perhaps with & ⟦one obtains⟧ a still more elegant system.[r] In any case no trace of a law of excluded middle is to be noticed. *6 is perhaps problematic*, but not when $\supset$ is interpreted as provable [Łukasiewicz $p \supset q \,.\!\supset p : \supset p$].[s]

4. For every proposition of number theory ⟦which is built up by $\supset$ and ( )⟧ $\neg\neg A \supset A$ ⟦holds. The proof rests on the theorem of propositional logic $(p \supset q) \supset (((p \supset a) \supset a) \supset ((q \supset a) \supset a))$, which is easy to prove with the help of Axiom 1.⟧[t]

1. For atomic formulae clear, also if they contain free variables.
2. If for $F(x)$, $\neg\neg F(x) \supset F(x)$ has been proved, then also $\neg\neg(x)F(x) \supset (x)F(x)$:

$$\begin{aligned} (x)F(x) &\supset F(y) \\ \neg\neg(x)F(x) &\supset \neg\neg F(y) \\ &\supset F(y) \\ &\supset (y)F(y), \end{aligned}$$

3. And likewise ⟦one proves⟧ the transition from p, q to $p \supset q$:

$$\begin{aligned} p \supset : p \supset q . &\supset q \\ p \supset . \neg\neg(p \supset q) &\supset \neg\neg q \\ &\supset q \\ \neg\neg(p \supset q) \,.\!\supset . p &\supset q \end{aligned}$$

5. *That this goes so easily rests on the fact that Heyting's system violates all essential requirements on* constructivity, but perhaps it is still to be made

---

[s]The reference to Łukasiewicz might be to *Łukasiewicz and Tarski 1930*. In section 4 of that paper the authors discuss a semantically characterized, implicational fragment of sentential logic. They prove a completeness theorem (their Theorem 29) for a calculus whose axioms are Gödel's axioms 1,6 and this formula, i.e., Peirce's law.

[t]Here we follow *Konzept*, p. 6. See the textual notes.

⟦ist es⟧ doch irgendwie konstruktiv ⟦zu⟧ machen.—Weswegen erkennen es die Intuitionisten überhaupt an? Sie denken an eine konstruktive Deutung:

$$p \supset q \text{ bedeutet “}\textit{q ist aus p ableitbar}\text{”},$$

“ableitbar” verstanden nicht in einem bestimmten System, sondern im absoluten Sinn (d. h., man kann es evident machen); und “ableitbar” im konstruktiven Sinn verstanden, wie ich oben E eingeführt habe, d. h. man hat eine Ableitung. Und bei dieser Interpretation ⟦sind⟧ tatsächlich ⟦die⟧ *obigen Axiome plausibel*, aber nicht für “ableitbar” in einem bestimmten System.

6. Auf dieser Idee basierend habe ich eine Interpretation ⟦in *1933f*⟧ mittels B gegeben, ohne auf konstruktive Forderungen Wert zu legen⟦, und⟧ den gewöhnlichen Aussagenkalkül ergänzt⟦e ich⟧ durch B (“ist beweisbar im absoluten Sinn”) und ⟦durch⟧ Axiome 1[u] und 2. $Bp \rightarrow BBp$; 3. $Bp \rightarrow p$; 4. B kann hinzugefügt werden. *Intuitionismus* ⟦*ist*⟧ *daraus ableitbar.*[v]

7. Merkwürdiges Ergebnis, obwohl diese Axiome sämtlich außerordentlich plausibel ⟦sind⟧, ⟦sind⟧ *trotzdem daraus Sätze über B ableitbar, welche sicher für jedes definierte B falsch sind.* Nämlich $B \sim B(0 = 1)$.[w] Auf etwas ähnliches müssen wir auch jetzt gefaßt sein.

8. Damals ⟦hatte ich⟧ keinen Wert darauf gelegt, mittels *des B ein konstruktives System zu erhalten.* ⟦Das⟧ B ist ja nicht konstruktiv, und außerdem ⟦wurde⟧ darauf der gewöhnliche Aussagenkalkül angewendet. *Aber das ist vermeidbar*: ⟦den⟧ Grundbegriff zBp, q, ⟦d. h.,⟧ z ist eine Ableitung von q aus p, ⟦kann man⟧ mit dem nötigen guten Willen als entscheidbar annehmen.

Axiome: z. B. Transitivität der Implikation: $zBp, q \,\&\, uBq, r \rightarrow f(z,u)Bp, r$.

Andere Axiome: $zB\varphi(x,y) \rightarrow \varphi(x,y)$, $uBv \rightarrow u'B(uBv)$; ferner, wenn
9 q bewiesen und a der Beweis, so daß ist anzuschreiben “aBq”, | würde wie oben beweisbar $aB[(u) \sim uB(0 = 1)]$.

9. *Es fragt sich nun, ist dieses System konstruktiv im obigen Sinn?*

A. Es ist *kein Einwand, daß* auf alle Aussagen ja jetzt doch logische Operationen (nämlich das B) angewendet werden, was gerade verboten war, denn die Aussage tritt hier in der *suppositio materialis* als Gegenstand auf, unter Anführungszeichen.

---

[u]By “Axiom 1” Gödel must mean $Bp \rightarrow . B(p \rightarrow q) \rightarrow Bq$. Indeed, in his *1933f*, Gödel used the axioms $Bp \rightarrow p$, $Bp \rightarrow . B(p \rightarrow q) \rightarrow Bq$, and $Bp \rightarrow BBp$; in addition, he had the rule that allows the inference from p to Bp. The “interpretation” alluded to above is given there by the following translation:

$\neg p$	is translated as	$\sim Bp$,
$p \supset q$	as	$Bp \rightarrow Bq$,
$p \vee q$	as	$Bp \vee Bq$
and $p \wedge q$	as	$p \,.\, q$.

somehow constructive. For what reason do the intuitionists recognize it at all? They are thinking of a constructive interpretation:

$$p \supset q \text{ means "}q \textit{ is derivable from } p\text{"},$$

⟦with⟧ "derivable" understood not in a particular system, but in the absolute sense (that is, one can make it evident), and "derivable" understood in the constructive sense, as I introduced E above, that is, one has a derivation. And on this interpretation the *above axioms* are actually *plausible*, but not for "derivable" in a particular system.

6. Relying on this idea, I gave ⟦in *1933f*⟧ an interpretation by means of B, without laying stress on constructive requirements, and I supplemented the usual propositional calculus with B ("is provable in the absolute sense"), and Axioms 1[u] and 2. $Bp \to BBp$; 3. $Bp \to p$; 4. B can be added. *Intuitionism is derivable from this.*[v]

7. A curious result, although these axioms are all extraordinarily plausible: *nevertheless propositions about B are derivable from them which are surely false for every defined B*: namely, $B \sim B(0 = 1)$.[w] Now as well, we have to be prepared for something similar.

8. At that time ⟦I hadn't⟧ put any value on *obtaining a constructive system* by means *of B*. B is not constructive, and moreover the usual propositional calculus was applied to it. *But that can be avoided*: the basic notion zBp, q, that is, z is a derivation of q from p, can be viewed as decidable with enough good will.

Axioms: for example transitivity of implication: $zBp, q \,\&\, uBq, r \to f(z, u)Bp, r$.

Other axioms: $zB\varphi(x, y) \to \varphi(x, y), uBv \to u'B(uBv)$; furthermore, if q has been proved and a is the proof, so that "aBq" is to be written down, then as above $aB[(u){\sim}uB(0 = 1)]$ would be provable.

9. *The question now arises, is this system constructive in the above sense?*
   A. It is *no objection that* logical operations (namely B) are now applied to all statements, just what was forbidden, because the statement occurs here in *suppositio materialis* as an object, in quotation marks.

---

[v]The meaning of the remark "Intuitionismus ist daraus ableitbar" is not entirely clear. One possibility is that he means simply that on the interpretation given, the theorems of intuitionistic propositional logic are derivable. Another suggestion is that he is alluding to the converse of this result obtained by McKinsey and Tarski (*1948*, cf. these *Works*, Volume I, p. 296). Gödel conjectured this result in *1933f*.

[w]In *Konzept* Gödel argues also that these axioms—in spite of their plausibility from an intuitionistic standpoint—are false for each formalized provability predicate. He does this as follows: the first axiom gives $B(p\&\neg p) \supset p\&\neg p$; so we have by logic ${\sim}B(p\&\neg p)$, and thus with the rule $B{\sim}B(p\&\neg p)$. The same argument is indicated at the end of *1933f*, with $0 \neq 0$ replacing $p\&\neg p$.

B. ⟦Sind die⟧ Voraussetzungen erfüllt? So wie es steht, sind 3 und 4 sicher nicht erfüllt [aber sehr hohe Evidenz], denn Beweise ⟦sind⟧ unüberblickbar; und die Grundsätze sind nicht Definitionen und folgen auch nicht aus den oben angeführten Schlußregeln.
Die Voraussetzungen 1 und 2 sind bei nötigem guten Willen erfüllt.

C. Vielleicht aber ist es möglich, das System so zu präzisieren, daß tatsächlich alle Voraussetzungen bis auf 3 erfüllt ⟦sind⟧, indem man sich auf Beweise des Systems selbst beschränkt. Das ⟦ist⟧ vielleicht nicht widerspruchsvoll, denn die äquivalente arithmetische Aussage ⟦ist⟧ nicht beweisbar. Dann ⟦können⟧ die meisten von den Axiomen tatsächlich auf Definitionen zurückgeführt ⟦werden⟧.
Im Wesentlichen nicht das Unterstrichene[x]—Das ist im Wesentlichen die Widerspruchsfreiheit des Systems.

10. Ein anderer Weg, um zu versuchen, ob man aus diesem System etwas Vernünftiges bekommt, ist die Typeneinführung:[y]

$$B_{n+1} \sim B_n(0=1)$$

beweisbar. Aber ⟦solch ein Versuch macht⟧ große Schwierigkeiten, obwohl Gentzen angibt, daß dies der heuristische Gesichtspunkt ⟦für seinen Widerspruchsfreiheitsbeweis sei,⟧ jedenfalls *indem man es ins Transfinite fortsetzt.*

11. Von den drei Wegen ⟦ist⟧ dieser am schlechtesten, sogar vielleicht schlechter als der zu beweisende, daher führt er auch am weitesten (Analysis wahrscheinlich zu bekommen). Aber ⟦er ist⟧ durchaus nicht müßig, sondern vielleicht heuristisch sehr wertvoll. Daher so ausführlich obwohl außerhalb durch Gentzen...[z]

10 |

VI

1. Schon innerhalb *der finiten Zahlentheorie* (oder wenigstens des Systems mit f) ⟦sind⟧ gewisse Ordinalzahlen der zweiten Klasse definierbar, und ⟦es ist⟧ nachweisbar, daß Beweise und Definitionen nach diesen Ordinalzahlen möglich sind, z. B. $\omega + \omega$, und beweisbar, daß Ordinalzahl ...

2. *Was heißt das, finit?* $\Phi(f)$ definierbar, so daß

$$\sim\{f(\Phi(f)+1) \underset{R}{<} f(\Phi(f))\}.$$

[x]All text underlined in the manuscript has here been rendered in italics. We conjecture that "das Unterstrichene" may refer to the last line of 8 above, where in the original text each occurrence of "B" had a dot underneath.

B. Are the presuppositions satisfied? As it stands, 3 and 4 are surely not satisified [but ⟦there is⟧ very high evidence], for proofs are unsurveyable, and the principles are not definitions and also don't follow from the rules of inference cited above.
The presuppositions 1 and 2 are, with enough good will, satisfied.

C. Perhaps, however, it is possible to specify the system in such a way that actually all presuppositions except 3 are satisfied, by limiting oneself to proofs of the system itself. That is perhaps not contradictory, for the equivalent arithmetic statement is not provable. Then most of the axioms ⟦can⟧ actually be reduced to definitions.
Essentially not the underlined[x]—that is essentially the consistency of the system.

10. Another way to understand whether one gets something reasonable from this system is the introduction of types:[y]

$$B_{n+1} \sim B_n(0 = 1)$$

⟦is⟧ provable. But ⟦such an attempt has⟧ great difficulties, although Gentzen claims that this is the heuristic point of view ⟦for his consistency proof⟧, at any rate *if one continues it into the transfinite.*
11. This is the worst of the three ways, even perhaps worse than the one to be proved; thus it also leads furthest (analysis is probably obtainable), but it is definitely not idle but perhaps heuristically very valuable, thus so detailed although except by Gentzen ...[z]

VI

1. Already within *finitary number theory* (or at least the system with f), certain ordinals of the second number class are definable, and it can be shown that proofs and definitions according to these ordinals are possible, for example $\omega + \omega$, and provable, that ordinal number ...
2. *What does this mean,* ⟦*in*⟧ *finitary* ⟦*terms*⟧*?* $\Phi(f)$ ⟦is⟧ definable, so that

$$\sim\{f(\Phi(f)+1) \underset{R}{<} f(\Phi(f))\}.$$

[y] In *Konzept*, p. 8, Gödel wrote more explicitly that B_n means "beweisbar im finiten System n-ten Typs" ("provable in the finitary system of nth type"). As to the latter notion, see section 5 of the introductory note.

[z] What follows "daher" in this sentence is evidently incomplete. It seems Gödel means that what he has just said is a reason for treating the approach more fully, but that only Gentzen has made use of its heuristic value.

Daraus folgt dann:

1. Schluß durch Induktion, d. h. wenn E eine Aussagenfunktion ⟦ist⟧, ⟦und⟧ es ist aus der Annahme $x \underset{R}{<} a \rightarrow E(x)$ ableitbar $E(a)$ im System S_1, dann ist $(x)E(x)$ beweisbar im System S_1;
2. Definition durch Induktion: Wenn $g_1(x), \ldots, g_v(x)$ irgendwelche Funktionen sind, so daß $g_i(x) \underset{R}{<} x$, und man schreibt ⟦für⟧ irgendeinen Ausdruck ⟦A⟧ ein

$$\varphi(x) = A(\varphi(g_1(x)), \ldots, \varphi(g_v(x))),$$

so gibt es ein φ⟦, das diese Gleichung erfüllt⟧.

3. Dasselbe ⟦ist⟧ auch für höhere Ordinalzahlen, z. B. ω^2, ω^ω, ω^{ω^2}, noch möglich. Also ⟦sind⟧ Induktionsschlüsse beweisbar aus den anderen Axiomen. ω^ω *schon nicht mehr in der finiten Zahlentheorie* (*aber alle* ⟦$< \omega^\omega$⟧ *definierbar in der finiten Zahlentheorie*).
4. Aber es gibt gewisse Zahlen, für welche das nicht möglich ⟦ist⟧ zu beweisen und zwar nicht einmal in der *transfiniten Arithmetik.* Eine solche Zahl ist die erste ϵ-Zahl—definiert[aa] durch ⟦$\alpha_1 = 2^{\omega+1}$ und⟧ $\alpha_{n+1} = 2^{\alpha_n}$. *Anschauliches Bild dieser Zahl.* Zunächst ⟦betrachte ich den⟧ Prozeß 2^α:

$$2^1 = 2, \quad 2^\beta = \sum_{x<\beta} 2^x,$$

11 | d. h., jedem Element *wird die Summe der vorhergehenden zugeordnet*; ⟦das ist ein⟧ sehr anschauliches Bildungsverfahren.[bb]

5. Schwierigkeit: *Warum das nicht in der Arithmetik formalisiert* werden kann und auch nicht in der durch endliche finite Systeme ergänzten? Angenommen es sei schon bewiesen, α ist eine Ordinalzahl in dem schärfsten Sinn, d. h., ein $\Phi(f)$ ⟦ist⟧ definiert, so daß ⟦jede Teilfolge⟧ abbricht.—Wir wollen versuchen, zunächst eine *Anordnung von* natürlichen Zahlen in der Arithmetik zu definieren, welche 2^α darstellt, und dann zu beweisen, daß diese Anordnung eine Wohlordnung ist (d. h. ein Φ dazu definieren). Selbstverständlich ⟦werden⟧ beide durch Induktion nach der schon als Ordinalzahl nachgewiesenen Anordnung α ⟦definiert⟧.
6. *Definition* geht: das sieht man am leichtesten, wenn man daran denkt, daß 2^α erhalten werden kann durch absteigende n-Tupel von Zahlen $< \alpha$, lexikographisch geordnet:

$$(\text{Beweis} : \beta < 2^\alpha \rightarrow \beta = 2^{\gamma_1} + \ldots + 2^{\gamma_k}, \quad \gamma_i < \alpha)$$

[aa]The base case for the definition of the fundamental sequence for ϵ_0 was omitted in the text, but given by Gödel in *Konzept*, p. 9.

From that then follows:

1. Inference by induction, that is, if E is a propositional function, and $E(a)$ is derivable in the system S_1 from the assumption $x \underset{R}{<} a \rightarrow E(x)$, then $(x)E(x)$ is provable in the system S_1;
2. Definition by induction: If $g_1(x), \ldots, g_v(x)$ are any functions such that $g_i(x) \underset{R}{<} x$, and for any expression ⟦A⟧, one writes down

$$\varphi(x) = A(\varphi(g_1(x)), \ldots, \varphi(g_v(x))),$$

then there is a φ ⟦that satisfies this equation⟧.

3. The same is also still possible for larger ordinals, for example ω^2, ω^ω, ω^{ω^2}. Therefore inferences by induction are provable from the other axioms. *ω^ω already no longer in finitary number theory (but all ⟦$< \omega^\omega$⟧ definable in finitary number theory).*

4. But there are certain numbers for which it is not possible to prove that, not even in *transfinite arithmetic.* One such number is the first ϵ-number—defined[aa] by ⟦$\alpha_1 = 2^{\omega+1}$ and⟧ $\alpha^{n+1} = 2^{\alpha_n}$. ⟦An⟧ *intuitive picture of this number.* First ⟦I consider the⟧ process 2^α:

$$2^1 = 2, \quad 2^\beta = \sum_{x<\beta} 2^x,$$

that is, to every element *is assigned the sum of the previous ones.* Very intuitive construction procedure.[bb]

5. Difficulty: *Why can that not be formalized in arithmetic,* also not if it is extended by finite finitary systems? Suppose it has already been proved that α is an ordinal number in the sharpest sense, that is, a $\Phi(f)$ is defined, so that ⟦every subsequence⟧ terminates. Next we will try to define in arithmetic an *ordering* of natural numbers which represents 2^α, and then to prove that this ordering is a well-ordering (that is, to define a Φ for it). Obviously both ⟦will be defined⟧ by induction according to the ordering α, already shown to be an ordinal.

6. The *definition* goes through: That is most easily seen if one considers the fact that 2^α can be obtained by decreasing n-tuples of numbers $< \alpha$, lexicographically ordered:

$$\text{(Proof: } \beta < 2^\alpha \rightarrow \beta = 2^{\gamma_1} + \ldots + 2^{\gamma_k}, \gamma_i < \alpha).$$

[bb] One might think from the end of *Konzept*, p. 9, that at this point Gödel left out the successor case. However, his following comment and the argument of 7 support the hypothesis that he has exactly what he intended. If we call the function defined by this recursion 2[exp]x, then for $1 < x < \omega$, 2[exp]$x = 2^{x-1}$; otherwise 2[exp]$x = 2^x$.

Nun kann man innerhalb der Arithmetik die n-Tupel abbilden auf die Zahlen (z. B. durch Primzahlen), und ⟦diese⟧ Anordnung ⟦ist⟧ auch definierbar.
7. *Beweis* geht nicht: denn wie ist *das gewöhnlich* ⟦*zu*⟧ *beweisen*? Man würde beweisen, 2^1 ist eine Ordinalzahl. Angenommen 2^γ $(\gamma < \beta)$ ist eine Ordinalzahl, dann ist ja $2^\beta = \sum_{\gamma<\beta} 2^\gamma$. ⟦Die⟧ Summe einer wohlgeordneten Folge von Ordinalzahlen ist eine Ordinalzahl. Die Eigenschaft, nach der man Induktion anwendet, ist also das "*Ordinalzahlsein*", d. h. aber, *jede Teilmenge* hat ein erstes Element, *oder jede* Teilfolge bricht ab. ⟦Diese Eigenschaft ist⟧ also imprädikativ, also nicht mehr formulierbar in der Zahlentheorie (*Existentiale Operationen für Klassen* nötig und sogar ⟦das⟧ Reduzibilitätsaxiom), und das ist *wesentlich, wie aus dem Gentzenschen Beweis hervorgeht.*
8. Vor Gentzen zeigte es sich nur in dem Fehlschlag der Versuche, z. B. ⟦des⟧ Versuchs, ein Φ zu definieren. $\Phi(f)$ führt auf eine Rekursion nach α, aber in den g_i kommt das Φ selbst wieder vor.
12 |9. Durch abzählbare Iteration dieses Übergangs von α zu 2^α entsteht dann ϵ_o. Diese ⟦Ordinalzahl wäre⟧ also ohne weiteres gegeben, falls dieser Übergang gegeben ⟦wäre⟧. Nichtsdestoweniger wird man dem Induktionsschluß nach dieser so definierten Zahl ϵ_o einen hohen Grad von Anschaulichkeit nicht absprechen, wie überhaupt dem Verfahren, eine *Ordinalzahl durch Induktion nach Ordinalzahlen zu definieren* (obwohl ⟦es⟧ ein imprädikatives Verfahren ⟦ist⟧).
10. Es macht, wie gesagt, keine Schwierigkeit, eine Anordnung vom Typus ϵ_o *innerhalb eines finiten Systems* (*sogar* S_1) *zu definieren.* Nur der Beweis der vollständigen Induktion ist unmöglich. Man kann also diese als neue Schlußregel hinzufügen, d. h. also folgendes: Sei R_o die Ordnung, und es sei gelungen aus $(x)[x \underset{R_o}{<} a \rightarrow E(x)]$ abzuleiten $E(a)$, dann darf man daraus $(x)E(x)$ schließen. Oder andere Systeme würde man erhalten, indem man das für andere Ordinalzahlen tut, die durch Induktion nach Ordinalzahlen definiert sind.
11. Ist dieses System finit? Es *genügt allen unseren Bedingungen* außer 3. Die hatte ja gelautet, es soll kein anderer Schluß als ⟦die⟧ gewöhnliche Induktion vorkommen. Dieser ⟦neue⟧ Schluß kann aber als eine Verallgemeinerung der gewöhnlichen Induktion aufgefaßt werden, und insofern ist die Abweichung von der Forderung 3 vielleicht keine so tiefgehende.
12. Ich möchte übrigens bemerken, daß Gentzen *einen "Beweis"* für diesen Schluß zu geben suchte, und ⟦er⟧ sag⟦te⟧ sogar, daß dies der wesentliche Teil seines Widerspruchsfreiheitsbeweises sei. In Wirklichkeit handelt es sich dabei aber gar nicht um einen Beweis, sondern um eine Berufung auf Evidenz; was ja auch klar ist. Ich möchte vorlesen, was Gentzen selbst über diesen Beweis sagte.[cc] Ich glaube es hat mehr Sinn, ein Axiom präzis zu

Now one can, within arithmetic, map the n-tuples onto the numbers (for example by primes), and this ordering is also definable.

7. The *proof* does not go through: for how is *this usually to be proved?* One would prove that 2^1 is an ordinal. Supposing that 2^γ is an ordinal ($\gamma < \beta$), then so is $2^\beta = \sum_{\gamma<\beta} 2^\gamma$. The sum of a well-ordered sequence of ordinals is an ordinal. The property to which induction is applied is therefore that of "*being an ordinal*", that is that *every subset* has a first element, *or every* subsequence breaks off. ⟦This property is⟧ therefore impredicative, thus no longer formulable in number theory (*existential operators for classes* are necessary, and even the axiom of reducibility), and that is *essential, as is evident from Gentzen's proof.*

8. Before Gentzen this came to light only in the failure of attempts, for example the attempt to define a Φ. $\Phi(f)$ leads to a recursion on α, but Φ itself occurs again in the g_i.

9. By countable iteration of this transition from α to 2^α, ϵ_0 is generated. This ⟦ordinal would be⟧ therefore given immediately, once this transition is given. Nonetheless, one will not deny a high degree of intuitiveness to the inference by induction on ϵ_0 thus defined, as in general to the procedure of *defining an ordinal by induction on ordinals* (even though this is an impredicative procedure).

10. As we have said, there is no difficulty in defining an ordering of type ϵ_0 *within a finitary system (even S_1).* Only the proof of complete induction is impossible. Therefore one can add this as a new rule of inference, that is the following: Let R_0 be the ordering, and suppose one has succeeded in deriving $E(a)$ from $(x)[x \underset{R_o}{<} a \rightarrow E(x)]$, then from this one may conclude $(x)E(x)$. Or other systems would be obtained by doing this for other ordinal numbers that are defined by induction on ordinals.

11. Is this system finitary? It *satisfies all our conditions* except 3. Recall that this had stated that no other inference should occur than ordinary induction. This ⟦new⟧ inference can be considered as a generalization of ordinary induction, and in this respect the deviation from the requirement 3 is perhaps not such a drastic one.

12. I would like to remark by the way that Gentzen sought to give *a "proof"* of this rule of inference and even said that this was the essential part of his consistency proof. In reality, it's not a matter of a proof at all, but of an appeal to evidence—what is after all also clear. Let me read what Gentzen himself says about this proof.[cc] I think it makes more sense to formulate

---

[cc] At this point Gödel evidently read a passage from Gentzen which he does not identify, almost certainly from *Gentzen 1936*. Gentzen's proof of transfinite induction is in §15.4 of the paper, but possibly Gödel read from the comments on the proof in §16.11.

formulieren und zu sagen, daß ⟦es⟧ eben nicht weiter reduzierbar ist. Aber
13 es tritt wohl hier wieder | das Bestreben der Hilbert Schüler, aus nichts etwas zu deduzieren, hervor.

13. *Wie weit kommt man damit?* Gentzen hat die Widerspruchsfreiheit der Zahlentheorie ⟦bewiesen⟧, wahrscheinlich auch des Weylschen *Kontinuums* ⟦*1918*⟧. Gentzen hoffte ferner, die Analysis und Teile der Mengenlehre und höher zu erhalten; bei genügend grossen Ordinalzahlen ist das wohl nicht zu bezweifeln. Ob man aber mit diesem Prinzip, ⟦d. h.⟧ Definition von *Ordinalzahlen durch* transfinite *Induktion nach bereits definierten Ordinalzahlen*, das ja immerhin einen hohen Grad von Anschaulichkeit besitzt, auskommt, ist wohl fraglich; hier vielleicht negativ. [Insbesondere $\epsilon_{[\![0]\!]}$ vielleicht in S_ω.]

14. Verhältnis zu den ersten Erweiterungen. Vielleicht ⟦geht⟧ dieses ganze Prinzip (es erscheint sehr anschaulich) nicht über die Systeme S_i hinaus.

15. Jetzt möchte ich noch einiges über den Gentzenschen Beweis sagen. Zunächst muß man sagen, daß seine Arbeit sehr kompliziert ⟦und⟧ durchaus nicht ein Muster an Klarheit ist. Es gibt ferner bis jetzt anscheinend nur sehr wenige Mathematiker, die sie kontrolliert haben. Es ist also die Möglichkeit eines Fehlers nicht auszuschließen. Ich glaube das eigentlich nicht, sondern ich habe den Eindruck, daß die Reduktion stimmt.

14 |16. Gentzen gibt an, er habe seinen Beweis in dem Bestreben, den modalitätslogischen zu präzisieren, gefunden. Die Verwandtschaft des fertigen Beweises ist aber eine ziemlich entfernte. Es ist von jeder beweisbaren Formel eine $\neg\neg$ bewiesen, aber in einem anderen Sinn als dem modalitätslogischen. Nehmen wir eine zahlentheoretische Formel in der Normalform

$$\Phi = (x)(Ey)(z)(Eu)A(x, y, z, u),$$

wobei A ein elementarer (entscheidbarer) Ausdruck ist.—Was wäre ein konstruktiver Beweis einer solchen Formel? ⟦Für gewisse gegebene⟧ $f(x)$ ⟦und⟧ $g(x, z)$ ⟦ein konstruktiver Beweis von⟧ $(x, z)A(x, f(x), z, g(x, z))$. Wie sieht die Negation der obigen Formel aus, wenn man sie auf die Normalform bringt? $(Ex)(y)(Ez)(u){\sim}A(x, y, z, u)$. Ein konstruktiver Beweis der Negation wäre als⟦o⟧ die Angabe einer Zahl c und einer Funktion $f(y)$, so daß für alle $(y, u){\sim}A(c, y, f(y), u)$. ⟦Ein Beweis von⟧ $\neg\neg\Phi$ wäre der Beweis, daß es kein solches f, c geben kann: man kann zu jedem f, c; y, u angeben, ⟦so daß⟧ $(f, c)A(c, \underline{y_{f,c}}, f(y_{f,c}), \underline{u_{f,c}})$. Solche Funktionsfunktionen y, u nennt er eine Reduktion.[1]

[1]Der uferlose Begriff des "Beweis" wird also hier ersetzt durch den ebenso uferlosen Begriff der Funktionsfunktion.—Das geht, wo das Ersetzen durch bestimmten Beweis⟦begriff⟧ eben nicht geht.

an axiom precisely and to say that it is just not further reducible. But here again the drive of Hilbert's pupils to derive something from nothing stands out.

13. *How far does one get with this?* Gentzen proved the consistency of number theory ⟦and⟧ probably also of Weyl's *Kontinuum* ⟦*1918*⟧. Gentzen hoped further to obtain analysis and parts of set theory and higher; with sufficiently large ordinals, that is no doubt true. Whether one gets by with this principle, that is *definition of ordinals by* transfinite *induction on previously defined ordinals*, which admittedly still has a high degree of intuitiveness, seems questionable; here perhaps negative. [In particular $\epsilon_{[\![0]\!]}$ perhaps in S_ω.]

14. Relation to the first extensions. Perhaps this whole principle (it seems very intuitive) doesn't go beyond the systems S_i.

15. Now I would like to say something about Gentzen's proof. First one must say that his paper is very complicated and not at all a model of clarity. There are, furthermore, so far apparently only very few mathematicians who have checked it. Thus the possibility of a mistake ⟦in the proof⟧ can't be ruled out. I don't really believe that; rather I have the impression that the reduction is correct.

16. Gentzen claims he found his proof in the endeavor to make the modal-logical one more precise. The affinity of the finished proof is, however, rather remote. A $\neg\neg$ of every provable formula is proved, but in a different sense from the modal-logical. Consider a number-theoretic formula in the normal form

$$\Phi = (x)(Ey)(z)(Eu)A(x, y, z, u)$$

where A is an elementary (decidable) expression.—What would a constructive proof of such a formula be? ⟦For certain given⟧ $f(x)$ and $g(x, z)$, ⟦a constructive proof of⟧ $(x, z)A(x, f(x), z, g(x, z))$. How does the negation of the above formula look, if one brings it into the normal form $(Ex)(y)(Ez)(u){\sim}A(x, y, z, u)$? Thus a constructive proof of the negation would be the giving of a number c and a function $f(y)$, so that for all $(y, u){\sim}A(c, y, f(y), u)$. ⟦A proof of⟧ $\neg\neg\Phi$ would be the proof that there can be no such f, c: one can for every f, c give y, u so that $(f, c)A(c, \underline{y_{f,c}}, f(y_{f,c}), \underline{u_{f,c}})$. He calls such functionals y, u a reduction.[1]

[1]The vast notion of "proof" is thus here replaced by the equally vast notion of functional.—That works, where the replacement by a definite ⟦notion of⟧ proof does not.

17. Es wird gezeigt, wie man aus dem Beweis für eine Formel eine Reduktion finden kann. Zunächst ist es ganz leicht, für die Axiome Reduktionen anzugeben. Auch für den Satz vom ausgeschlossenen Dritten ⟦ist es⟧ ganz leicht:[dd]

$$(Ex)(y)[\varphi(x) \vee {\sim}\varphi(y)] \qquad (f)[\varphi(\overbrace{\psi(f)}^{x}) \vee {\sim}\varphi(f(\overbrace{\psi(f)}^{x}))] \qquad [\![*]\!]$$

$$\begin{array}{ll} {\sim}\varphi(x).\varphi(f(x)) & \overbrace{\psi(f)}^{x} = f(0) \mid \varphi(f(0)) \\ {\sim}\varphi(0).\underline{\varphi(f}(0)) & \qquad [\![=]\!] \;\; 0 \mid {\sim}\varphi(f(0)) \\ {\sim}\varphi\underline{(f(0))}.\varphi(ff(0)) & \end{array}$$

|18. Die Methode, wie dieses ψ definiert wird, ist die Probiermethode. Man setzt zuerst $\psi(f) = 0$. Geht es damit nicht, so muß $\varphi(f(0))$ gelten. Dann geht es mit $\psi(f) = f(0)$.

Schlußregeln. Die ganze Schwierigkeit des Beweises liegt in der Schlußregel $P \quad P \supset Q$ ⟦$/Q$⟧. Hat man die, so ist insbesondere der Induktionsschluß klar. ⟦Seien die Formeln⟧

$$F(0) \qquad\qquad F(n) \supset F(n+1)$$

bewiesen. Daher hat man für jedes $F(n)$ ein Reduktionsverfahren, ⟦d. h. man hat⟧ jedem n ein Reduktionsverfahren ⟦für⟧ $F(n)$ zugeordnet; ⟦das⟧ heißt, man hat ein Reduktionsverfahren für $(n)F(n)$.

19. Der Beweis für $P \supset Q$ geht folgendermaßen: den Funktionsfunktionen, welche finit definiert sind (d. h. für jedes konkret vorgelegte f berechenbar), kann man ja Ordinalzahlen der zweiten Klasse zuordnen (Souslinsches Schema). Die reduzierende Funktion für Q wird definiert durch transfinite Induktion nach der Ordinalzahl der reduzierenden Funktion für P, und wenn man sich die Ordinalzahl, welche der reduzierenden Funktion für Q zugeordnet ist, ausrechnet, so tritt die von P im Exponenten auf. Es ist also genau der Schluß, daß man eine gewisse neue Ordinalzahl einführt durch Rekursion nach der bereits als Ordinalzahl erkannten und dann nach dieser neuen wieder rekursive Definition anwendet.

[dd]That is, we set x (i.e, $\psi(f)$) $= f(0)$ if $\varphi(f(0))$ holds, $= 0$ otherwise. Thus if $\varphi(f(0))$ holds, the first disjunct of ⟦*⟧ holds, otherwise the second disjunct holds. That is, one

17. It is shown how one can find a reduction for a formula from the proof. First, it is quite easy to give reductions for the axioms. Even for the law of the excluded middle it is very easy:[dd]

$$(Ex)(y)[\varphi(x) \vee \sim\varphi(y)] \quad (f)[\varphi(\overbrace{\psi(f)}^{x}) \vee \sim\varphi(f(\overbrace{\psi(f)}^{x}))] \qquad [\![*]\!]$$

$$\begin{array}{ll} \sim\varphi(x).\varphi(f(x)) & \overbrace{\psi(f)}^{x} = f(0) \mid \varphi(f(0)) \\ \sim\varphi(0).\underline{\varphi(f}(0)) & \quad [\![=]\!] \;\; 0 \mid \sim\varphi(f(0)) \\ \sim\varphi\underline{(f(0))}.\varphi(ff(0)) & \end{array}$$

18. The method by which this ψ is defined is the method of trial and error. First one sets $\psi(f) = 0$. If that doesn't work, then $\psi(f(0))$ must be true. Then $\psi(f) = f(0)$ works.

Rules of inference. The whole difficulty of the proof lies in the rule of inference $P\ P \supset Q[\![/Q]\!]$. If we have that then in particular the induction rule is clear. [[Suppose that the formulae]]

$$F(0) \qquad F(n) \supset F(n+1)$$

have been proved. Then we have a reduction procedure for every n, that is, we have correlated to each n a reduction procedure for $F(n)$; that means that we have a reduction procedure for $(n)F(n)$.

19. The proof for $P \supset Q$ goes as follows: We can assign ordinals of the second number class to the functionals that are defined in a finitary way (that is, computable for every concretely presented f) (Souslin's schema). The reducing function for Q is defined by transfinite induction on the ordinal of the reducing function for P, and if we compute the ordinal that is assigned to the reducing function for Q, then that for P occurs in the exponent. It is therefore exactly the inference of introducing a certain new ordinal by recursion on an ordinal already recognized as such and then again applying recursive definition on this new [[ordinal]].

---

of the two last formulas on the left must be false, thus ruling out f as a Skolem function for the negated prenex of the orginal formula. Essentially the same argument occurs in *Kreisel 1951*, p. 257. Cf. the discussion in section 7 of the introductory note.

16 |
VII

Zum Schluß möchte ich auf die historische und erkenntnistheoretische Seite der Frage zurückkommen und also fragen, (1) ob einem Widerspruchsfreiheitsbeweis mittels der drei erweiterten Systeme ein Wert im Sinn einer Tieferlegung der Fundamente zukommt; (2) was damit zusammenhängt, ob das Hilbertsche Programm dadurch, daß über die finite Zahlentheorie notwendig hinausgegangen wird, in einem wesentlichen Punkt vereitelt ist.

Dazu kann man zweierlei sagen: (1) Falls das ursprüngliche Hilbertsche Programm durchführbar gewesen wäre, so wäre das zweifellos von ungeheurem erkenntnistheoretischem Wert gewesen. Es wären nämlich beide Forderungen erfüllt worden: (A) Die Mathematik wäre auf einen sehr kleinen Teil von sich reduziert worden (also eine große Anzahl von unabhängigen Annahmen ⟦wären⟧ überflüssig geworden). (B) Es wäre wirklich alles auf eine konkrete Basis reduziert worden, auf die alle sich müssen einigen können.[ee] ⟦(2)⟧ Bei den Beweisen mittels des erweiterten Finitismus ist das erste gar nicht mehr der Fall, denn man muß ja immer, um etwas als widerspruchsfrei ⟦zu⟧ beweisen, gewisse andere Annahmen an die Stelle der als widerspruchsfrei ⟦zu beweisenden⟧ setzen, so daß man keine Reduktion ⟦im obigen Sinne⟧[ff], sondern eine Ersetzung oder Verschiebung hat. Das zweite (Reduktion auf die konkrete Basis, d. h. also Erhöhung des Evidenzgrades) ⟦ist⟧ bei den verschiedenen Systemen in verschiedenen Graden der Fall, also z. B. bei den modalitätslogischen gar nicht, bei den höheren Funktionstypen am meisten, bei den transfiniten Ordinalzahlen—soweit man nur
17 das oben ausgesprochene Prinzip anwendet—auch | noch in einem ziemlich hohen Grad. Man wird in diesem Sinn dem Gentzenschen Beweis nicht absprechen können, daß er das Operieren mit dem transfiniten E auf etwas evidenteres (die erste ϵ-Zahl) zurückführt. Auf jeden Fall scheint mir, daß die erkenntnistheoretische Bedeutung, im Sinn einer besseren Fundierung, dadurch daß sie ⟦die verschiedenen Systeme⟧ nicht in der finiten Zahlentheorie enthalten sind, sehr vermindert wird. ⟦Davon⟧ ganz unbeschadet ⟦ist⟧ die mathematische Bedeutung dieser Untersuchung. Diese scheint hier tatsächlich außerordentlich groß zu sein, und ich bin überzeugt, daß die dabei verwendeten Methoden in der Grundlagenforschung und auch außerhalb ihrer zu sehr interessanten Resultaten führen werden.

[ee]B is a paraphrase of *Hilbert 1926*, p. 180, ll. 9–10; see *van Heijenoort 1967*, p. 384 ll. 6–7.

VII

In conclusion, I would like to return to the historical and epistemological side of the question and then ask (1) whether a consistency proof by means of the three extended systems has a value in the sense of laying the foundations more securely; (2) what is closely related, whether the Hilbert program is undermined in an essential respect by the fact that it is necessary to go beyond finitary number theory.

To this we can say two things: (1) If the original Hilbert program could have been carried out, that would have been without any doubt of enormous epistemological value. The following requirements would both have been satisfied: (A) Mathematics would have been reduced to a very small part of itself (therefore a large number of independent assumptions would have become superfluous). (B) Everything would really have been reduced to a concrete basis, on which everyone must be able to agree.[ee] (2) As to the proofs by means of the extended finitism, the first is no longer the case at all, since in order to prove something consistent, one must always put other assumptions in the place of those ⟦to be proved⟧ consistent, so that one doesn't have a reduction ⟦in the above sense⟧,[ff] but rather a replacement or shifting. The second (reduction to the concrete basis, which means increase of the degree of evidence) obtains for the different systems to different degrees, thus for example for the modal-logical not at all, for the higher function types the most, for the transfinite ordinal numbers—insofar as one applies only the principle stated above—also to a rather high degree. In this sense, one will not be able to deny of Gentzen's proof that it reduces operating with the transfinite E to something more evident (the first ϵ-number). In any case, it seems to me that the epistemological significance, in the sense of a better foundation, is very much diminished by the fact that ⟦the different systems⟧ are not contained in finitary number theory. The mathematical significance of this investigation is totally unaffected. The latter seems to me in fact to be extraordinarily great, and I am convinced that the methods applied here will lead to very interesting results in foundational research and also outside it.

[ff]Without the addition there is (the possibility of seeing) a conflict with the subsequent remarks and also the discussion in section 1.